Leveraging Distribution Quantiles to Detect Gene Interactions in the Pursuit of Personalized Medicine

LEVERAGING DISTRIBUTION QUANTILES TO DETECT GENE INTERACTIONS IN THE PURSUIT OF PERSONALIZED MEDICINE

By Akram Alyass, M.Sc

A Thesis submitted to
the School of Graduate Studies in
Partial Fulfilment of the Requirements
for the Degree Doctor of Philosophy

McMaster University

Doctor of Philosophy (2018) McMaster University
School of Computational Science and Engineering Hamilton, Ontario, Canada

TITLE:	Leveraging Distribution Quantiles to Detect Gene Interactions in the Pursuit of Personalized Medicine
AUTHOR:	Akram Alyass, MSc
SUPERVISOR:	Dr. David Meyre
NUMBER OF PAGES:	xviii, 294

Dedicated to my mother, Fouzieya Al-shabik, and my father, Mohamed Alyass

Abstract

Anticipations of personalized medicine are primarily attributed to the recent advances in computational science and high-throughput technologies that enable the ever-more realistic modeling of complex diseases. These diseases result from the interplay between genes and environment that have limited our ability to predict, prevent, or treat them. While many envision the utility of integrated high-dimensional patient-specific information, basic research towards developing accurate and reliable frameworks for personalized medicine is relatively slow in progress. This thesis provides a state-of-the-art review of current challenges towards personalized medicine. There is a need for global investment in basic research that includes 1) cost-effective generation of high-quality high-throughput data, 2) hybrid education and multidisciplinary teams, 3) data storage and processing, 4) data integration and interpretation, and 5) individual and global economic relevance; to be followed by global investments into public health to adopt routine personalized medicine. This review also highlights that unknown or unadjusted interactions result in true heterogeneity in the effect and relevance of patient data. This limits our ability to integrate and reliably utilize high-dimensional patient-specific data. This thesis further investigates the true heterogeneity in marginal effects of known BMI genetic variants. This involved the development of the novel statistical method, meta-quantile regression (MQR), to identify variants with potential gene-gene / gene-environment interactions. Applying MQR on public and local data (75,230 European adults) showed that *FTO*, *PCSK1*, *TCF7L2*, *MC4R*, *FANCL*, *GIPR*, *MAP2K5*, and *NT5C2* have potential interactions on BMI. In addition, a gene score of 37 BMI variants shows that the genetic architecture of BMI is shaped by gene-gene and gene-environment interactions. The computational cost of fitting MQR models was greatly reduced using unconditional quantile regression. The utility of MQR was further compared to variance heterogeneity tests in identifying variants with potential interactions. MQR tests were found to have a higher power of detecting synergetic and antagonistic interactions for skewed quantitative traits while maintaining nominal Type I error rates compared to variance heterogeneity tests. Overall, MQR is a valuable tool to detect potential interactions without imposing assumptions on the nature of interactions.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor, Dr. David Meyre, for his unwavering support, constructive feedback, and warm encouragement over the years. Without his guidance, consistent help, and engaging teaching style, this thesis would not have been possible. I am grateful for him.

I would also like to extend my appreciation to my Advisory Committee, Dr Ben Bolker, and Dr. Russell de Souza, for their constructive feedback and helpful support throughout this project. I am very thankful to have worked with such knowledgeable and insightful faculty throughout this project.

I would also like to thank my colleagues, Arkan Abadi, Amel Lamri, Aihua Li, and Sébastien Robiou du Pont, for their constant support and encouragements for this research project and stimulating discussions.

Lastly, I would like to thank my parents, my wife, Mada Mohammed, and all my close family and friends. I would like to apologize for all the time this 'thesis' has taken to complete, and I am grateful for all the support I've received.

Declaration of Academic Achievement

All chapters are separate manuscripts except for the introduction and conclusion of this thesis. Chapter 2 and 3 have been published, while Chapter 4 is prepared for submission and publication post reviews from co-authors.

All scripting, analytical development, statistical analysis, manuscript write-up of each of the chapters were primarily completed as an individual effort. Contributions to the data preparation, analysis, and edits for Chapter 3 and 4 have been made by Arkan Abadi, while David Meyre has provided contributions to edits in Chapter 2, 3, and 4. Ben Bolker has provided edits and critical reviews to Chapter 3 and 4.

Notation and Abbreviations

BF	- Brown-Forsythe
BGI	- Beijing Genomics Institute
BMI	- Body Mass Index
CI	- Confidence Interval
CPU	- Centeral Processing Units
CQR	- Conditional Quantile Regression
dbGaP	- Database of Genotypes and Phenotypes
DNA	- Deoxyribonucleic acid
EA	- European Ancestry
GS	- Gene Score
GPU	- Graphic Processing Units
GWAS	- Genome Wide Association Study
GWIS	- Genome Wide Interaction Study
HIV	- Human Immunodeficiency Virus
HWE	- Hardy Weinberg equilibrium
IDB	- Identity by descent
MAF	- Minor Allele Frequency
MCQR	- Meta Conditional Quantile Regression
MODMatcher	- Multi-Omics Data Matcher
MONA	- multi-level ontology analysis
MQR	- Meta-Quantile Regression
MR	- Meta-Regression
MUQR	- Meta Uconditional Quantile Regression
NW	- Normal Weight
OLS	- Ordinary least-squares
OW	- Over Weight
QC	- Quality Control
QR	- Quantile Regression
RNA	- Ribonucleic acid
SN	- Skew-Normal
SNP	- Single nucleotide polymorphism
T2D	- Type 2 Diabetes
UQR	- Unconditional Quantile Regression

Contents

List of Tables

List of Figures

Chapter 1

Introduction and Problem Statement

Multifactorial complex diseases such as obesity, clinical depression, anxiety, type 2 diabetes, hypertension, dyslipidemia and cancer have all turn into global epidemics in recent years [1, 2, 3, 4, 5, 6, 7]. They all include genetic, environment and genetic-environment interaction components that all contribute to the incident and prognosis of disease cases [8]. The recent advances in our ability to measure and record large multidimensional amount of individual data (e.g. demographic, clinical, environmental, and genetic profiles) have enabled us to envision the emergence of personalized medicine. That is routine evidence-based medicine via the integration of ones unique genetics and environmental exposures to proactively optimize our well-being [9]. Hence, the primary progress steps towards personalized medicine include the integration of heterogeneous data in the from of "bench to bed" and "bed to bench" frameworks.

The advent of personalized medicine depends in large part on the availability of accurate and reliable predictive models that incorporate the influence of relevant genetic, environmental factors, and their corresponding interactions. While much progress has been made identifying the genetic components of complex traits, the amount of phenotypic variance (i.e. heritability) explained by genome-wide significant associations remains minor [10]. On

the other hand, heritability estimates of complex traits could be inflated by genetic interactions (i.e. single or multiple gene x gene / gene x environment interactions) that require large sample sizes to detect. Genetic data embody a large mixture of modest signals and random errors, where our current ability to identify genetic interactions comes at the cost of tolerating larger error thresholds. Both unknown and or unadjusted genetic interactions are problematic as they result in true heterogeneity in the marginal effect of variants on complex diseases. Increasing the sample size only bypasses random error, but true heterogeneity between samples can only be bypassed via standardizations and calibrations that limit generalizability. It is nontrivial to distinguish random from systematic differences in effect of variants on complex traits between well-phenotyped and high-quality sample studies. Our inability to reliably detect genetic interactions limits the predictive value of models for complex diseases. There is an urgent need for robust statistical methods to reliability detect genetic interactions under heterogeneous systems of gene-gene, gene-environment interactions.

This thesis includes two peer-reviewed published articles and one unpublished paper. The second chapter of this thesis is the first paper published in BMC Medical Genomics. It is a state-of-the-art review on current challenges and opportunities for personalized medicine given that it is a broad and rapidly advancing research field. There is a large diversity in recent advances towards personalized medicine that make it difficult to follow and assess the status-quo of current challenges and recent advances. The primary notable challenges include overcoming the growing gaps in 1) socioeconomics and scientific progress between developed and undeveloped nations that threaten the social pillars of stability, and 2) our ability to generate compared to analyzing interoperating that is currently stagnating information from omic data (i.e. genomic, methylomic, metabolomic,..., etc). These challenges are discussed in Chapter 2 in more details.

The third chapter includes the paper published in the American Journal of Human Genetics. It investigates the heterogeneity in the impact of genetic variants across the sample

distribution on BMI as a complex trait. It proposes the use of a novel framework based on meta-regression and quantile regression to demonstrate the BMI includes a large genetic-environment interactions, and show that four BMI variants show potential interactions in which no-interactions are ruled out.

The fourth chapter contains unpublished work that is to be submitted to Plos Genetics or Genetic Epidemiology. The paper provides the novel method with a name, Meta-Quantile Regression (MQR), and it expands on its utility by substantially reducing the computational cost. It further provides a simulate study to compare the power of MQR to detect potential interactions compared with variance-heterogeneity tests. MQR is shown to have higher power than variance heterogeneity tests for asymmetric distributions and antagonistic interactions while maintaining nominal Type I error rates

Chapter 2

From big data analysis to personalized medicine for all: challenges and opportunities

Abstract

Recent advances in high-throughput technologies have led to the emergence of systems biology as a holistic science to achieve more precise modeling of complex diseases. Many predict the emergence of personalized medicine in the near future. We are, however, moving from two-tiered health systems to a two-tiered personalized medicine. Omics facilities are restricted to affluent regions, and personalized medicine is likely to widen the growing gap in health systems between high and low-income countries. This is mirrored by an increasing lag between our ability to generate and analyze big data. Several bottlenecks slow-down the transition from conventional to personalized medicine: generation of cost-effective high-throughput data; hybrid education and multidisciplinary teams; data storage and processing; data integration and interpretation; and individual and global economic relevance. This review provides an update of important developments in the analysis of big data and forward

strategies to accelerate the global transition to personalized medicine.

Introduction

Access to large omics (genomics, transcriptomics, proteo-mics, epigenomic, metagenomics, metabolomics, nutrio-mics, etc.) data has revolutionized biology and has led to the emergence of systems biology for a better understand-ing of biological mechanisms. Systems biology aims to model complex biological interactions by integrating in-formation from interdisciplinary fields in a holistic man-ner (holism instead of the more traditional reductionism). In contrast to treating a mixture of factors as single entities leading to an endpoint, systems biology relies on experimental and computational approaches in order to provide mechanistic insights to an endpoint [9]. Trad-itional observational epidemiology or biology alone are not sufficient to fully elucidate multifaceted heterogeneous disorders and this directly limits all prevention and treat-ment pursuits for such diseases [11, 12]. It is widely recog-nized that multiple dimensions must be considered simultaneously to gain understanding of biological sys-tems [13]. Systems approaches are driving the leading-edge of biology and medicine [14, 15]. The use of deterministic networks for normal and abnormal phenotypes are thought to allow for the proactive maintenance of wellness specific to the individual, that is predictive, preventive, personalized, and participatory medicine (P4, or more generally speaking, personalized medicine) [9].

Many predict the emergence of personalized medicine in the near future, but it is not likely to come about as quickly as the scientific community and the media may think [16]. In parallel to an escalating two-tiered health system at the global level, a similar two-tiered phenomenon is observed with regard to our ability to generate and analyze omics data that may delay even further the transi-tion to personalized medicine. The generation and manage-ment (storage, and computational resources) of omics data remain expensive despite technological progress. This im-plies that personalized medicine could be restricted to the

wealthier countries [17]. This is mirrored by a growing gap in our abilities to generate and interpret omics data. The bottleneck in omics approaches is becoming less and less about data generation and more and more about data man-agement, integration, analysis, and interpretation [18]. There is an urgent need to bridge the gap between advances in high-throughput technologies and our ability to manage, integrate, analyze, and interpret omics data [19, 20, 21]. This review addresses the growing gaps in socioeconomic and scientific progress toward personalized medicine.

Review

The rich get richer and the poor get poorer

The developing world is home to 84 % of the worlds population, yet accounts for only 12 % of the global spending on health [22]. There is a large disparity between the distribution of people and global health expenditures across geographical regions (Figure 3.1). While public financing of health from domestic sources has increased globally by 100 % from 1995 to 2006, a majority of low and middle-income countries experienced a reduction of funding during the same time [23]. Several life-threating but easily preventable or treatable diseases are still prevalent in developing countries (e.g. malaria). Personalized medicine will further increase these dispar- ities and many low and middle-income countries may miss the train of personalized medicine [24, 25, 26], unless the international community devotes important efforts towards strengthening health systems of the most disadvantaged nations.

Systems medicine, the application of systems biology to human diseases [27], requires investments in infrastructures with cutting-edge omics facilities and analytical tools, advanced digital technologies (high computing performance and storage resources), and highly-qualified multi-disciplinary teams (clinicians, epidemiologists, biologists, computer scientists, statisticians and mathematicians) in addition to investments in security and privacy. On the bright side, technology is evolving quickly and new developments are producing data more

efficiently. A few examples include the development of high-throughput next generation sequencing and microarrays in genomics and transcriptomics, mass spectrometry-based flow cytometer in proteomics, real-time medical imaging, and more recently, lab-on-a-chip technologies [28]. Some predict that a technological plateau may be reached for different reasons (reliability, cost-effectiveness), but these projections are not validated by historical trends in science as novel technological developments can always occur [29]. However, there is a consensus that most of the cost in omics studies will come from data analysis rather than data generation [18].

The economic value of omics networks as personalized tests for future disease onset or response to specific treatments / interventions remains largely unknown. A recent study by Philips et al. reflects this issue and highlights a lag between clinical and economical value assessment of personalized medical tests in current research [30]. Very few studies have incorporated an economic aspect in the evaluation of personalized tests. These tests range from those available in clinical use or in advanced stage of development, genetic tests with Food and Drug Administration labels, tests with demonstrated clinical utility, and tests examining conditions with high mortality or high health-associated expenditures. Economic evaluations of personalized tests are needed to guide investments and policy decisions. They are an important pre-requisite to hasten the transition to personalized medicine. In addition, those few personalized tests that included economic information were found to be relatively cost-effective, but only a minority of them were cost-saving, suggesting that better health is not necessarily associated with lower expenditures [30]. In summary, the costs associated with personalized medicine transition remain unclear, but personalized medicine may further widen the economic inequality in health systems between high and low-income countries. This jeopardizes social and political pillars of stability, and highlights the need for a broader translation-oriented focus across the globe [31].

Several ideas for stimulating sustainable innovations in developing nations include micro-grants as proposed by Ozdemir V. et al. [32]. Although $1,000$ micro-grants are relatively

small, they far exceed the annual income of individuals below the poverty line of 1.25/day as de-fined by the World Bank. Recipients of these grants may go a long way in connecting and co-producing know-ledge based innovations to broaden translational efforts. Type 1 micro-grants which are awarded through funding agencies may support small labs and local scholars to connect personalized medicine with new models of dis-covery and translation [32]. Type 2 micro-grants funded by science observatories and/or citizens through crowd-funding mechanisms may facilitate developments of glo-bal health diplomacy to share novel innovations (i.e. therapeutics, diagnostics) in areas with similar burdens [32]. There is an overall need to support local scholars in promoting knowledge and innovation within low and middle-income countries [33]. This includes for ex-ample, the case of advocating for treatment of persons with Human Immunodeficiency Virus (HIV) infections where their peers may not recognize their illness as an endemic that affects society [33]]. One successful ex-ample of personalized medicine for HIV patients in low and middle-income countries include personal text mes-sages for improving adherence to antiretroviral therapy in Kenya and Cameroon [34].

Interdisciplinary programs for global translational science such as the Science Peace Corps are another promising catalyzing agent for research and developments in low and middle-income countries (http://www.peace-corps.gov/) [31]. The present Peace Corps program entails volunteer work (6 weeks minimum and up to 2 years) in various regions across the globe to serve as a steady flux of knowledge for translational research. Junior or senior scientists may cover topics from life sciences, medicine, surgery, and psychiatry. This program is bi-directional as it serves both the rich and poor to elucidate the concept of health and integrate personal-ized medicine within various environments. Lagging developments in low and middle-income countries are in fact open opportunities with rewards for intellectual individuals given the simple fact that it is where the majority of the human populations reside.

The tragedy of the commons is a conceptual economic problem where the benefits of

common and open re-sources are jeopardized by individuals self-interest to optimize personal gains [26]. The 2009 Economics Nobel Laureate, Elinor Ostrom, has shown that this issue is not actually common among humans since individuals work through establishing trust, and tend to find solutions to common problems themselves [35]. Societies do systematically develop complex sustainable regulations to collectively benefit each other where assurance is a critical factor for cooperation [36]. There is a need to understand institutional diversity if humans are to act collectively to benefit each other. Diverse applications of personalized medicine can be envisioned to cope with the diversity of the world by allowing multi-tier personalized health care systems at multiple scales and avoiding a single top-tier health care system that may instead compromise resource management. This also brings about the need for nested regulation systems for both science and ethics (i.e. ethics-of-ethics) as the assurance factor for cooperation [37, 38]. Transparency and accountability need to be imposed on all scientists, practitioners, ethicists, sociologists, and policymakers. No one should be above the fray for account-ability if a sustainable transition towards personalized medicine is to occur.

Omics data: the shifting bottlenecks

In parallel to the gap in health systems between rich and poor countries that personalized medicine may widen, an increasing lag has been observed in our ability to generate versus integrate and interpret omics data these last ten years [18]. New technologies and knowledge emerging from the Human Genome Project, fueled by biotechnology companies, led to the omics revolution in the beginning of the 21th century [39]. Using high-throughput technologies, we are now able to perform an exhaustive number of measurements over a short period of time giving access to individuals DNA (genomics), transcribed RNA from genes over time (transcriptomics), DNA methylation and protein profiles of specific tissues and cells (epigenomics and proteomics), metabolites (metabolomics), among other types of omics data [40]. Even histopathological and radiological images which are traditionally evaluated and

scored by trained experts are now subjected to computational quantifications (i.e. imaging informatics) [19, 20, 21, 41]. Business models based on returns on investments have driven ongoing technological developments to accelerate the generation of omics data at increased affordability in comparison with existing technologies. As a consequence, omics platforms and individual omics profiles are expected to be-come fairly affordable and data generation is no more a bottleneck for most laboratories, at least in the middle and high-income countries [42].

Initially, there were great expectations for omics data to provide clues on the mechanisms underlying disease initiation and progression as well as new strategies for disease prediction, prevention and treatment [9]. The idea was to translate omics profiles into subject-specific care based on their disease networks (Figure 3.2). However, our ability to decipher molecular mechanisms that regulate complex relationships remains limited despite growing access to omics profiles. Biological processes are very complex, and this coupled with the noisy nature of experimental data (e.g. cellular heterogeneity) and the limitations of statistical analyses (e.g. false positive associations) poses many challenges to detecting interactions between net-works and networks of networks. As an illustration, only a minority of the genetic variants predisposing to type 2 diabetes have been identified so far, despite large-scale studies involving up to 150,000 subjects [9, 43]. It becomes more and more obvious that the bottleneck in laboratories has shifted from data generation to data management and interpretation [44].

Personalized medicine needs hybrid education

Although solutions for the challenges of big data already exist and are adopted by companies such as Google, Apple, Amazon, and Facebook to tackle the fairly homogenous big data (i.e. user data) [45], the heterogeneous nature of omics data presents a new challenge that requires sufficient understanding of the underlying biological concepts and analysis algorithms to carry out data integration and interpretation [46]. It is important for the working scientist

to understand 1) the underlying problem, 2) the methods of data analysis, and 3) the advantages, and disadvantages of different computational platforms to carry out explorations and draw inference. Expertise in biology provides a foundation to contextualize causal effects and guide identification and interpretation of interaction signals from noise. There is also no uniformly most powerful method to analyze omics data and the use of various approaches to infer biological interactions requires modeling expertise [47]. Otherwise, research quality is sacrificed to avoid the logistical challenges of modeling in exchange for the use of more straightforward approaches [48]. Lastly, computer programing skills are necessary to navigate explorations and analyze omics data accordingly. There is a need for reliable and maintainable computer codes through best prac-tices for scientific computing [49]. Approximately 90% of scientists are self-taught in developing software and one may lack basic practices such as task automation, code review, unit testing, version control and issue tracking [50, 51]. Barriers between disciplines still exist between informaticians, mathematicians, statisticians, biologists, and clinicians due to a too divergent scientific background. Cutting-edge science is integrative by essence and innovative strategies in universities to educate and train future researchers at the interface of traditionally partitioned disciplines is urgently needed for the transition to personalized medicine. Johns Hopkins University is leading this evolution by changing the teaching plans and establishing new programs in the school of medicine that integrate the notion of personalized medicine [52]. Although increased know-ledge at the population level is a key factor in development of modern societies, there is an upper limit to the wealth of knowledge and expertise a single individual can hold [53]. This is the reason why, in addition to multidisciplinary individual training, initiatives by universities, research funding agencies, and governments are encouraged to connect researchers from diverse scientific backgrounds on interface topics related to systems biology and personalized medicine. The recent shift by the Canadian Institutes of Health Research from distinct discipline (e.g. genetics) to multidisciplinary ex-pert panels in funding biomedical research is a step in the right direction. The creation of interdisciplinary research institutes, such as

the Steno Diabetes Center in Denmark that combine clinical, educational and multifaceted re-search activities to lead translational research in diabetes care and prevention, is another sensible initiative that could prefigure what may become personalized medicine institutes in the future.

Management and processing of omics data

Major investments need to be made in bioinformatics, biomathematics, and biostatistics by the scientific community to accelerate the transition to personalized medicine. Classic research laboratories do not possess sufficient storage and computational resources for processing omics data. Laboratory-hosted servers require investments in informatics support for configuring and using software. Such servers are not only expensive to setup and maintain, but do not meet the dynamic requirements of different workflows for processing omics data, leading to either extravagant or sub-optimal servers. One promising technology to close the gap between generation and handling of omics data is cloud computing [54, 55]. It is an adaptive storage and computing service that exploits the full potential of multiple computers together as a virtual resource via the Internet [56]. Examples include the EasyGenomics cloud in Beijing Genomics Institute (BGI), and Embassy clouds as part of ELIXIR project in collaboration with multiple European countries (UK, Sweden, Switzerland, Czech Republic, Estonia, Norway, the Netherlands, and Denmark) [57]. The focus is currently placed on developing cloud based toolkits and workflow platforms for high-throughput processing and analysis of omics data [57, 58, 59, 60]. More recently, Graphics Processing Units (GPUs) have been proposed for general-purpose computing in a cloud environment [61]. GPUs provide faster computations as accelerators by one or two orders of magnitudes compared to general Central Processing Units (CPUs) and have been exploited to cope with exponentially growing data [62, 63, 64]. MUMmerGPU for example, processes queries in parallel on a graphics card, achieves more than a 10-fold speedup over a CPU version of the sequence alignment kernel, and outperforms the CPU version of MUMmer by 3.5-fold in total

application time when aligning reads [65]. However, a significant amount of work will be required for developing parallelization algorithms considering the heterogeneous framework of omics data that present challenges in communications and synchronizations [45]. There are tradeoffs between computational cost (floating-point operations), synchronization, and communications to consider while developing parallelization algorithms [66]. Moreover, developing error-free and secure applications is a challenging and labor-intensive, yet critically important task. Examples of programming errors and studies outlining wrongly mapped SNPs in commercial SNP chips have been reported in literature [67, 68, 69]. There is a need to validate the reliability of research platforms before considering the clinical utility of omics data. For instance, ToolShed, a feature of the Galaxy project that draws in software developers worldwide to upload and validate software tools, aims to enhance the reliability of bioinformatics tools. Novel tools and workflows with demonstrated usefulness and instructions are publically available (http://toolshed.g2.bx.psu.edu/) [70]. Both storage and computing platform such as Bioimbus [71], Bioconductor [72], CytoScape [73], are made available by scientists to exchange algorithms and data. There are many questions and methodologies that researchers may wish to consider, and this continuously drives on novel bioinformatics tools. Ultimately, light-weight programing environments and supporting programs with diverse cloud-based utilities are essential to enable those without or with limited programing skills to investigate biological networks [74]. Figure 3.3 illustrates a cloud-based framework that may help to implement personalized medicine. Much more programing efforts are still needed for the integration and interpretation of omics data in the transition to personalized medicine. Potential downstream applications are not always apparent when data are generated, promoting sophisticated flexible programs that may be regularly updated [75].

Integrative methods of omics data

Lastly, the depiction of biological systems through the integration of omics data requires appropriate mathematical and statistical methodologies to infer and describe causal links between different subcomponents [48]. The integration of omics data is both a challenge and an opportunity in biostatistics and biomathematics that is an increasing reality with the decreasing costs of omics profiles. Aside from the computational complexity of analyzing thousands of measurements, the extraction of correlations as true and meaningful biological interactions is not trivial. Biological systems include non-linear interactions and joint effects of multiple factors that make it difficult to distinguish signals from random errors. Caspase-8 for example, has opposing biological functions as it promotes cell death by triggering the extrinsic pathway of apoptosis, while having beneficial effects on cell survival through embryonic development, T-lymphocyte activation, and resistance to necrosis induced by tumor necrosis factor- (TNF-) [76]. Genes may carry out different functions in different cell types / tissues, which adds to the already substantial inter-individual variability. Biological complexity presents a challenge in extracting useful information within high-dimensional data [77]. Both computational and experimental methodologies are needed to fully elucidate biological networks. However, in contrast to experimental assays, computational models rely on biologic-ally driven variables and have inherent pitfalls of omics data.

Coping with to the curse of dimensionality

High-dimensionality is one of the main challenges that biostatisticians and biomathematicians face when deciphering omics data. It is the issue of large p, small n, where the number of measurements, p, is far greater than the number of independent samples, n [41, 77]. The analysis of thousands of measurements often leads to results with poor biological interpretability and plausibility. The reliability of models decreases with each added dimension (i.e. increased model complexity) for a fixed sample size (i.e. bias-variance dilemma, see Figure 2.4) [77]. All estimate instability, model overfitting, local convergence, and large

standard errors compromise the prediction advantage provided by multiple measures. A better understanding of these inherent caveats comes from the key concept behind statistical inference that is the distribution of repeated identical experiments. This distribution can be characterized by parameters such as the mean, and variance that quantify the average value (i.e. effect size), and degree of variability (i.e. biological or experimental noise). These parameters are estimated from observed data drawn from the true distribution (i.e. a finite number of independent samples). The reliability of estimates from a small sample size is low where it is more likely to observe estimates that deviate from the true distribution parameters. The chance of encountering such deviations also increases with the number of different measurements in a fixed sample. It is difficult to reliably estimate many parameters, and correctly infer associations from multiple hypotheses tested simultaneously. As a result, the analysis of both single and integrative omics data is prone to high rates of false-positives due to chance alone. This requires researchers to adjust for multiple testing to control for type 1 error rate using various methods based on the family-wise error rate (e.g. Bonferroni corrections, Westfall and Young permutation), and the false-positive rate (e.g. Benjamin and Hochberg) that are under strict assumptions [78, 79, 80, 81, 82, 83]. Another solution to overcome multiple testing issues is to reduce dimensionality via sparse methods that provide sparse linear combinations from a subset of relevant variables (i.e. sparse canonical correlation analysis, sparse principal components analysis, sparse regression) [84, 85]. Both mixOmics and integrOmics are publically available R packages for utilizing sparse methods on omics data [85, 86]. There are several approaches to derive optimal tuning parameters to dictate the number of relevant variables to pursue [87, 88]. However, stochastic processes to select best subsets of variables inferred from a given sample population may not contain the best information on another independent study, and certainly not at an individual level (i.e. selection-bias) [89, 90]. Reducing dimensionality is problematic as key mechanistic information could be lost. There is an overall tradeoff between false positive rates and the benefit of identifying novel associations within biological process that align with that of bias

and variance (Figure 2.4) [78].

The multi-level ontology analyses (MONA) is one ap-proach that bypasses the high-dimensionality as de-scribed by Sass et al. [91]. This method integrates multiple omics information (DNA sequence, mRNA and protein expressions, DNA methylation, and other regulation factors) and copes with redundancies related to multiple testing problems by approximating marginal probabilities using the expectation propagation algorithm [92]. The MONA approach allows for biological insights to be incorporated into the defined network as prior knowledge. This can address overfitting or uncertainty issues though reducing the solutions space to biological meaningful regions [93, 94]. This approach, however, relies on predefined known biological networks (i.e. proteinprotein interactions) or on the accuracy of mechanistic models (i.e. network models). Another strategy to analyze omics data involves integrating multiple data types into one single data set that holds maximum information. This reduces the complexity of omics data to the analysis of a single high-dimensional data set. Co-inertia analysis for example, has been used to integrate both proteomic and gene expression data to visualize and identify clusters of networks [95, 96]. It was initially introduced by Culhane et al. to compare gene expression data provided by different platforms, but has been further generalized to assess similarities between omic data sets [97]. The basic principal is to apply within and between principal component analysis, correspondence analysis, or multiple correspondence analysis while maximizing the sum of squares of covariances between variables (i.e. maximizing co-inertia between hyperspaces). The omicade4 package in R is available for exploring omics data using multiple co-inertia analysis [98]. Other similar, but conceptually different approaches include generalized singular value decom-position [99], and integrative bioclustering methods [100, 101]. An integrative omics study by Tomescu et al., have utilized all three approaches to characterize networks within Plasmodium faclicparum at different stages of life cycles [102]. Although the basic mathematical assumptions are different, the overlap in their results was considerable. The relative

importance and incremental value of individual omics data on one another may also be considered when predicting specific outcomes. For instance, Hamid et al. recently pro-posed a weighted kernel Fisher discriminant analysis that accounts for both quality and informativity of each individual omics data to integrate [103]. Significant improvements however, may not occur when data are redundant (i.e. correlated) or of low quality.

Mixing apples and oranges

Another challenge for integrating omics data lies in deriving meaningful interpretable correlations. For ex-ample, direct correlation analyses between transcriptomics and proteomics profiles are not valid in eukaryotic organisms. No high correlations between the two do-mains were observed as reported by multiple studies, and this was attributed to post-transcriptional and post-translational regulations [104, 105, 106, 107]. The advantage of inte-grating transcriptomic and proteomic data may diminish without accounting for regulation factors as the resulting inflated variability may limit reliability and reproducibility of findings [108]. Many complex traits are tightly reg-ulated and incorporating regulation factors may explain a relevant portion of observed variations due to true heterogeneity (i.e. true differences in effect sizes). Unlike the impact of noise on estimate precision which could be minimized by increasing the sample size, true heterogeneity may only be adjusted for during analysis when possible or via standardizations that limit generalizability. True heterogeneity poses a problem given biological complexity in the pursuit of precise effect size estimations (Figure 2.5). Hence, there is a need for network analysis to account for protein-protein and protein-DNA interactions in the context of integrating transcriptomics and proteomics data alone. An early study by Hwang et al. utilized network models to identify protein-protein and DNA-protein interactions with experimental verifications [109].

Bayesian networks are graphical models that involve structure and parameter optimization steps to represent probabilistic dependencies [110]. This modeling strategy that elucidates biological networks has been utilized in various studies [111, 112]. A seminal example

includes the use of dynamic Bayesian networks trained on chroma-tin data to identify expressed and non-expressed DNA segments in a myeloid leukemia cell line [113]. This was done by integrating position of histone modifications, and transcription factors binding sites at multiple intervals. It is however, a computationally demanding approach that requires advanced computing methods such as parallel computing and acceleration via GPUs [114]. Network models may serve as meaningful statistical results to be integrated with the biological domain. It has the potential to generate insight and a number of hypotheses on biological interactions to be experimentally and/or independently verified through a follow-up validation set. The ultimate goal is to continuously provide insight into biological interactions to subsequently build upon.

Separate the wheat from the chaff

It is important to minimize sources of error with omics data as it is challenging to distinguish between random error and true interaction signals. Hence, it is necessary to utilize statistical methods to account for sources of error. For example, the quality of omics data may vary between high-throughput platforms. Hu et al. have pro-posed quality-adjusted effect size models that were used to integrate multiple gene-expression microarray data given heterogeneous microarray experimental standards [115]. Omic studies are also prone to errors such as sample swapping and improper data entry. New methodologies for assessing data quality include Multi-Omics Data Matcher (MODMatcher) [116]. Moreover, complex diseases are often evaluated using a single phenotype that compromises statistical analysis by introducing errors such as misclassifications, and/or lack of account-ability for disease severity [117]. Modeling images for example, requires multiple phenotypes to properly capture image features [118]. Joint modeling of multiple responses to accurately capture complex phenotypes has been shown to increase power of discovery in genome-wide association studies [119]. There are even novel net-work methodologies to account for within-disease heterogeneity [120, 121]. Network approaches in model-ing complex diseases may provide a map of

disease pro-gression and play a major role in the proactive maintenance of wellness [122]. All reproducibility and validations of complex interaction signals are essential in the pursuit of personalized medicine. This highlights the growing need for metadata as the science of hows (i.e. data about data) to help harmonize omics studies and enable proper reproducibility of research results [123]. Examples of a metadata checklist and a metadata publication are available [124, 125]. Metadata may also serve as open innovations for integrative sciences, and may prove to be valuable for diversifying models of discovery and translation in high, and more importantly, low and middle-income countries. Altogether, validations on multiple data sets are required as evidence of stability, and that theoretically sound new methods outperform existing ones [126]. Both descriptive and mechanistic models for determining relevant biological networks require handling with care [127]. Software that integrate and interpret omics data are currently developed by competing companies in the private sector (e.g. Anaxomics, LifeMap), which may rapidly advance the field in the near future.

Conclusion

This review aims to stimulate research initiatives in the field of big data analysis and integration. Omics data embody a large mixture of signals and errors, where our current ability to identify novel associations comes at the cost of tolerating larger error thresholds in the con-text of big data. Major investments need to be made in the fields of bioinformatics, biomathematics, and biostatistics to develop translational analyses of omics data and make the best use of high-throughput technologies. New generations of multi-talented scientists and multidisciplinary research teams are required to build accurate complex disease models and permit effective personal-ized prevention, diagnosis and treatment strategies. Our ability to integrate and interoperate omics data is an important limiting factor in the transition to personalized medicine. Overcoming these limitations may boost the nation-wide implementation of omics facilities in clinical settings (Figure 2.6). The subsequent economies of scale

may in turn favor the access to personalized medicine to disadvantaged nations, repelling the growing shadow of two-tiered personalized medicine.

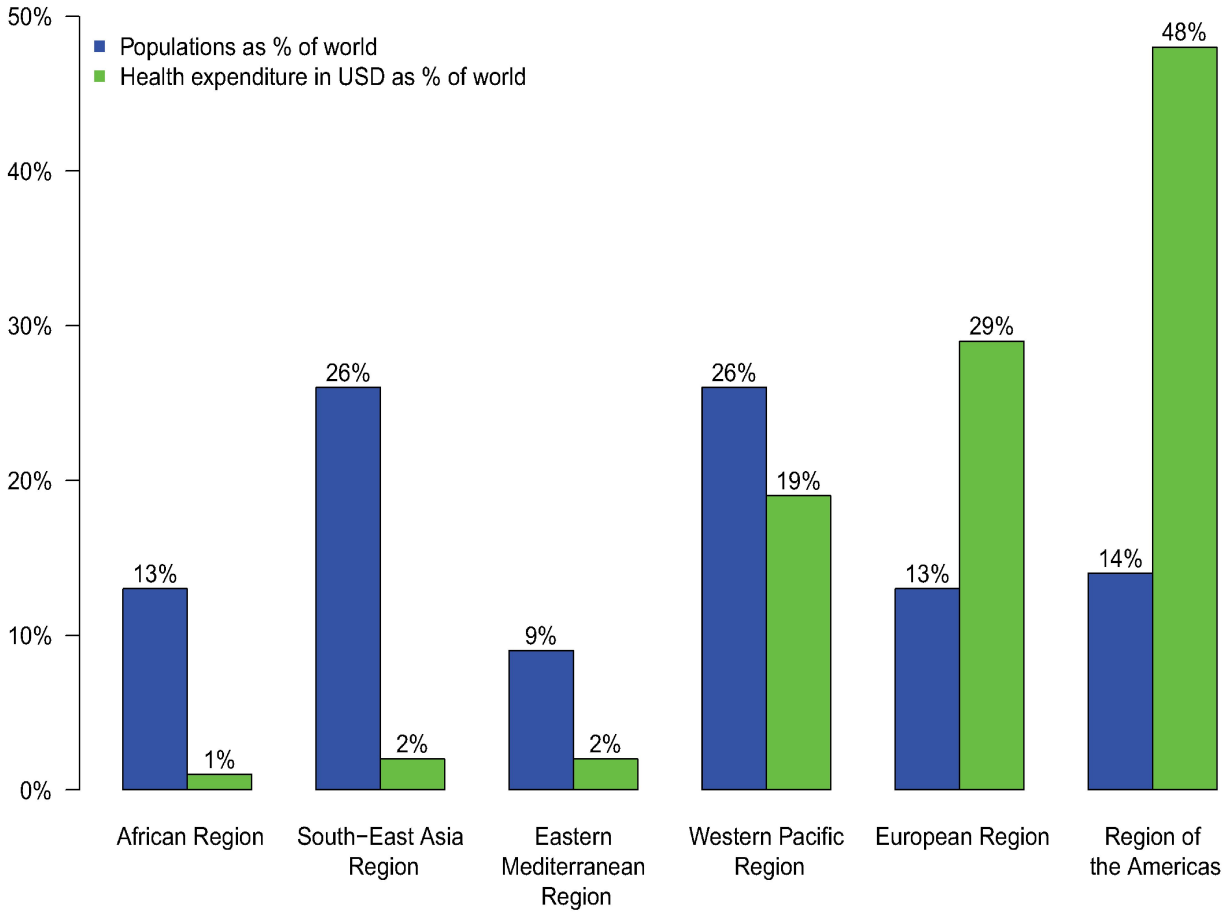


Figure 2.1: Distributions of populations and global health expenditure according to WHO 2012

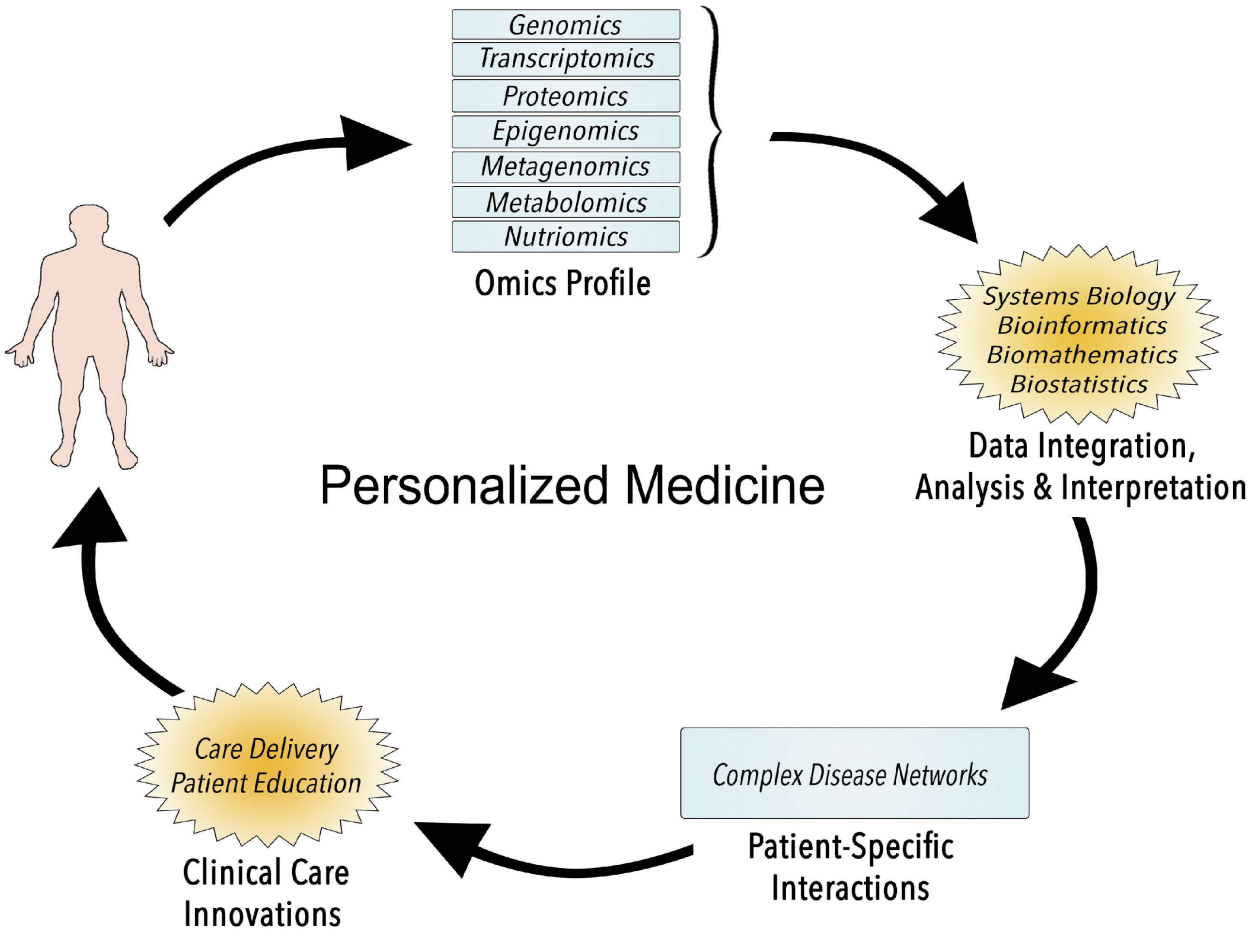


Figure 2.2: A basic framework of personalized medicine. The integration of omics profiles permit accurate modeling of complex diseases and opens windows of opportunities for innovative clinical applications to subsequently benefit the patient

Figure 2.3: An interdisciplinary cloud-based model to implement personalized medicine. The consecutive knowledge and service swapping between modeling and software experts in research and development units is essential for the management, integration, and analysis of omics data. Thorough software and model development will derive updates upon knowledge bases for complex diseases, in addition to clinical utilities, commercial applications, privacy and access control, user-friendly interfaces, and advanced software for fast computations within the cloud. This translates into personalized medicine via personal clouds that upload wellness indices into personal devices, electronic databases for health professionals, and innovative medical devices

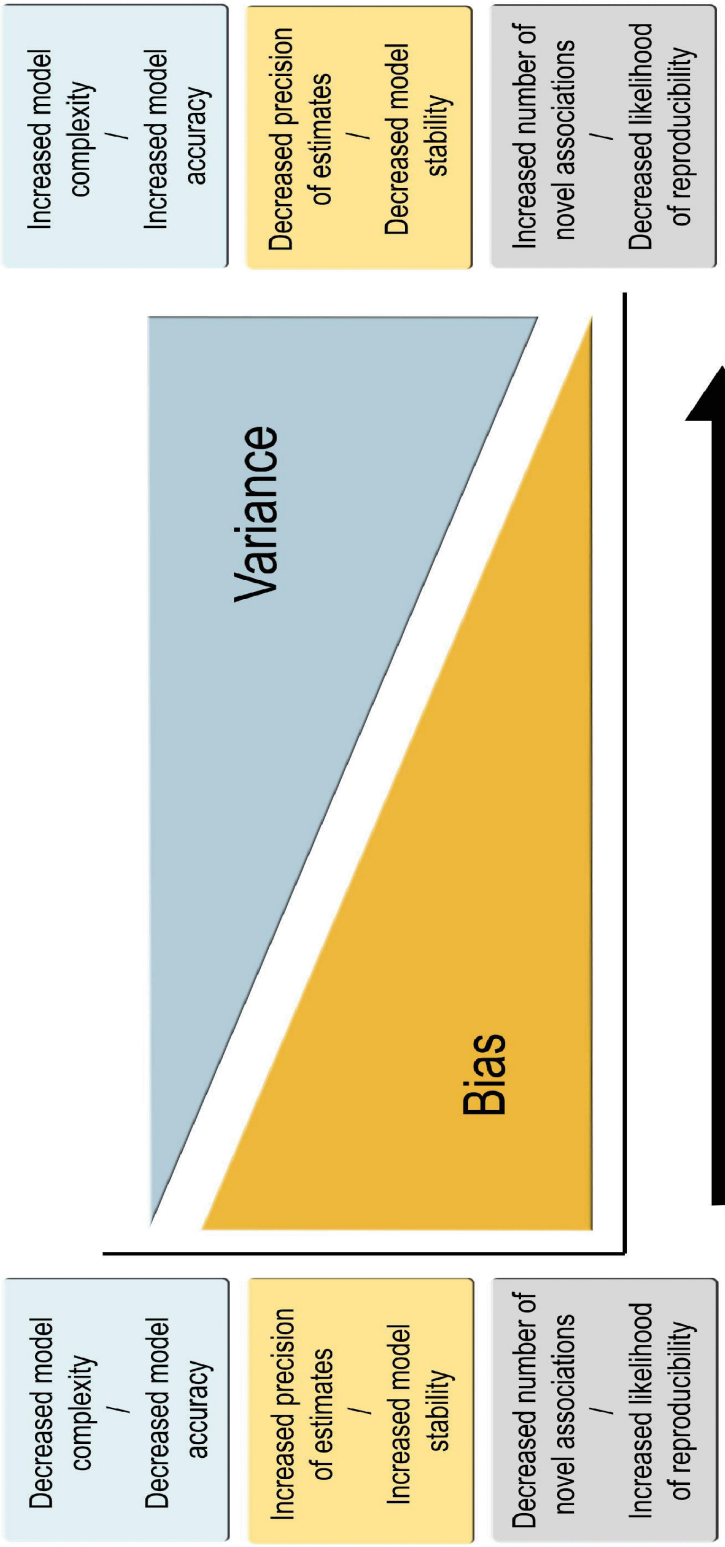


Figure 2.4: The bias-variance tradeoff with increasing model complexity

Figure 2.5: Noise and true heterogeneity within complex systems. Source of noise include measurement error and sampling variability. True heterogeneity however, is the result of true differences of effect sizes due to 1) the dynamic biological nature which encompasses feedback loops and temporal associations; and 2) multi-factorial complexity. Increasing the sample sizes is one solution to bypass noise and attain precise effect sizes, but true heterogeneity can only be adjusted during analysis when possible and via standardizations and calibrations that limit generalizability of the conclusions

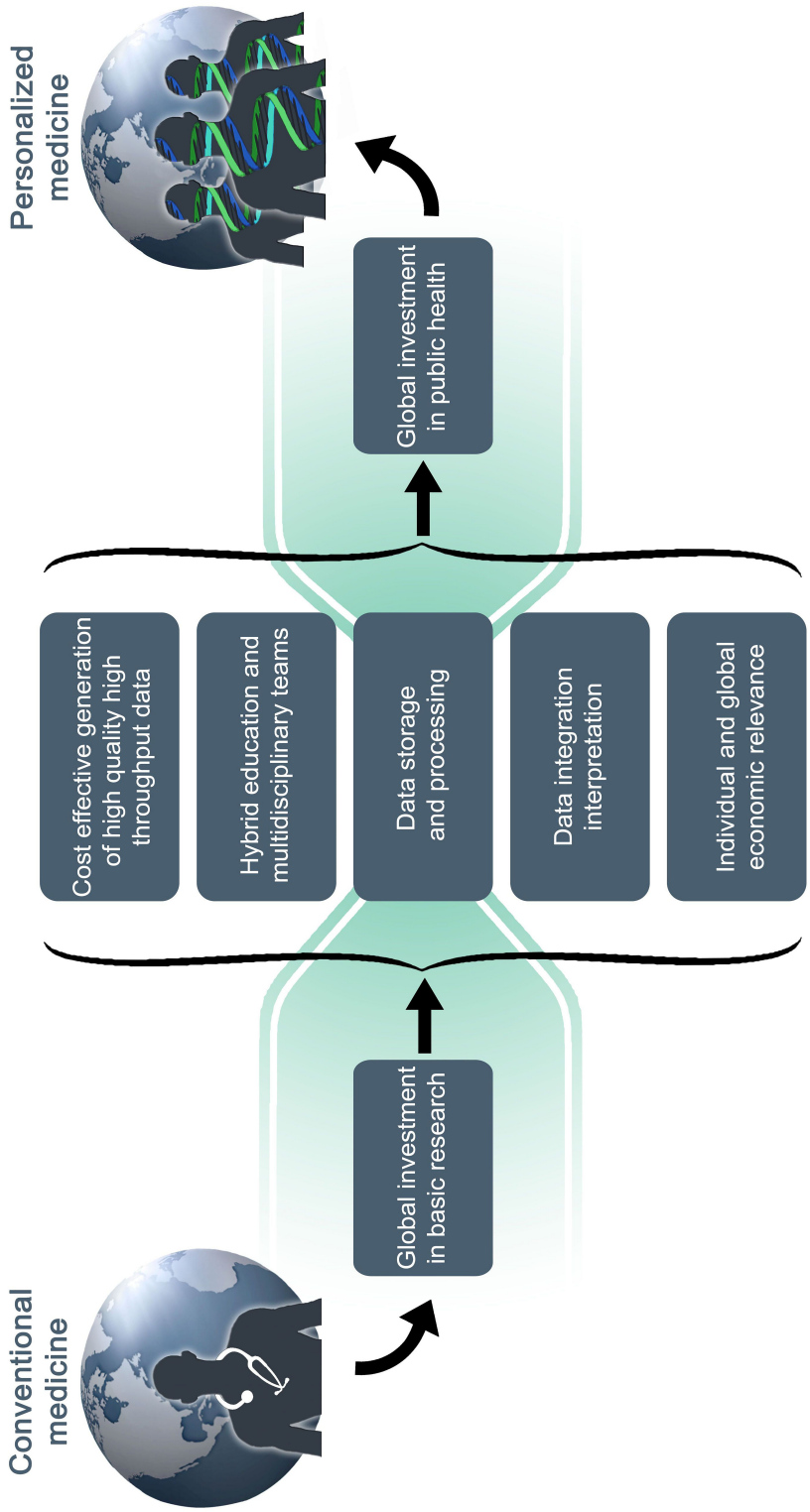


Figure 2.6: Bottleneck toward personalized medicine. The collective challenges to make the transition from conventional to personalized medicine include: i) generation of cost-effective high-throughput data; ii) hybrid education and multidisciplinary teams; iii) data storage and processing; iv) data integration and interpretation; and v) individual and global economic relevance. Massive global investment in basic research may precede global investment in public health for transformative medicine

Chapter 3

Penetrance of polygenic obesity susceptibility loci across the body mass index distribution

Abstract

A growing number of single nucleotide polymorphisms (SNPs) have been associated with body mass index (BMI) and obesity, but whether the effect of these obesity susceptibility loci is uniform across the BMI distribution remains unclear. We studied the effects of 37 BMI/obesity-associated SNPs in 75,230 adults of European ancestry along BMI percentiles using conditional quantile regression (CQR) and meta-regression (MR) models. The effects of 9 SNPs (24%) increased significantly across the sample BMI distribution including, FTO (rs1421085, $p = 8.69 \times 10^{-15}$), PCSK1 (rs6235, $p = 7.11 \times 10^{-6}$), TCF7L2 (rs7903146, $p = 9.60 \times 10^{-6}$), MC4R (rs11873305, $p = 5.08 \times 10^{-5}$), FANCL (rs12617233, $p = 5.30 \times 10^{-05}$), GIPR (rs11672660, $p = 1.64 \times 10^{-4}$), MAP2K5 (rs997295, $p = 3.25 \times 10^{-4}$), FTO (rs6499653, $p = 6.23 \times 10^{-04}$) and NT5C2 (rs3824755, $p = 7.90 \times 10^{-4}$). We showed that such increases stem from unadjusted gene interactions that enhanced the effects of SNPs in persons with

high BMI. When 125 height-associated were analyzed for comparison, only one ($< 1\%$), IGF1 (rs6219, $p = 1.80 \times 10^{-4}$), showed effects that varied significantly across height percentiles. Cumulative gene scores of these SNPs (GS-BMI and GS-Height, respectively) showed that only GS-BMI had effects that increased significantly across the sample distribution (BMI: $p = 7.03 \times 10^{-37}$, Height: $p = 0.499$). Overall, these findings underscore the importance of gene-gene and gene-environment interactions in shaping the genetic architecture of BMI and advance a method to detect such interactions using only the sample outcome distribution.

Introduction

Obesity is a prominent risk factor for osteoarthritis, hyper-tension, type 2 diabetes (T2D), cardiovascular disease, and certain psychological disorders and cancers [128, 129].The rise in obesity has coincided with 'obesogenic' societal and environmental changes that include increased consumption of high-calorie foods, an increasingly sedentary lifestyle, and urbanization [129, 130, 131]. Genetic factors are also known to play an important role in obesity, given that 50% - 80% of body mass index (BMI) variation can be ascribed to genetics (heritability) [132, 133]. Moreover, genome-wide association studies (GWASs) have identified 140 polygenic loci that are directly associated with BMI or obesity [134].

The role of individual and compound gene-environment (G×E) and gene-gene (G×G) interactions in determining BMI has not been fully elucidated. The study of BMI-associated G×G interactions has been impeded by statistical and computational limitations, although promising new approaches have recently been proposed [135, 136, 137]. On the other hand, several lines of evidence suggest that G×E interactions could play an important role in shaping BMI. First, estimates of the heritability of BMI are influenced by environmental exposures [138]. One study reported that the heritability of BMI is increased in persons born after the obesogenic transition, whereas another reported that the heritability of BMI

is correlated with the population prevalence of obesity [139, 140]. More recently, the cumulative gene score from 29 BMI-associated single-nucleotide polymorphisms (SNPs) showed a positive interaction effect with birth year [141]. Interactions between the genetic determinants of BMI and obesogenic environmental factors readily explain why both estimates of BMI heritability and cumulative SNP effects are enhanced in permissive environments. Second, specific interactions between BMI-associated SNPs and environmental factors have been documented [138]. Physical activity and energy intake have been reported to modify the effects of SNPs within the fat-mass- and obesity-associated gene FTO (MIM: 610966) [142, 143, 144, 145, 146]. Importantly, FTO (rs1421085) has been shown to jointly interact with diet, physical activity, salt and alcohol consumption, and sleep duration [147]. Thus, a subset of genetic variants could affect BMI through a mixture of direct effects and compound interactions. As such, investigating individual environmental factors might not capture the full range of environmental modification for a given SNP [148, 149].

In this report, we advanced a statistical framework to assess the effects of single and mixed G×E and G×G interactions on the association between SNPs and BMI. Specifically, we applied conditional quantile regression (CQR) to investigate the effects of 37 BMI-associated SNPs at multiple percentiles of the sample BMI distribution in 75,230 adults of European ancestry (EA) [150, 151]. Variability in SNP effects across these BMI percentiles was demonstrated to result from unadjusted interactions and was modeled by meta-regression (MR) [152, 153]. In this way, we used CQR and MR to collect evidence of unadjusted interactions directly from the sample distribution of BMI without measures of specific environmental factors. A secondary analysis of 125 established height-associated SNPs is also included for comparison.

Subjects and Methods

Participants and Phenotypes

The sample population included participants from the following studies: Atherosclerosis Risk in Communities (ARIC; phs000280.v3.p1), Coronary Artery Risk Development in Young Adults (CARDIA; phs000285.v3.p2), Cardiovascular Health Study (CHS; phs000287 .v6.p1), EpiDREAM, the Framingham Cohort (phs000007.v29.p10), Multi-Ethnic Study of Atherosclerosis (MESA; phs000209.v13.p3), Genetic Epidemiology of COPD (COPDGene; phs000179.v5.p2), Electronic Medical Records and Genomics (eMERGE) II (phs000888.v1.p1), and the Women's Health Initiative (WHI; phs000200.v10.p3). Measurements collected from participants below the age of 18 years or above the age of 92 years were excluded ($< 1\%$ collectively). For studies with repeated measures across multiple time points or visits, the median height and the median weight were extracted along with the corresponding age at these median values. We calculated BMI by dividing the median weight (in kg) by the square of the average measures of height (m). Diabetic status was indicated by one of the following criteria: (1) physician report or self-report of physician diagnosis, (2) report of taking diabetes medication, (3) fasting plasma glucose $\geq$ 126 mg/dL (7 mM), or (4) 2 hr glucose $\geq$ 200 mg/dL (11 mM) during an oral glucose-tolerance test [154]. Obesity categories including normal weight (NW) and over-weight(OW),as well as obesity classes I, II, and III (Ob-I,Ob-II,and Ob-III, respectively), were specified according to World Health Organization guidelines [155]. Analyses were restricted to participants of self-reported EA with a combined sample size of $n = 75,230$. Summary statistics are presented in Table A1. This project was approved by a local ethics committee (Hamilton Integrated Research Ethics Board), and participant-level data access was granted through the Database of Genotypes and Phenotypes (dbGaP) after approval was provided by study-specific data-access committees. All analyses are consistent with study-specific data-use certifications.

Sample Quality Control

Detailed genotyping procedures for EpiDREAM and studies from the Candidate Gene Association Resource (CARe) project, including ARIC (phs000557.v2.p1), CARDIA (phs000613. v1.p2), CHS (phs000377.v4.p1), the Framingham Cohort (phs000282. v17.p10), and MESA (phs000283.v7.p3), are presented elsewhere [156, 157]. Genotyping was performed with the gene-centric HumanCVD Genotyping BeadChip with 49,320 markers concentrated in ∼2,100 loci related to metabolism and cardiovascular disease [158]. This limited scope of analysis was motivated by the availability of a greater sample size, as well as the high computational cost of fitting CQR models. Samples with sex discordance, an array-wide call rate below $95\% - 98\%$, and/or an average heterozygosity beyond 3 standard deviations of the mean heterozygosity were removed [159, 160]. Family members were defined by identity by descent (IBD, $\hat{\pi}$) above 0.5, and those with a lower call rate were removed so that only one member of each family group was retained for analysis (Table A2). Samples from COPDGene (phs000765.v1.p2) were genotyped with the Illumina HumanHap550 (v3) genotyping BeadChip (Illumina) with 561,466 markers, and QC procedures were performed as above except that cryptic relatedness was defined by IBD $\hat{\pi} > 0.1875$ [161, 162]. Genotypes from the WHI study (phs000746.v1.p3) and eMERGE II (phs000888.v1.p1) were composed of an imputed dataset, and samples from related or duplicate participants were removed. Analyses of the WHI dataset were conducted on each sub-study (WHI Memory Study [WHIMS], WHI Genomics and Randomized Trials Network [GARNET], HIPFX [Hip Fracture GWAS], MOPMAP, and Genetics and Epidemiology of Colorectal Cancer Consortium [GECCO]). A summary of sample quality control (QC), along with a complete list of datasets (and accession numbers) and additional details on these studies, is provided in Table A2.

SNP Selection and Marker QC

We identified SNPs that had previously been associated with BMI, obesity, and height by searching the GWAS Catalog and GIANT Consortium data files and screening the literature

[163, 164, 165, 166]. A.A. and D.M. conducted literature screening independently to maximize SNP attainment. For GWAS SNPs, only associations with $p < 5 \times 10^{-8}$ were considered. These SNPs were sorted into correlated linkage disequilibrium (LD, $R^2 > 0.1$) blocks on the basis of genomic sequences from EA populations (1000 Genomes Project phase 3), and the strongest association SNP on the HumanCVD Genotyping BeadChip was selected [158, 167]. Proxy SNPs ($R^2 > 0.9$) were identified for SNPs not represented on the array. Thus, 39 BMI- and 129 height-associated SNPs were identified. For studies that used different genotyping platforms, the original association SNPs (39 BMI and 129 height) were screened and proxied as described above on each genotyping platform. For SNPs that mapped to the same gene, we screened them jointly with conditional regression analysis to test for independent associations with quantitative traits (BMI or height), and only SNPs that maintained associations were retained [168]. However, SNPs in FTO (rs1421085 and rs6499653) and PCSK1 (MIM: 162150; rs6232 and rs6235) were exempted from exclusion as a result of prior evidence in the literature of independent associations with BMI [169, 170, 171]. In total, 37 BMI- and 125 height-associated independent SNPs were identified and selected for further analysis. SNP call rate, minor allele frequency (MAF), and exact tests of Hardy-Weinberg equilibrium (HWE) in EA populations are presented in Tables A3 and A4. Within each study, SNPs with a call rate $< 90\%$ or HWE p value $< 1 \times 10^{-6}$ were excluded from analysis. In addition, only SNPs imputed with high quality were retained for analysis ($R^2 > 0.7$ for WHI and info score > 0.7 for eMERGE II) [172]. SNP genotypes were encoded per the effect alleles and modeled additively for individual analyses.

Gene Scores

The cumulative gene score (GS) was calculated for all BMI- and height-associated SNPs (GS-BMI and GS-height, respectively). An un-weighted GS was utilized because weights can be biased and context dependent [173, 174]. No GS was calculated for participants with more than 10% missing genotypes; otherwise, missing SNP genotypes were imputed

with the arithmetic average genotype at each missing SNP. In addition to being associated with BMI, GIPR (MIM: 137241; rs10423928, LD $R^2 = 1$ with rs11672660 in EA), TCF7L2 (MIM: 602228; rs7903146), TOMM40 (MIM: 608061) and APOE (MIM: 107741) (both rs2075650), HMGCR (MIM: 142910; rs4604177, LD $R^2 = 0.63$ with rs6453133 in EA), PCSK1 (rs6235), CDKAL1 (MIM: 611259; rs9356744), and KCNQ1 (MIM: 607542; rs2283228) have also been associated with several co-morbidities of obesity, including glucose homeostasis, T2D, increased lipid levels, and heightened C-reactive protein (CRP) levels [175, 176, 177, 178, 179, 180, 181, 182]. To mitigate potential biases stemming from these co-morbidities at higher BMI percentiles, we also calculated a GS excluding these seven SNPs: GS-BMI (stringent). Finally, GSs for both BMI and height were calculated without imputation of missing genotypes: GS-BMI (no imputation) and GS-height (no imputation). GS-BMI (stringent), GS-BMI (no imputation), and GS-height (no imputation) were tested by sensitivity analysis.

Statistical Analysis

A statistical framework combining CQR and MR was used to model variation in the effects of SNPs under single and mixed G×E and G×G interactions (see Supplemental Note) [151, 153]. Like ordinary least-squares (OLS) models, CQR models can assume a linear relationship and provide intercept and slope estimates for a series of pre-specified percentiles [150, 151]. Therefore, CQR can be applied to produce a comprehensive evaluation of the effects of a SNP across the sample distribution of a quantitative trait (e.g., BMI or height). A piecewise linear plot for the series of CQR estimates at different percentiles provides a useful visual summary of their variation along the sample distribution [150, 151]. Figure 3.1 shows a working example of CQR and MR in comparison with OLS for FTO (rs1421085) in the ARIC CARe study.

Under conditions where true single and mixed G×E and G×G interactions are unadjusted, SNPs will shift both the location and scale (variance) of the sample outcome distribution (see Supplemental Note) [183]. These shifts in scale result in detectable variations of CQR estimates collected from percentiles across the sample outcome distribution. It follows that CQR estimates for a SNP are constant (i.e., equal) across percentiles if all unadjusted interaction effects are zero. Thus, the association between SNPs and an outcome under unadjusted interactions essentially reduces to modeling variability in CQR estimates. This can be effectively achieved with MR [152, 153]. In this context, MR is basically a regression model where the CQR estimates from across the sample outcome distribution represent the dependent variable, and the percentiles at which these CQR estimates were calculated represent the independent variable (Figure 3.1). Additional details on CQR and MR, as well as simulations and an analytic description of this statistical framework, are presented in the Supplemental Note and Figures A1 and A2.

OLS models were used to verify the associations of SNPs and GSs with BMI and height in the sample populations included in this study. CQR models were fitted at every fifth percentile of the distribution of BMI and height for each SNP. We used a total of 10,000 Markov-chain marginal-bootstrap replicates to compute confidence intervals (CIs) and the cross-percentile variance-covariance matrix for CQR estimates [184, 185, 186]. The proportion of the trait variance explained by GS-BMI and GS-height in CQR models was also calculated [187]. We computed hypothesis test statistics in MR (by assuming normality) to estimate the effects of percentiles on changes in mean CQR estimates for each SNP. The set of percentiles ($5^{\text{th}} - 95^{\text{th}}$) was re-centered at the 50^{th} percentile so that the intercept of the MR models corresponded to the main effect of the SNP at the median. Lastly, the effects of each SNP and the GS on the risk of specific BMI categories (NW versus OW, NW versus Ob-I, NW versus Ob-II, and NW versus Ob-III) were estimated with logistic regression.

All regression models were performed by one-step individual- participant-data meta-analysis (also known as 'joint-analysis' or 'mega-analysis') [188, 189]. This method was

chosen on the basis of access to individual participant data and the fact that CQR analyses refer to the conditional sample distribution [190]. This means that analyses on separate studies correspond to their conditional distributions, and it would not be appropriate to combine them by using meta-analysis of their summary statistics. All models were adjusted for age (years), sex (female = 0, male = 1), and study (factor). For BMI analysis, age was modeled quadratically (age and age squared) as in previous reports [141, 147]. Analyses of the associations of SNPs and GSs with BMI (37 SNPs + GS = 38) and height (125 SNPs + GS = 126) were subject to multiple-testing correction using Bonferroni-adjusted p value thresholds of $p < 0.05/38 = 1.32 \times 10^{-3}$ and $p < 0.05/126 = 3.97 \times 10^{-4}$, respectively. 64 QC and statistical analyses were conducted with PLINK v1.90b3.42 and R v3.3.2 [159, 160, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201]. CQR models were fitted with quantreg, and MR models were fitted with metafor [202, 203]. Additional packages used in the analysis include pracma, doParallel, foreach, and data.table [204, 205, 206, 207]. An extended version of this work appears online [208].

Results

Figure 3.1 depicts a step-by-step analysis of FTO (rs1421085) in the ARIC CARe study. In the top left panel, we fitted an OLS model (green) to determine the mean effects of the FTO genotype on BMI (β_{OLS}, kg/m^2 per effect allele) and fitted CQR models (gray) evenly across the sample BMI distribution (every fifth percentile) to determine the effects of the FTO genotype at each BMI percentile (β_{CQR} , kg/m^2 per effect allele). In the middle right panel, the estimates (β_{OLS} and β_{CQR}) and 95% CIs from these models are collected and plotted against the BMI percentile at which they were fitted. In the bottom left panel, MR analysis (magenta) models variation in the CQR estimates across the sample BMI distribution, and MR estimates (β_{MR} , kg/m^2 per effect allele per BMI percentile) are plotted along with 95% CIs. Presenting the results of OLS, CQR, and MR in this way is useful for

summarizing the purpose of each analysis and contrasting possible differences between them.

Initially, OLS models were fitted for each of 37 BMI-associated SNPs, and all but one were verified to increase BMI in this study sample (Table 3.1). We then fitted CQR models at regular intervals of the BMI distribution to explore whether the effects of SNPs on BMI varied across the sample distribution (Table A5). We plotted CQR estimates for each SNP against the BMI percentiles at which they were produced to provide a visual summary of the CQR results (Figure 3.1 and Figure A3). Several SNPsincluding rs1421085 (FTO), rs6235 (PCSK1), rs7903146 (TCF7L2), rs11873305 (MC4R [MIM: 155541]), rs12617233 (FANCL [MIM: 608111]), rs11672660 (GIPR), rs997295 (MAP2K5 [MIM: 602520]), rs6499653 (FTO), and rs3824755 (NT5C2 [MIM: 600417])had effects that appeared to increase across the distribution of BMI.

Single or mixed SNP interactions that are not adjusted in regression models will produce variability in CQR estimates along the distribution of the outcome (see Supplemental Note). This variability can be detected and quantified with MR [152, 153]. Simulations showed that the power to detect such interactions by using CQR and MR was not affected by the MAF or the main effects of the SNPs, but it increased with the number of interactions as well as the main effects of the interacting covariate (see Supplemental Note and Figure A1). Yaghootkar et al. recently showed that differences in the prevalence of disease outcomes (e.g., the outcome of T2D) between sample and general populations can bias regression estimates of the main effects of SNPs on risk factors (e.g., BMI) [209]. However, the variability of CQR estimates across the sample distribution is not affected by biased main effects when CQR models are adjusted for disease status (see Supplemental Note). This was supported by simulations showing that the prevalence of disease outcomes in sample populations had negligible effects on the power and type I error rate for detecting unadjusted interactions when CQR models were adjusted for disease status (see Supplemental Note and Figure A2).

We fitted MR models to assess the variability in the CQR estimates of BMI-associated

SNPs along the sample distribution of BMI (Table 3.2, Figure 3.2, and Figure A3). Significant positive associations ($p < 1.32 \times 10^{-3}$) between BMI percentile and CQR estimates were detected for 9 of 37 SNPs (24%): rs1421085 (FTO; β_{MR} [95% CI] = 0.49 [0.37, 0.62], p = 8.69×10^{-15}), rs6235 (PCSK1; 0.32 [0.18, 0.46], 7.11×10^{-6}), rs7903146 (TCF7L2; 0.30 [0.17, 0.44], 9.60×10^{-6}), rs11873305 (MC4R; 0.60 [0.31, 0.89], 5.08×10^{-5}), rs12617233 (FANCL; 0.26 [0.13, 0.39], 5.30×10^{-5}), rs11672660 (GIPR; 0.29 [0.14, 0.45], 1.64×10^{-4}), rs997295 (MAP2K5; 0.23 [0.10, 0.35], 3.25×10^{-4}), rs6499653 (FTO; 0.25 [0.11, 0.40], 6.23×10^{-4}), and rs3824755 (NT5C2; 0.36 [0.15, 0.57], 7.90×10^{-4}). The estimates from MR (β_{MR}) quantify changes in the impact of each SNP on BMI across the sample distribution. For these 37 SNPs, the median β_{MR} value [Q_1, Q_3] was 0.135 [0.094, 0.217] kg/m^2 per effect allele per BMI percentile. In this statistical framework, β_{MR} is equal to zero if all SNP interaction effects are also equal to zero (see Supplemental Note). Positive β_{MR} estimates indicate that the effects of SNPs vary systemically by BMI percentile because unadjusted interactions are inflating the effects of SNPs in participants with a high BMI.

Given that height is known to be highly heritable, analyses were extended to height for comparison with the BMI results [149, 210, 211]. OLS models were fitted for each of 125 height-associated SNPs, and all but two were verified to increase height (Table A6). CQR and MR were used to estimate variation in the effects of these SNPs on height as described previously (Figure A4 and Table A7). Only one height-associated SNP, rs6219 (IGF1 [MIM: 147440], β_{MR} = [95% CI] = 0.48 [0.23, 0.73], p = 1.80×10^{-4}), showed significantly ($p < 3.97 \times 10^{-4}$) increased effects along the sample height distribution (Table A8). For height-associated SNPs, the median β_{MR} value [Q_1, Q_3] was 0.002 [-0.056, 0.085] cm per effect allele per height percentile. Thus, CQR estimates for height-associated SNPs were predominantly consistent across height percentiles, and < 1% showed evidence of unadjusted interactions, whereas 24% of BMI-associated SNPs did.

We combined BMI- and height-associated SNPs into GSs (GS-BMI and GS-height, respectively) to examine the overall association of these SNPs across the sample distribution.

OLS models were used to verify the positive association between GS-BMI and GS-height and their respective traits (Table 3.3). CQR models for GS-BMI showed steadily increasing effects with increasing percentiles, whereas CQR models for GS-height did not vary across percentiles (Figure 3.3). MR analysis indicated that percentiles were significantly and positively associated with CQR estimates for GS-BMI (β_{MR} [95% CI] = 0.15 [0.13, 0.17], 7.03 $\times 10^{-37}$) but not GS-height (0.01 [-0.01, 0.02], 0.499) (Table 3.3). At the 10^{th} and 90^{th} BMI percentiles, each additional effect allele of GS-BMI increased BMI by 0.054 and 0.167 kg/m^2 (3.1-fold increase), respectively, whereas each additional allele of GS-height increased height by 0.172 and 0.180 kg/m^2, respectively (Tables A5 and A7). Thus, in 1.73-m-tall persons at the tenth BMI percentile, carrying ten additional BMI-increasing alleles was associated with 1.6 kg of extra weight, whereas at the 90^{th} BMI percentile, this was associated with 5.0 kg of extra weight. Furthermore, at the 10^{th} and 90^{th} BMI percentiles, the proportion of trait variance explained by GS-BMI increased (2.7-fold from 0.130% to 0.357%), whereas that of GS-height was stable (1.825% to 1.822%) (Tables A5 and A7). These results support the conclusion that the impact of BMI-associated SNPs was larger for individuals with high BMI, whereas the impact of height-associated SNPs varied little by height.

Excluding seven SNPs that have also been associated with comorbidities of obesity from the gene score GS-BMI (stringent) did not alter the pattern of increasing effects across the sample BMI distribution (Figure A5) [175, 176, 177, 178, 179, 180, 181, 182]. Moreover, MR analysis indicated that BMI percentile was significantly and positively associated with the CQR estimates for GS-BMI (stringent) (β_{MR} [95% CI] = 0.14 [0.11, 0.16], p = 2.18×10^{-23}). In addition, CQR models were re-fitted with adjustment for diabetic status because this had been shown to mitigate the effects of possible stratification within the sample population (see Supplemental Note and Figure A2). Of the nine SNPs whose effects showed significant increases across the sample BMI distribution (Table 3.3 and Figure 3.2), three have also been associated with glucose homeostasis and T2D, namely, GIPR (rs11672660), TCF7L2 (rs7903146), and PCSK1 (rs6235) [175, 177, 180]. Refitting CQR models with adjustment for

diabetic status had little impact on the results from MR analysis of these SNPs or GS-BMI (Table A9). Additional sensitivity analysis that included linearly modeling the effects of age or testing fewer percentiles (i.e., every 10^{th} percentile from the 5^{th} to 95^{th} BMI percentiles) also showed no substantial changes to MR results (Table A9). Furthermore, calculating the GS for each trait without imputing missing genotypes did not affect results for GS-BMI or GS-height (Figure A5). Finally, the results from CQR were compared with those obtained from conventional subgroup analysis. To this end, the effect of genotype on the risk of OW, Ob-I, Ob-II, and Ob-III was evaluated separately with logistic regression (Table A10). The odds ratios of each SNP for each category were plotted against the BMI percentiles of the corresponding category, and CQR estimates were then overlaid on these bar plots. The patterns from logistic regression models across BMI categories were qualitatively consistent with the patterns from CQR models at comparable BMI percentiles (Figure A6).

Discussion

The aim of this study was to investigate variations in the effect of 37 BMI-associated SNPs across the distribution of BMI. We introduced a method that applies CQR to model the effects of SNPs at different percentiles of the sample BMI distribution and estimates variability in these effects by using MR. CQR estimates at different percentiles were shown to be uniform if all unadjusted SNP interactions were zero (see Supplemental Note). It follows that SNPs whose CQR estimates vary significantly across the sample BMI distribution are regulated by such interactions.

CQR analysis revealed distinct profiles of associations of BMI SNPs across the sample BMI distribution. Several of these SNPs had effects that increased steadily at higher BMI percentiles, whereas others had uniform effects that varied little across BMI percentiles (Figure 3.2 and Figure A3). One other study has used CQR to investigate the association between BMI and FTO (rs1558902) and a GS in a modest sample of adults [212]. The

patterns reported by that study are consistent with the results reported here [212]. Two other studies used CQR to investigate the effects of SNPs on BMI in European children, and their results are also comparable with those here [213, 214]. Overall, the high degree of correspondence between previously reported CQR results from European children and those from adults presented here emphasizes the robustness of these findings. Furthermore, the patterns observed with CQR analysis were compared with those from conventional logistic regression (subgroup analysis), given that Berndt et al. have demonstrated that the genetic architecture of BMI strongly overlaps BMI categories (Table A10) [215]. Across BMI categories, the patterns from logistic regression were largely consistent with those from CQR (Figure A6). CQR overcomes several of the limitations of subgroup analysis by utilizing all sample data to estimate regression parameters on the same scale as the continuous outcome, and comparing CQR estimates from different quantiles is relatively intuitive and easy [150, 215]. 23,89

MR was applied in order to model changes in the effects of BMI SNPs across the sample BMI distribution [152, 153]. Results from MR showed that BMI percentile was positively and significantly associated with CQR estimates for 9 of 37 SNPs (24%). In addition, nominal associations were also observed for several other SNPs, and the median β_{MR} [Q_1, Q_3] was 0.135 [0.094, 0.217] kg/m^2 per effect allele per BMI percentile (Table 3.2 and Figure A3). This is supported by the GS-BMI analysis, which also showed significantly increasing effects across the sample BMI distribution (Figure 3.3 and Table 3.3). These findings indicate that unadjusted interactions enhanced the effects of BMI-associated SNPs at higher BMI levels. Modeling the effects of age linearly or considering fewer BMI percentiles (i.e., every tenth rather than every fifth percentile) had minimal effects on these results (Table A9).

There is evidence that differences in disease prevalence (e.g., in T2D) between sample and general populations can result in the stratification of secondary traits (e.g., BMI) that are risk factors for disease [209]. This stratification can compromise regression estimates of the main effects of SNPs on secondary traits, and naively adjusting regression models for

disease status might not adequately address this [209]. Although the main effects of SNPs from disease-adjusted regression models are susceptible to stratification bias, the variation of SNP effects across the sample distribution is not (see Supplemental Note). This was evident in simulations showing that stratification had little effect on the power and type I error rate of MR analysis when CQR models were adjusted for disease status (Figure A2). Because GIPR (rs11672660), TCF7L2 (rs7903146), and PCSK1 (rs6235) have been associated with glucose homeostasis and T2D, CQR models were refitted with adjustment for diabetic status and analyzed by MR [175, 177, 180]. These SNPs and the GS continued to show significantly increasing effects across the sample BMI distribution with this adjustment, demonstrating that the results were not an artifact of possible sample stratification (Table A9). Although estimating the variability of disease-adjusted CQR estimates across the sample distribution by using MR is robust to stratification bias, future studies aimed at estimating the main effects of SNPs by using CQR should implement methods to address this potential source of bias [216]. A total of 7 of the 37 obesity-predisposing loci that were selected for analysis have also been associated with comorbidities of obesity, including glucose homeostasis, T2D, increased lipid levels, and heightened CRP levels. Excluding these SNPs from the GS did not alter the pattern observed across the sample BMI distribution or affect the results from MR analysis, suggesting that these findings do not stem from the influence of comorbidities at high BMI levels (Figure A5).

Although BMI was the primary focus of this report, these analyses were also applied to height. This was important because analysis of height could shed light on the nature of the unadjusted interactions that were detected. BMI is a composite of both height and weight-height is one of the most heritable complex human traits, and weight is strongly influenced by environmental exposures and behavior [138, 217]. If unadjusted interactions in the effects of BMI-associated SNPs are predominantly due to G×G interactions, then it is reasonable to suppose that these unadjusted interactions would be detected at a similar frequency in other quantitative traits such as height. On the other hand, if G×E interactions predominate, then

these unadjusted interactions might be less frequently detected in quantitative traits with a smaller environmental component (i.e., height). CQR models for 125 height-associated SNPs were mostly uniform and exhibited little variability across height percentiles (Figure A4). Only one significant association between height percentiles and CQR estimates for height SNPs was detected by MR, and the median β_{MR} [Q_1, Q_3] was 0.002 [-0.056, 0.085] cm per effect allele per height percentile (Table A8). Moreover, the effects of GS-height did not vary along the sample height distribution, which suggests that unadjusted interactions do not affect the genetic architecture of height to the same extent that they do for BMI (Table 3.3 and Figure 3.3). The simplest explanation for the discrepancy between the results for GS-BMI and GS-height is that the unadjusted interactions detected from GS-BMI were predominantly G×E interactions.

G×E interactions for SNPs in FTO have been reported for physical activity, food intake, dietary salt,alcohol consumption, and sleep duration [218, 219, 220, 221].In addition, the association between TCF7L2 (rs12255372) and BMI was modulated by fat intake in a weight-loss trial [222]. Our analyses also pointed to significant interactions for FTO (rs1421085) and TCF7L2 (rs7903146) but suggested that such interactions might extend to additional BMI-associated SNPsincluding rs6235 (PCSK1), rs11873305 (MC4R), rs12617233 (FANCL), rs11672660 (GIPR), rs997295 (MAP2K5), rs6499653 (FTO), and rs3824755 (NT5C2)and GS-BMI. This is entirely consistent with a report showing that the effects of GS-BMI (29 SNPs) were enhanced by increased greater exposure to obesogenic environments and another demonstrating interactions between GS-BMI (69 SNPs) and several obesogenic drivers, including socio-economic status, TV watching, 'Westernized' diets, and physical activity [141, 223]. These reports also support the argument that the unadjusted interactions detected for BMI SNPs are predominately G×E interactions. Environmental modification of the effects of genetic variants raises the possibility that preventive measures, sustained lifestyle modifications, and therapeutic interventions could attenuate some of the genetic predisposition to unhealthy BMI. Indeed, the overall effect of BMI SNPs is minimal at low BMI

levels (Figures 3.2 and 3.3). If weight gain leads to a genetically driven 'vicious circle,' then weight loss can lead to a genetically driven 'virtuous circle.' Investigating additional BMI-associated SNPs by using CQR and MR to uncover the full extent of unadjusted interactions in the architecture of BMI will be the focus of future studies.

This study is the largest yet to apply CQR to examine how the effects of SNPs vary with BMI, and it establishes quantitative support for hitherto qualitative descriptions of CQR. The combined utility of CQR and MR presents a contemporary statistical framework to cue hypotheses on gene interactions, better define clinical risks associated with genetic profiles, and prioritize clinical targets. Future studies aimed at distinguishing variants whose effects are modified by unadjusted interactions from those with fixed effects could advance the field of precision medicine. With the combined application of CQR and MR, this can now be achieved solely with information contained within the sample outcome distribution.

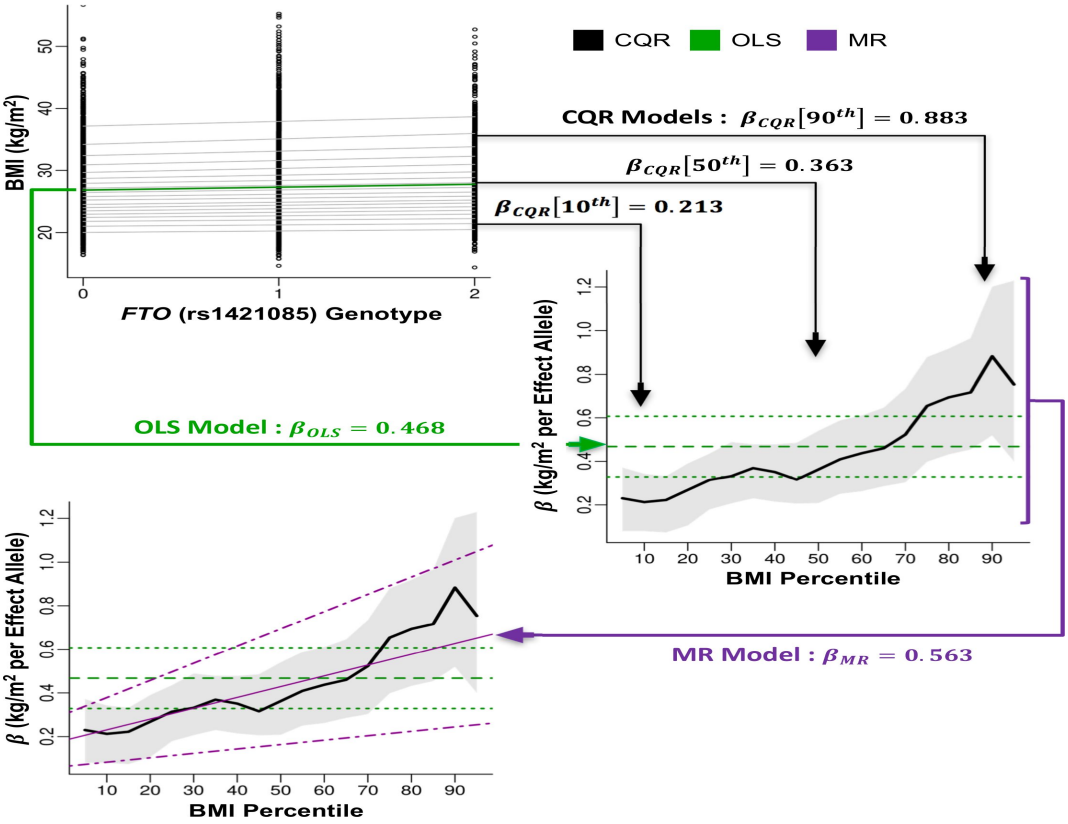


Figure 3.1: Working example of conditional quantile regression. BMI (kg/m^2) was plotted against the number of effect alleles of FTO (rs1421085) in the ARIC CARe study (top-left). An ordinary least squares (OLS) model of the mean effect of this SNP on BMI was plotted (solid-green line). Conditional quantile regression (CQR) models, fitted every 5^{th} percentile of BMI, show the effects of this SNP at these BMI percentiles (solid-grey lines). The slopes (β_{OLS}, horizontal-dashed-green line; β_{CQR}, thick-black line; kg/m^2 per Effect Allele) from these models were then plotted against BMI percentile at which they were fitted (middle-right). 95% confidence intervals for these estimates are also plotted (OLS, horizontal-dotted-green line; CQR, shaded-grey region). The change in CQR estimates across BMI percentiles was modeled using meta-regression (MR). The MR slope (β_{MR}, kg/m^2 per Effect Allele per BMI percentile, thin-magenta line) and the 95% confidence intervals (dotdashed-magenta lines) were plotted (bottom-left).

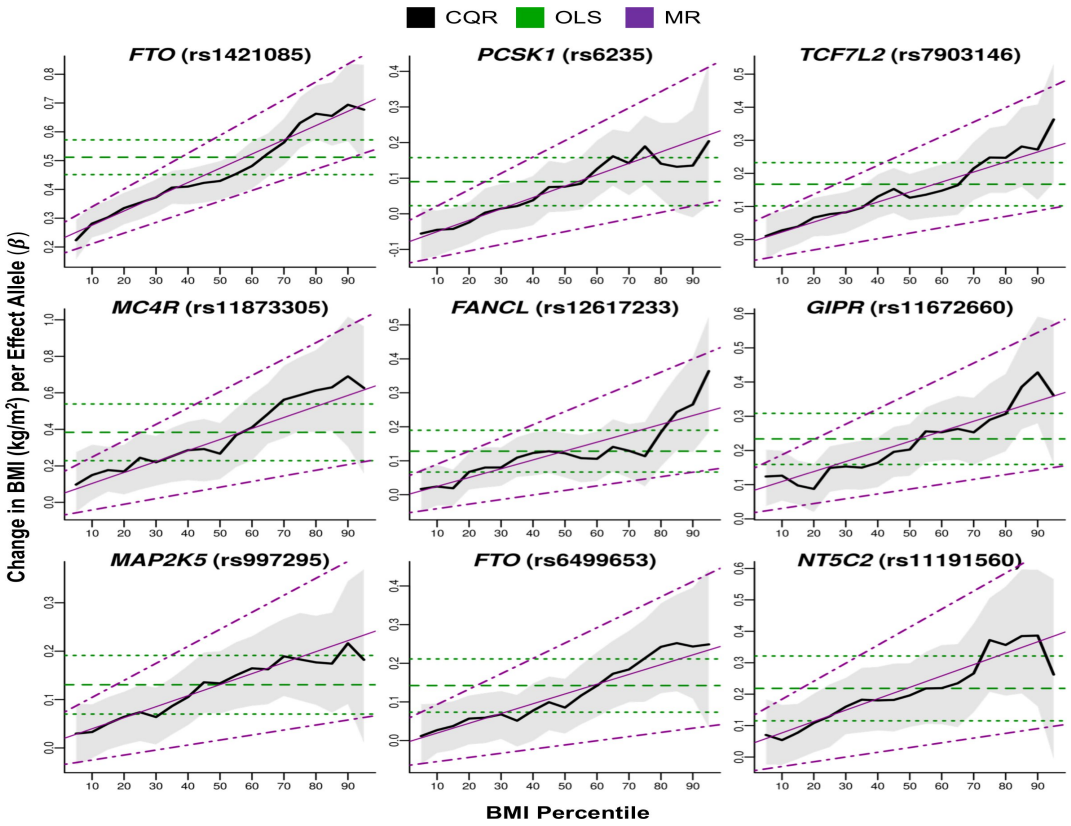


Figure 3.2: The effects of BMI/obesity-associated SNPs across the sample BMI distribution. CQR models of BMI-associated SNPs were fitted every fifth percentile of BMI and adjusted for age, age squared, sex, and study. Estimates of the change in BMI (kg/m^2) per effect allele (β_{CQR}) from these models were plotted against the BMI percentile (thick black line) along with the 95% confidence intervals (shaded gray region). The results from OLS models (β_{OLS}, kg/m^2 per effect allele, horizontal dashed green line) and the 95% confidence intervals (horizontal dotted green lines) were also plotted for comparison. The change in CQR estimates across BMI percentiles was modeled with MR, and estimates from MR (β_{MR}, kg/m^2 per effect allele per BMI percentile, thin magenta line) and the 95% confidence intervals (dotted magenta lines) were plotted. MR analysis detected significant ($p < 1.32 \times 10^{-3}$) increases in the effects of these SNPs across the sample BMI distribution.

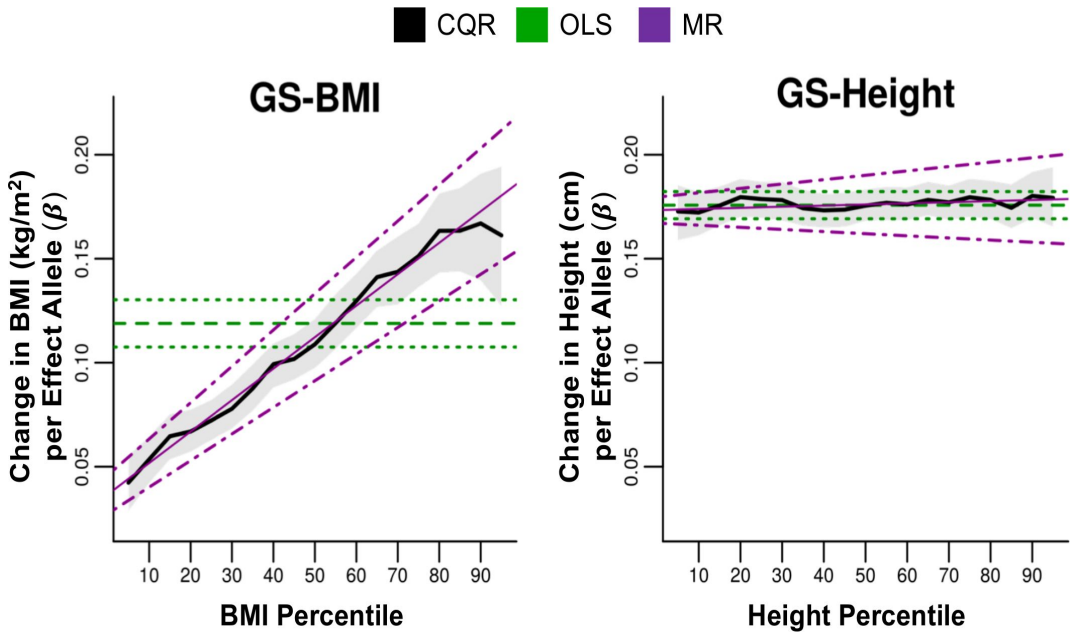


Figure 3.3: The effects of GS-BMI and GS-Height across the sample distribution of BMI and height, respectively. As in Figure 2, CQR models of the GS-BMI and GS-Height were plotted against the BMI percentile and height percentile, respectively. The thick-black line is the estimated change in each trait per effect allele (GS-BMI, β_{CQR}, kg/m^2 per Effect Allele; GS-Height, β_{CQR}, cm per Effect Allele) and shaded-grey region represents the 95% confidence intervals. Also plotted are the OLS regression estimates (GS-BMI, β_{OLS} in kg/m^2 per Effect Allele; GS-Height, β_{OLS}, cm per Effect Allele, horizontal-dashed-green line) and 95% confidence intervals (horizontal-dotted-green lines). The change in CQR estimates across outcome percentiles was modeled using meta-regression (MR). Estimates from MR (GS-BMI, β_{MR}, kg/m^2 per Effect Allele per BMI Percentile; GS-Height, MR, cm per Effect Allele per Height Percentile; thin-magenta line) and the 95% confidence intervals (dotdashed-magenta lines) were also plotted.

Table 3.1 BMI-Associated SNP Information and Results from OLS Models. 37 BMI-predisposing SNPs were selected for analysis. The effect and other (E and O, respectively) alleles were based on original discovery studies (PMID), and SNPs were coded by BMI-increasing or obesity-predisposing alleles. The indicated positions are based on GRCh37, and all alleles are on the positive strand. The association between these SNPs and BMI was assessed by OLS models that were adjusted for age, age squared, sex, and study. b OLS is the effect size (kg/m^2 per effect allele), and 95% CIs are the 95% confidence intervals.

SNP	Gene (OMIM)	Chromosome Position	E/O	PMID	β_{OLS} [95% CI]	p Value
rs1421085	FTO (610966)	chr16: 53,800,954	C/T	17658951	0.512 [0.451, 0.572]	5.88×10^{-62}
rs10767664	BDNF (113505)	chr11: 27,725,986	A/T	20935630	0.246 [0.172, 0.319]	5.89×10^{-11}
rs11672660	GIPR (137241)	chr19: 46,180,184	C/T	25673413	0.234 [0.159, 0.309]	8.16×10^{-10}
rs4788099	SH2B1 (608937)	chr16: 28,855,727	G/A	23001569	0.180 [0.113, 0.246]	1.13×10^{-7}
rs7903146	TCF7L2 (602228)	chr10: 114,758,349	C/T	25673413	0.167 [0.102, 0.232]	5.36×10^{-7}
rs2075650	TOMM40 (608061)	chr19: 45,395,619	A/G	23001569	0.218 [0.131, 0.305]	9.75×10^{-7}
rs11873305	MC4R (155541)	chr18: 58,049,192	A/C	25673413	0.384 [0.229, 0.539]	1.23×10^{-6}
rs997295	MAP2K5 (602520)	chr15: 68,016,343	T/G	23001569	0.131 [0.070, 0.191]	2.40×10^{-5}
rs3824755	NT5C2 (600417)	chr10: 104,595,849	C/G	25673413	0.218 [0.115, 0.321]	3.32×10^{-5}
rs12617233	FANCL (608111)	chr2: 59,039,998	C/T	23001569	0.128 [0.067, 0.190]	4.34×10^{-5}
rs6499653	FTO (610966)	chr16: 53,877,592	T/C	25673413	0.142 [0.073, 0.211]	5.19×10^{-5}
rs1788826	NPC1 (607623)	chr18: 21,154,024	G/A	25673413	0.124 [0.061, 0.186]	1.08×10^{-4}
rs17066846	MC4R (155541)	chr18: 58,044,818	G/T	25673413	0.144 [0.068, 0.220]	2.09×10^{-4}
rs6453133	HMGCR (142910)	chr5: 74,692,776	A/G	25673413	0.124 [0.058, 0.189]	2.18×10^{-4}
rs739564	IQCK	chr16: 19,740,237	A/G	25673413	0.147 [0.067, 0.227]	2.97×10^{-4}
rs2272903	TFAP2B (601601)	chr6: 50,786,571	G/A	23001569	0.173 [0.076, 0.270]	4.77×10^{-4}
rs7553158	TNNI3K (613932)	chr1: 75,005,238	G/A	25673413	0.102 [0.042, 0.162]	8.40×10^{-4}
rs11570094	SPI1 (165170)	chr11: 47,359,706	A/C	25673413	0.107 [0.041, 0.172]	1.37×10^{-3}
rs4946932	FOXO3 (602681)	chr6: 108,974,746	C/A	25673413	0.107 [0.041, 0.174]	1.57×10^{-3}
rs2819347	LMOD1 (602715)	chr1: 201,884,288	G/C	25673413	0.101 [0.037, 0.165]	1.89×10^{-3}
rs2836754	ETS2 (164740)	chr21: 40,291,740	C/T	25673413	0.099 [0.033, 0.164]	3.20×10^{-3}
rs2984618	TAL1 (187040)	chr1: 47,690,438	T/G	25673413	0.087 [0.026, 0.148]	5.17×10^{-3}
rs11208662	LEPR (601007)	chr1: 65,987,164	C/G	23563609	0.139 [0.037, 0.242]	7.66×10^{-3}
rs6235	PCSK1 (162150)	chr5: 95,728,898	G/C	18604207	0.090 [0.023, 0.158]	8.82×10^{-3}
rs9356744	CDKAL1 (611259)	chr6: 20,685,486	T/C	22344219	0.071 [0.005, 0.137]	0.035
rs7988412	MTIF3	chr13: 28,000,282	T/C	25673413	0.090 [0.005, 0.175]	0.037
rs1780050	NEXN (613121)	chr1: 78,400,540	A/C	25673413	0.063 [0.002, 0.124]	0.042
rs526134	USP37	chr2: 219,402,371	G/A	25673413	0.066 [0.000, 0.132]	0.049
rs980828	NOS1AP (605551)	chr1: 162,306,415	G/T	25133637	0.050 [0.010, 0.110]	0.100
rs17001561	SCARB2	chr4: 77,096,118	A/G	25673413	0.070 [0.017, 0.157]	0.113
rs6232	PCSK1 (162150)	chr5: 95,751,785	C/T	18604207	0.095 [0.041, 0.232]	0.172
rs749767	KAT8 (609912)	chr16: 31,124,407	A/G	25673413	0.042 [0.022, 0.105]	0.199
rs1211166	NTRK2 (600456)	chr9: 87,285,992	A/G	23001569	0.041 [0.034, 0.116]	0.289
rs2535633	ITIH4 (600564)	chr3: 52,859,630	G/C	24861553	0.024 [0.037, 0.085]	0.437
rs10144353	PRKCH (605437)	chr14: 61,911,157	T/C	23563609	0.044 [0.067, 0.155]	0.441
rs1561288	ADCY3 (600291)	chr2: 25,369,002	C/T	23669352	0.024 [0.047, 0.095]	0.507
rs2283228	KCNQ1 (607542)	chr11: 2,849,530	C/A	24861553	0.037 [0.159, 0.085]	0.550

Table 3.2: Quantifying the Effect of BMI Percentile on CQR Estimates. MR was used to model variability in the CQR estimates across BMI percentiles. Note that the percentiles were re-centered around the 50 th percentile so that the intercept from MR models would correspond to the main effect of the SNP at the median. Asterisks (*) denote statistical significance at the Bonferroni-adjusted threshold of $p < 1.32 \times 10^{-3}$, RI 50 is the re-centered intercept of the MR models, b MR is the effect of BMI percentile on CQR estimates (kg/m^2 per effect allele per BMI percentile), and 95% CIs are the 95% confidence intervals.

SNP	Gene(MIM)	RI_{50}	β_{MR}	$[95\% CI]$	p Value
rs1421085	FTO(610966)	0.473	0.495	[0.370, 0.620]	8.69×10^{-15}*
rs6235	PCSK1(162150)	0.078	0.320	[0.180, 0.459]	7.11×10^{-6}*
rs7903146	TCF7L2(602228)	0.144	0.303	[0.169, 0.437]	9.60×10^{-6}*
rs11873305	MC4R(155541)	0.344	0.603	[0.311, 0.895]	5.08×10^{-5}*
rs12617233	FANCL(608111)	0.129	0.261	[0.134, 0.387]	5.30×10^{-5}*
rs11672660	GIPR(137241)	0.227	0.294	[0.141, 0.447]	1.64×10^{-4}*
rs997295	MAP2K5(602520)	0.131	0.228	[0.103, 0.352]	3.25×10^{-4}*
rs6499653	FTO(610966)	0.121	0.253	[0.108, 0.398]	6.23×10^{-4}*
rs3824755	NT5C2(600417)	0.222	0.362	[0.151, 0.574]	7.90×10^{-4}*
rs7553158	TNNI3K(613932)	0.099	0.196	[0.071, 0.322]	2.12×10^{-3}
rs10767664	BDNF(113505)	0.247	0.217	[0.064, 0.370]	5.50×10^{-3}
rs4788099	SH2B1(608937)	0.151	0.194	[0.057, 0.332]	5.59×10^{-3}
rs17066846	MC4R(155541)	0.124	0.215	[0.063, 0.367]	5.61×10^{-3}
rs9356744	CDKAL1(611259)	0.063	0.186	[0.050, 0.322]	7.35×10^{-3}
rs6453133	HMGCR(142910)	0.13	0.177	[0.040, 0.314]	0.011
rs2819347	LMOD1(602715)	0.111	0.137	[0.004, 0.269]	0.044
rs2075650	TOMM40(608061)	0.283	0.161	[0.019, 0.341]	0.079
rs4946932	FOXO3(602681)	0.106	0.120	[0.016, 0.256]	0.084
rs2984618	TAL1(187040)	0.069	0.108	[0.019, 0.235]	0.095
rs980828	NOS1AP(605551)	0.024	0.095	[0.030, 0.220]	0.135
rs1788826	NPC1(607623)	0.109	0.094	[0.036, 0.224]	0.156
rs11570094	SPI1(165170)	0.103	0.096	[0.039, 0.231]	0.163
rs7988412	MTIF3	0.088	0.109	[0.062, 0.280]	0.212
rs2283228	KCNQ1(607542)	0.003	0.147	[0.094, 0.388]	0.232
rs739564	IQCK	0.122	0.100	[0.065, 0.265]	0.234
rs526134	USP37	0.062	0.079	[0.055, 0.212]	0.247
rs2272903	TFAP2B(601601)	0.145	0.113	[0.084, 0.310]	0.261
rs2836754	ETS2(164740)	0.086	0.073	[0.060, 0.206]	0.28
rs2535633	ITIH4(600564)	0.016	0.068	[0.059, 0.194]	0.296
rs11208662	LEPR(601007)	0.142	0.111	[0.105, 0.327]	0.314
rs6232	PCSK1(162150)	0.075	0.133	[0.137, 0.404]	0.334
rs749767	KAT8(609912)	0.048	0.058	[0.075, 0.191]	0.39
rs1561288	ADCY3(600291)	0.027	0.037	[0.185, 0.112]	0.627
rs10144353	PRKCH(605437)	0.043	0.049	[0.171, 0.269]	0.662
rs1211166	NTRK2(600456)	0.029	0.027	[0.179, 0.126]	0.731

Continued on next page

Table 3.2 – *Continued from previous page*

SNP	Gene(MIM)	RI_{50}	β_{MR} $[95\%CI]$	p Value
rs17001561	SCARB2	0.068	0.020 $[0.194, 0.154]$	0.824
rs1780050	NEXN(613121)	0.045	0.010 $[0.117, 0.136]$	0.883

Table 3.3 Analysis of GS-BMI and GS-Height

	OLS Models			MR Models		
SNP	β_{OLS}	$[95\%CI]$	p Value	RI_{50}	β_{MR} $[95\%CI]$	p Value
GS-BMI	0.119	$[0.108, 0.130]$	3.48×10^{-93}	0.112	0.151 $[0.128, 0.175]$	$7.03 \times 10^{-37}*$
GS-height	0.176	$[0.169, 0.182]$	2.2×10^{-308}	0.176	0.005 $[0.010, 0.021]$	0.499

Chapter 4

Meta Quantile Regression: A Novel Method for Detecting Potential Gene Interactions using Sample Distributions of Complex Traits

Abstract

Background: The effect of genetic variants on complex traits includes interaction components that are challenging to reliably detect. Meta-Quantile Regression (MQR) is a framework that combines conditional quantile regression (CQR) and meta-regression to infer potential interactions by modeling variations in genetic effects across the sample distribution of quantitative traits.

Objectives: Compare the utility of MQR and variance heterogeneity tests for detecting potential interactions.

Methods: The relationships between variance per genotype and MQR were analytically investigated. MQR fitted using CQR were termed as MCQR to differentiate them from MQR

models fitted using unconditional quantile regression (UQR) which were termed as MUQR. The computational cost and asymptotic convergence rate of MCQR and MUQR estimates were compared using simulations. Variance heterogeneity tests investigated include Levene's and Brown-Forsythe F-tests for a total of 4 tests of potential interactions (Levene, Brown-Forsythe, MCQR, and MUQR). Simulations were conducted to compare their type I error and power by 1) the number of genotype group levels; 2) symmetric, asymmetric error, and inverse-normal rank transformation to treat skewness; and 3) synergistic and antagonistic interactions.

Results QR estimates were analytically shown to be influenced by unadjusted interactions that capture the change in distribution spread by genotypes. MUQR models were found to use less CPU time to fit and provide estimates that asymptotically converge faster than MCQR models. Furthermore, rank-transformations were shown to inflate type I error rates for 4 four tests on genotypes with main effects. Both MCQR and MUQR were found to have higher power of detecting potential interactions with the number of genotype group levels; under asymmetric error, and antagonistic interactions compared to variance heterogeneity tests.

Conclusions: MQR models are useful for identifying potential interactions and MUQR is a computationally feasible framework for genome-wide association studies.

Introduction

Complex traits are influenced by a combination of environmental and genetic components. Over the past decade, a growing number of genetic variants have been associated with complex traits using genome-wide association studies (GWAS) [165, 164, 224]. Nonetheless, a large proportion of the heritability for many complex traits remains unexplained [225, 226, 227]. For example, the heritability of body mass index (BMI) is estimated at 40-75%, but variants discovered using GWAS only account for 2.7% of the overall variability in

BMI [164, 132, 133, 228]. While many variants directly associated with such traits are yet to be discovered, it is likely that genetic interactions (e.g. gene×gene / gene×environment) constitute a substantial proportion of the missing of heritability [227, 229]. Identifying genetic interactions is important for elucidating the genetic architecture and biological networks that underlie complex traits.

Genetic interactions refer to circumstances where an interacting variable modifies the effects of a genetic variant on a phenotype. Interacting variables could be other genetic variants (i.e. epistasis), environmental (e.g. pollutants), biological (e.g. sex or age), behavioral exposures (e.g. smoking or unhealthy eating) or medical conditions (e.g. chronic diseases). Genome-wide interaction study (GWIS) designs have been developed to detect interactions by exploring complex genetic models with interaction components in a regression framework [230, 231, 232]. Interactions are modeled assuming multiplicative (i.e. classic two-way interactions), threshold-based effects (i.e. conditional effects of variants given an exposure threshold), or as part of an intermediate latent model (i.e. structural equation models via relational graphs). GWIS designs are conceptually appealing but face challenges in reproducibility that are exacerbated by heterogeneity across studies (i.e. differences by the degree in exposures of the interacting variable) [225, 233, 234, 235, 236, 237]. Several studies have shown that the detection of interaction effects requires larger sample sizes compared to main effects with similar effect sizes [238, 239, 240]. The power to detect interactions depends on multiple factors other than the magnitude of interaction effects. The nature (i.e. antagonistic vs synergetic interactions) and degree of exposure to the interacting variable(s) (i.e. low *vs* high variability of exposure to the interacting variables in the sample population) can significantly influence the power to detect them [241]. In addition, current approaches for the development of novel statistical methods to detect interactions are limited by prior assumptions to leverage power that include 1) independence of gene and environment for case-only analyses of binary traits, 2) the presence of main effects for interacting variants to enable filtering via marginal genetic associations [242, 243]. Lastly, accurate and reliable

measurement of environmental exposures remains an area of active investigation [218]. For example, collecting dietary intake data that is representative of real behavior faces critical methodological limitations [244]. As a net result, few GWIS have been successful in identifying interactions while many claims have failed to replicate [234, 245, 246]. There is an urgent need for robust statistical methods to detect genetic interactions at a more general level without the need for measurement of known or hypothesized interacting variables. Such methods would provide a means for detecting potential interactions reliably at the cost of limiting the knowledge and specifications on the source and nature of interactions. Examples for the nature of interactions include multiplicative or threshold-based interactions, while examples for the source of interactions is the unmeasured or unknown interaction variable (e.g. unmeasured diet or physical activity levels). The detection of variants with potential interaction would help draw further investigations into them without limiting the scope of explorations on all variants by a single type of interaction and interaction variable.

Differences in phenotype variance across genotypes have recently gained attention as a potential statistic for detecting interactions [247, 248, 183, 249]. For example, if a bi-allelic genetic variant (e.g. single nucleotide polymorphisms - SNPs) interacts with physical activity to affect BMI, then the variability in BMI will be higher in subjects carrying the BMI-increasing allele given that a portion of these carriers will engage in some level of physical activity. Hence, testing for differences in variance across genotypic categories can provide evidence of gene interactions. However, variance heterogeneity tests rely on subgroup location and variance estimates that are affected by sample size per level and group imbalances [250]. Group imbalance refers to the degree of disparity in proportions of factor levels may results in ill-representation of levels that is further jeopardized with increasing number of factor levels for a fixed sample size. Hence, variance heterogeneity tests will have lower power of detecting potential interacting variants with a high group imbalance (e.g. rare variants) or number of genotype levels (e.g. triallelic genotypes or genotypes of species with more than 2 chromosome copies). However, the variance is only one of several measures

for a distribution's spread which include the range and, interquartile range. It is therefore possible for phenotype distributions to have similar variances across genotypes but different interquartile range, minimum and maximum values by genotype. The variance may not be the most robust statistic for the spread of asymmetric distributions due to sensitivities towards distribution tails. Many phenotypes of interest have skewed or heavy-tailed distributions (e.g. BMI, plasma glucose, or protein expression data) for which the mean and variance are not necessarily the best measurements of location and shape.

An alternative approach to capture differences between distributions by genotype is to model the effect of variants across the sample distribution using quantile regression (QR). This is based on the principle that genetic interactions induce changes in the strength of genetic effects across percentiles of a trait. QR models the effect of a predictor on the position of a specified quantile of the outcome distribution [251]. By examining multiple quantiles, QR can produce a comprehensive picture for the effect of a predictor on the outcome distribution. First degree differences in QR estimates across percentiles denote changes to distribution spread (e.g. variance or the inter-quartile range) [186]. They can be modeled using meta-regression (MR). Combining QR and MR to model how QR estimates of genetic variants change across the response distribution similar to heteroscdasticity tests [186]. We term framework as Meta-Quantile Regression (MQR), and it is useful for identifying genetic variants with potential interactions. MQR does not require knowledge of the interacting variable(s). It utilizes phenotype and genotype information is necessary to infer potential interactions similar to variance heterogeneity tests. We have recently applied MQR using conditional quantile regression (CQR) to demonstrate that a substantial proportion of BMI/obesity-associated SNPs (24% - 9/37 SNPs) show evidence of genetic interactions [208]. While CQR is useful for investigating a limited number of genetic variants, the computational cost of fitting CQR models is prohibitive at scale. CQR parameters have no closed form solution and therefore require optimization methods for their estimation. In addition, the asymptotic variance-covariance matrix of estimates is challenging to evaluate

because it depends on an unknown response density and has a slow asymptotic convergence rate. Hence bootstrapping methods are required to reliably estimate the variance-covariance matrix, which limits the utility of QR methods in genome-wide analysis [252].

This study assesses the utility of modeling QR estimates from unconditional quantile regression (UQR) instead of CQR for identifying variants with potential interactions [253]. UQR parameters and their respective variance-covariance matrix have a closed solution and can potentially scale better with the number of variants, sample size, number of percentiles and number of covariates (Supplementary Material). A short description of the difference between CQR and UQR is provided in the supplementary material. This study applied simulations to assess type I error and the power to detect potential interactions compared with tests of variance heterogeneity tests.

Materials and Methods

Model Formulations

To formally introduce and illustrate the utility of QR in genome-wide association analysis, consider the response vector Y from n independent and identical distributed (i.i.d) samples with *cdf* $F_Y(y)$. Assume the following linear model

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 g_i + \beta_3 x_i g_i + \epsilon_i \tag{4.1}$$

where x_i corresponds to an interacting variable with $X \sim F_X(\mu_x, \sigma_x^2)$; g_i is the observed genotype of the genetic variant, G, under HWE with the population allele frequency, p, where $G \sim Bin(2, p)$; and ϵ_i is the random error with $\epsilon \sim F_\epsilon(0, \sigma_\epsilon)$. β_0, β_1, β_2 and β_3 correspond to the intercept, marginal effect of the interacting variable, marginal effect of the genotype, and the interaction effect between them, respectively. The conditional distribution of Y can be described as $F_{Y|X=x,G=g} \sim \left(\beta_0 + \beta_1 x + \beta_2 g + \beta_3 x g, \sigma_y^2\right)$. Assuming that X and

G are independent, then the conditional density of Y given G can be shown to have a mean and variance:

$$E\left[Y \mid G = g\right] = (\beta_0 + \beta_1\mu_x) + (\beta_2 + \beta_3\mu_x)g \tag{4.2}$$

$$Var(Y \mid G = g) = \sigma_x^2(\beta_1 + \beta_3 g)^2 + \sigma_\epsilon^2 \tag{4.3}$$

Note that $c = \beta_1 + \beta_3 g$ is an amplification factor for the conditional variance that corresponds to the remaining variability in Y given G. Under an additive genetic model, response variances by genotypes are all equal if and only if $\beta_3 = 0$. Hence, tests of differences in response variance per genotype can identify variants with potential interactions. Note that the conditional variance per genotype has a minimum at $\beta_3 = -\beta_1$. If the marginal and interaction effects are in opposite directions, then the power to detect potential interactions is reduced. Hence, failure to identify differences in variance by genotype effects does not rule out interactions effects [247]. Equation 4.3 indicates that gene interactions will increase or decrease variance consistently across genotype categories. identify true interaction under the genetic model of inHowever, tests for variance heterogeneity detect variance inconsistency rather than variance structure (i.e. direction of change). Therefore, variance heterogeneity is not specific to interaction signals, but includes conditions where no consistent direction for increasing or decreasing variance per genotype is observed. Modeling relationship between variance and genotypes assuming a structure (i.e. linear trend) could help improve power to detect variants with potential interactions if the assumtions are met. This can be done as a reformulation of heteroscedasticity in equation 4.3 into the QR framework. The linear scale model of heteroscedasticity [i.e. $\sigma(x) = (1 + x\theta)\sigma$] is a special case of QR models with linear conditional quantile functions. Note that other models of systematic heteroscedasticity can be approximated for small θ by a linear expansion [186]. This includes multiplicative heteroscedasticity in the form of $\sigma(x) = e^{x\theta}$ [254, 255]. Hence, assuming a linear heteroscedastic

function is reasonable for variants with small interaction effects. Nonetheless, modest interaction effects require the correct heteroscedasticity function (i.e. linear, or nonlinear)to detect, while small interaction effects may remain difficult to identify and/or non-clinically relevant.

To reformulate heteroscedasticity into the QR framework, let ϵ_x be the partial error for the unadjusted interacting variable X with mean zero and variance σ_x^2. Then the conditional distribution $F_{Y|G}(y \mid G = g)$ can be translated into a heteroscedastic linear model with partitioned residuals where $\sigma_{\epsilon_x}(g) = (\beta_1 + \beta_3 g)$. That is

$$y_i = \big(\beta_0 + \beta_1\mu_x\big) + \big(\beta_2 + \beta_3\mu_x\big)g_i + \big(\beta_1 + \beta_3 g_i\big)\epsilon_{x,i} + \epsilon_i \tag{4.4}$$

The conditional quantile function for the heteroscedastic model under i.i.d errors is

$$Q_Y(\tau \mid G = g) = \beta_0(\tau) + \beta_G(\tau)g \tag{4.5}$$

where

$$\beta_0(\tau) = \beta_0 + Q_\epsilon(\tau) + \beta_1\big[\mu_x + Q_{\epsilon_x}(\tau)\big]$$
$$\beta_G(\tau) = \beta_2 + \beta_3\big[\mu_x + Q_{\epsilon_x}(\tau)\big]$$

Again, assuming that X and G are independent (i.e. $E\big[X \mid G = g\big] = E\big[X\big]$), the effects of genetic variant G on the τ-th quantile of Y are equal (i.e. $\beta_G(\tau) = \beta_2$ for all $\tau \in (0, 1)$) if and only if $\beta_3 = 0$. $\beta_G(\tau)$ is in fact a linear function of $Q_{\epsilon_x}(\tau)$ with slope given by interaction coefficient β_3. Note that the independence assumption between X and G is used for simplification and illustration purposes. Overall, $\beta_G(\tau)$ inherits properties of $Q_{\epsilon_x}(\tau)$ that include monotonicity with τ and captures both variability, and skewness of the interacting variable X. An example is provided in the Supplementary Material.

Modeling QR Estimates

Modeling and testing the dependency structure of $\boldsymbol{\beta_G(\tau)}$ with $\boldsymbol{\tau}$ can be used as an overall robust indicator of potential interactions. Note that no parametric assumptions on partial errors due to interaction variables are required. Let $\widehat{\boldsymbol{\beta}}_G(\tau)$ and $\widehat{\boldsymbol{\Sigma}}_G$ denote the subset of QR regression estimates and their corresponding variance-covariance matrix of the genetic variant in the model. The relationship between genetic effects and percentiles $\boldsymbol{\tau}$ centered at the median can be investigated using MR (also known as generalized least squares). Let $A = \begin{pmatrix} \mathbf{1}' & \boldsymbol{\tau}^* \end{pmatrix}$ be the design matrix, where $\boldsymbol{\tau}^* = \boldsymbol{\tau} - 0.5$. The regression coefficients of MR models are given by

$$\begin{aligned} \widehat{\boldsymbol{\beta}}_M &= \left(A^t \widehat{\boldsymbol{\Sigma}}_G^{-1} A\right) A^t \widehat{\boldsymbol{\Sigma}}_G^{-1} \widehat{\boldsymbol{\beta}}_G(\tau) \\ &= \begin{pmatrix} \widehat{\beta}_G & \widehat{\beta}_\tau \end{pmatrix} \end{aligned} \tag{4.6}$$

with variance-covariance matrix of the estimates

$$\widehat{\boldsymbol{\Sigma}}_M = A^t \widehat{\boldsymbol{\Sigma}}_{G,\tau}^{-1} A$$

In this formulation, $\widehat{\beta}_G$ and $\widehat{\beta}_\tau$ correspond to the variant's marginal and change in marginal effects with percentiles respectively. We provide an analytical demonstration for the relevance of MR estimates on potential interactions using QR under the null hypothesis of no interactions (Supplementary Material). Moreover, CQR estimates have been shown to asympototically converge to a normal distribution under i.i.d errors [186, 256, 257]. These asymptotic properties of CQR estimates have varying converegnce rate that weaken with extreme distribution tails [258]. As a result, modeling extreme quantiles may require bias corrections and inference using simulations and suitable bootstrap methods [259]. Hence, bootstrap methods via re-sampling to estimate $\beta(\tau)$ and Σ_M provide robust inference and characterize the joint normal density of estimates across quantiles. Nonetheless, bootstrap methods can be computatally intensive for large models and sample data as in classic GWAS

study designs. Hence, we only focus on the original asymptotic estimator of $\widehat{\Sigma}_G$ by Koenker and Bassett 1978 for i.i.d error [251]. Up to our knowledge, there are no studies assessing the asymptotic convergence rate of $\widehat{\beta}(\tau)$ and $\widehat{\Sigma}_G$ from UQR models. Note that $\widehat{\Sigma}_G$ from UQR models using least-squares regression is the covariance-variance matrix of multivariate regression errors given as $\boldsymbol{\epsilon}'\boldsymbol{\epsilon}/(n-r-1)$, where r is the number of predictors. The estimated $\widehat{\Sigma}_G$, $\widehat{\beta}_G(\tau)$ follows a Student's t distribution with $df = n - r - 1$ which is approximately normal for large sample sizes. Hence, applying MR result in the modelling of normally distributed estimates with their corresponding variance-covariance matrix. Hence, provided that Σ_M was estimated correctly, a test for marginal effect of variants is simply testing for $H_0 : \beta_G = 0$:

$$Z_G = \frac{\widehat{\beta}_G}{SE_{\widehat{\beta}_G}} \sim N(0,1) \tag{4.7}$$

A linear trend with $\boldsymbol{\tau}$ can be tested using $H_0 : \beta_\tau = 0$ as,

$$Z_\tau = \frac{\widehat{\beta}_\tau}{SE_{\widehat{\beta}_\tau}} \sim N(0,1) \tag{4.8}$$

Supplementary Figure B3 shows the convergence of Z_τ from on CQR and UQR estimates toward nominal Type I error rate of 0.05. Deviations from the null distribution mainly occur when modeling extreme quantiles and by the number of quantiles modeled to a lesser extent. Nonetheless, hypothesis tests for Z_τ using UQR estimates converge much faster to the correct null distribution compared to CQR estimates.

Overall, the modeling of $\boldsymbol{\beta_G(\tau)}$ with $\boldsymbol{\tau}$ provides estimates and inference for marginal and potential interaction effects. β_G is interpreted similarly to the classic median change in the response by one unit change in the number of allele copies. On the other hand, β_τ is interpreted as the one unit change in the marginal effect with one unit change in percentile. That is, β_τ correspond to the inflation or deflation of marginal effects due to unadjusted interactions. Hence, when β_τ is zero, there is no evidence of potential interactions. A positive value indicates that marginal effects of the genetic variant are greater for samples at the upper

side of the response distribution compared to those at the lower side, while a negative value indicates the opposite. The linearity assumption of $\boldsymbol{\beta_G(\tau)}$ with $\boldsymbol{\tau}$ is used to summarize the overall inflation or deflation of marginal effects due to interacting variables that are present in the sample population. The magnitude and direction of β_τ reflect quantiles of the partial error(s) density, $\boldsymbol{\epsilon_X}$, which vary according to the variance-covariance matrix of interacting variables, $\boldsymbol{\Sigma_X}$ and multiplicatively amplified by interaction effects. The density of interacting variables is specific to sample population of interest, and one may consider that differences in β_τ between populations may reflect differences in the density of interacting variables (i.e. differences in exposures of different populations to interacting variables). For simplicity, the use of MR to estimate the variability of QR estimates by τ will henceforth be referred to as meta-quantile regression (MQR), meta-conditional quantile regression (MCQR), and meta-unconditional quantile regression (MUQR) as needed.

Simulations

Data Generation

Genotypes were generated from $\text{Bin}(\eta, p)$ centered at $\mu_g = 0$ with $\sigma_g = \eta p(1-p)$, where η is one plus the number of chromosomes (i.e. $\eta = 2$ for two levels of an indicator for whether the single chromosome contains the reference allele) and p is the allele frequency. For simplicity purposes, genotypes are generated from $\eta = 2$, unless otherwise specified. Genotypes were investigated under an additive genetic effect with allele frequencies ranging from 0.05 to 0.95. The interacting variable X was generated from a standard normal distirbution. The response variable, Y, was simulated from the linear model in equation 4.1. Coefficients of individual covariates were calculated as a function of pre-defined % variance explained (R^2) by individual marginal and interaction variables denoted as R^2_G, R^2_X, and $R^2_{G\times X}$. Regression coefficients were specified as $\beta_0 = 0$, $\beta_1 = \sqrt{R^2_X}$, $\beta_2 = \sqrt{R^2_G/\big(\eta p(1-p)\big)}$, and $\beta_3 = \sqrt{R^2_{G\times X}/\big(\eta p(1-p)\big)}$ [183, 249]. The random error ϵ was simulated from skew-normal distribution with means equal to zero and and variance equal to $1 - R^2_G - R^2_X - R^2_{G\times X}$. The

error was simulated with shape parameters $\alpha_\epsilon = 0$ and 20 to denote symmetric and asymetric distributions respectively. The treatment of skewness using rank-based inverse-normal transformation were also assessed. The variance explained by the interacting variable X was fixed at $R^2_X = 24\%$ for all simulations while R^2_G and $R^2_{G\times X}$ were varied from 0 to 0.4%. A total of $R = 10,000$ replications were performed for all simulated datasets with each having a sample size of $n = 10,000$ independent observations. Both sample size and percent variance explained by genetic variants are within the realistic contexts of GWAS and GWIS study designs [260]. Large sample sizes are often required due to severe multiple testing corrections.

Tests Statistics

Variance Heterogeneity Tests for equality in variance by genotype included Levene's mean-based F-test, and the Brown-Forsythe F-test [247]. Both test statistics are given by

$$T^2 = \frac{(n-k)\sum_{i=1}^{k} n_i(\bar{Z}_{i.} - \bar{Z}_{..})^2}{(k-1)\sum_{i=1}^{k}\sum_{j=1}^{n_i}(\bar{Z}_{ij} - \bar{Z}_{i.})^2} \tag{4.9}$$

where $n_1, ..., n_i$ are the sample size of the i-th group, y_{ij} is the j-th observation of the i-th group and $Z_{ij} = |y_{ij} - \bar{y}_{i.}|$. The global and subgroup location summary statistics correspond to the mean in the case of Levene's F-test, or the median for Brown-Forsythe's F-test . Both test statistics follow an F distribution under the null hypothesis of variance homogeneity with $df_1 = k-1$ and $df_2 = n-k$, where n is the sample size and k is the number of subgroups.

Heterogeneity of QR Estimates The two other test statistics for potential interactions include Z_τ from CQR and UQR. The process of computing Z_τ for UQR is different from CQR. UQR requires adjustment for the main effect of genotypes where Z_τ is computed on residuals of Y regressed against G. To explain why, consider take the earlier example in equation 4.5 and consider the special case where $\beta_0 = \beta_1 = 0$, $\beta_2 \neq 0$ and $\beta_3 \neq 0$. The implication of the heteroscedasticity function $\sigma_x(g) = \beta_3 g$ on the conditional quantile

function of Y is

$$\begin{aligned}\Pr\big[Y > q_\tau \mid G = g\big] &= \Pr\big[g(\beta_2 + \beta_3\epsilon_x) + \epsilon) > q_\tau\big] \\ &= \Pr\big[\epsilon_x > \frac{q_\tau - \beta_2 g - \epsilon}{\beta_3 g}\big] \\ &= 1 - F_{\epsilon_x}\big(\frac{q_\tau - \beta_2 g - \epsilon}{\beta_3 g}\big)\end{aligned} \tag{4.10}$$

where a change in G corresponds to a change in both the numerator and denominator given the directions and magnitudes of β_2 and β_3. The presence of both marginal and interaction effects complicates the relationship between genotypes and the 'unconditional' quantiles of Y that UQR models. UQR utilizes observed sample quantiles as approximations to the unobserved unconditional quantiles of Y for re-centered influence function (RIF) transformations at a given vector of percentiles, $\boldsymbol{\tau}$. This is problematic because sample quantiles are 'contaminated' by g where the mixture of marginal and interaction effects compromise trends in $\boldsymbol{\beta_G(\tau)}$ with $\boldsymbol{\tau}$ (See Suppelmentary). This can be overcomed using the residuals of the response adjusted for genotype effect. The marginal effect of g is treated as a nuisance parameter using a two-step estimation approach through modeling quantiles of $Y^* = Y - g\widehat{\beta}(\tau = 0.5)$, where $\widehat{\beta}(\tau = 0.5)$ is the marginal effect estimated using UQR. Note that CQR models the conditional quantile function of Y, and hence, do not suffer from this limitation of UQR. Nonetheless, UQR is more practical in genome-wide analysis given its computational efficiency over CQR.

Moreover, choosing the appropriate number and range of $\boldsymbol{\tau}$ to consider for CQR and UQR depend on the sample size given that our framework assumes that $\boldsymbol{\beta_G(\tau)}$ converges to the true population estimates and are normally distributed. All sample quantiles converge to normality at different rates, with the median being the fastest, and extreme quantiles having the slowest convergence rate with sample size [261]. Deviations from normality will occur when quantiles at the extreme distribution tails are modeled using an insufficient sample size. To help guide the selection of the number and range of τ, a diagnostic for

plots for deviations from normality can be made by constructing the null distribution. A simple strategy is to apply permutation tests for all variants include the random generation of genotypes and fitting CQR and UQR on the response of interest for different number and range of percentiles with $R = 10,000$ replicates to compute Z_τ and assess Type I error rates assuming normality. Supplementary Figure B3 shows the convergence of Type I error rate to the nominal level of 0.05. The true distribution of Z_τ deviates from normality when modeling extreme distribution tails. On the other hand, increasing the number of percentiles over-parameterizes the covariance matrix and slightly increases the rate of false positives. Both CQR and UQR were fitted for 10 percentiles across $5\% \leq \tau \leq 95\%$ given the convergence of Type I error rate to nominal level for $n = 10,000$.

Type I Error

Type I error rates for individual tests were computed as the proportion of false positives of $R = 10,000$ replications at a nominal level of 0.05. False positive rates were assessed while varying the symmetry of error distribution (symmetric or asymmetric), and on rank-based inverse normal transformation of the asymmetric error (i.e. skewed response variable). R^2_G was varied between 0% and 0.4%. All p, R^2_X and $R^2_{G\times X}$ were fixed at 5%, 24% and 0%, respectively. The false positive rate of unadjusted UQR for marginal effects of G was also provided to demonstrate the effect of mixtures of marginal and interaction effects on trends of $\beta_G(\tau)$ with τ for UQR. A direct interaction test using CQR is also applied as a reference for other tests.

Power to Detect Associations

The power of each test for detecting potential interactions was computed as the proportion of replications that correctly reject the null hypothesis at a nominal significance level of $\alpha = 0.05$. Power was computed to assess the impact of 1) the number of genotype levels, 2) skewness, and 3) antagonistic interactions on the ability of variance heterogeneity tests

(Levene's F-test, and BF test) and MQR-based tests (MCQR, MUQR) to detect a single unadjusted two-way interaction. The number of genotype levels was varied from 2 (i.e. single chromosome) to 5 levels (i.e. 4 chromosomes) where allele frequency is varied from 0.05 to 0.95. The effect of skewness on power was assessed by varying error distribution (symmetric, or asymmetric), and $R^2_{G\times X}$ (0% to 0.4%). The effect of antagonistic interaction effects was assessed similarly, but with an interaction coefficient having opposite direction compared to main effect of the interacting variable. The main effect of the interacting variable is set to be positive, while the coefficient for interaction effect is negative to correspond to antagonistic interaction effects.

Results

False Positives

Figure 4.1 shows Type I error rates for test statistics under symmetric and asymmetric errors in addition to rank-transformation to treat skewed response variables when $R^2_G = 0\%$ and 04%. Levene's F-test showed an increased Type I error rate for asymmetric error in comparison to BF F-test that was near nominal levels in all scenarios. MCQR showed a slightly elevated Type I error rates that were stable across the scenarios investigated. This results from the range and the number of percentiles considered which produced slight deviations from normality (Figure B3). The variance-covariance matrix for CQR estimates have a slow asymptotic convergence rate and require bootstrap methods to properly estimate [252]. Nonetheless, CQR was fitted for the same range and number of percentiles as UQR for valid comparisons. In addition, MUQR unadjusted for genotype effects had an increased Type 1 error rate resulting from marginal and interaction effects of G as discussed earlier. Adjusting for the marginal genotype effect resulted in a near nominal level of false positives for UQR. Lastly, rank-based inverse-normal transformation on skewed response variables where variants include a marginal effect at $R^2_G = 0.4\%$ have resulted in increase Type I error

rates for all reference CQR $G \times E$ interaction test, variance heterogeneity tests, and MQR tests.

Power Comparisons

By Group Levels Figures 4.2 and B4 show the power of detecting potential interactions with increasing number of group levels (e.g. mono, bi, tri, and quad allelic genotypes). MCQR and MUQR show higher power compared to variance heterogeneity tests when genotypes include more than two genotype levels. The difference in power is even greater for 4 and 5 group levels (e.g. tri and quad allelic genotypes) as shown in the Supplementary Figure B4. This can be explained by the fact that variance heterogeneity relies on subgroup analysis for variance per-genotype whereas MCQR and MUQR utilize the whole sample to compute Z_τ. The power remains constant across allele frequency due to the change in interaction effects with allele frequency to maintain the same variance explained by interaction effects (at $R^2_{G\times X} = 0.1\%$).

By Distribution Shape and Type of Interaction Figure 4.3 shows the power of detecting potential synergistic and antagonistic interactions under both symmetric ($\alpha_\epsilon = 0$) and asymmetric ($\alpha_\epsilon = 20$) error distributions. Genotypes correspond to allele copies from two chromosomes (i.e. 3 levels) under an additive genetic model effect. Both MCQR and MUQR were found to have more power of detecting unadjusted two-way interactions under symmetric error distribution compared to variance heterogeneity tests. This is due to the effect of genotype group levels. However, this difference in power increased greatly under an asymmetric error distribution. This is due to the sensitivity of variance estimates to skewness that increases their sampling error for variance estimates for skewed distributions [262]. In contrast, MCQR and MUQR, on the other hand, are not affected by skewness as much given that they rely on distribution quantiles instead. Moreover, the power of all four tests under antagonistic interactions have decreased compared to synergistic interactions.

However, the difference in power between MQR based tests and variance heterogeneity tests have increased for antagonistic compared to synergistic interactions. This difference is even larger under asymmetric error distributions. Note that treatment of skewness using a rank-based inverse-normal transformation was not applied since it increases in Type I error rates of all reference $G \times E$ interaction test and potential interaction tests (Levene, BF, MQR, and MUQR) for genotypes with main effects.

Discussion

This study expands the application of MQR to detect evidence of potential interactions in the context of genetic association analysis. We show that using MR to model heterogeneity in QR estimates across an outcome distribution is a robust and powerful approach for detecting potential interactions. We show that MUQR overcomes the computational limitation of MCQR. UQR was 2.5-8.5 times faster than CQR via Frisch-Newton approach after preprocessing. Hence, MUQR is more practical for large-scale genome-wide association analyses. MQR methods were also shown to maintain nominal Type I error rate and achieved greater power when compared to variance heterogeneity tests with increasing number of group levels, asymmetric error distributions, and antagonistic interactions.

This study shows how the shape of QR estimates captures information on the quantile density function of interaction variable(s) (Figure B1). Fitting MR models on QR estimates using re-centered percentiles around the median allows for estimation of 1) an intercept parameter that corresponds to the marginal effect of genotypes, and 2) a slope parameter that corresponds to the mean change in QR estimates with one-unit change in percentiles and denotes the presence of potential interactions. The effect of SNPs across the sample distribution is a function of error quantiles that assume the same density of the interacting variable(s) (Equation 4.5). In the case of a single interaction variable, the quantile function of a univariable error density is increasing by definition. In the case of multiple interacting

variables, the quantile function of a multivariate error density is equivariant and thus not strictly increasing in a linear fashion (Supplementary Material) [263, 264]. Identifying the best model to fit QR estimates across percentiles can, however, be challenging in the context of genome-wide analysis. Assuming a linear trend provides a simple analog of correlation, which implies associations while acknowledging that more complex relationships (i.e. multiple interacting variables with certain types of non-increasing quantile functions) may not be detected given the limited robustness of linear models for detecting non-linear relationships. The approach follows merely the natural association analysis using linear regression in which correlation implies association but not causation, while acknowledging that causation implies association but not correlation [265].

Heteroscedasticity resulting from gene interactions have been characterized previously, where testing for differences in variance per genotype was proposed as an indicator of potential interactions [247, 183]. Two well-known tests for variance heterogeneity include Levene's F-test and the Brown and Forsythe test [266, 267]. Brown and Forsythe highlighted that Levene's statistic is not robust when the underlying population distribution is skewed and proposed additional treatments to address this problem [267]. However, variance as a measure of distribution spread is less informative when applied to skewed phenotypes. As such, variance heterogeneity tests are prone to higher Type I error rates (in the case of Levene's F-test) and have decreased power of detecting differences in distribution spread for skewed distributions as confirmed by Figure 4.3. This is because the confidence intervals for variance estimates are larger for skewed distributions [262]. Instead, modeling distribution quantiles using QR is more suitable for testing differences in the distribution spread for asymmetric phenotypes by genotype levels. Many biological traits are skewed including those that are harmful at low quantities but tolerated at high quantities (i.e. right skewed blood glucose measures, or BMI) or vice versa.

Location-scale-shift models of QR estimates to assess differences in distribution by regressors have been previously proposed [268, 269, 270]. These approaches enable inference on

global differences between distributions by genotype but do not compare location and scale shifts separately. This problem is shared by other non-parametric tests such as Kolmogorov-Smirnov test and other derivatives [271, 272]. Current quantile-based tests rely on computed reference null distributions (i.e. permutations) that provide p-values with 4 or less significant digits or, else, require approximations of extreme distribution tails [273]. To our knowledge, there are currently no quantile-based tests that exclusively focus on differences in scale shifts. Hence, MQR as first utilized by Abadi et al. 2017 and elaborated further here in this study, is the first to reliably and efficiently enable robust statistical inference on scale differences [208]. It allows for the objective modeling of scale change with genotypes (i.e. linear heteroscedasticity due to one or multiple two-way interactions) as well as more complex terms. Its utility may not only be limited to genetics but to research at large as it provides estimates of both marginal and potential interaction effects. MQR is not restricted to genetic epidemiology alone in the pursuit of precision medicine but applies to all research fields alike.

This study includes simulations to compare the power of the proposed approach against tests for variance heterogeneity by genotype with 1) number of group levels, and 2) symmetric and asymmetric error distributions, synergistic and antagonistic interactions. The simulation results show that MQR-based tests are robust to skewness and maintain nominal Type I error rates under asymmetric, symmetric error distributions. Treating distribution skewness using inverse-normal rank-transformation was found to inflate Type I error for genotypes with main effects. Furthermore, MCQR and MUQR were shown to have higher power of detecting potential interactions compared to variance heterogeneity tests for genotypes with 3 or more group levels. Variance heterogeneity tests require sub-group estimates that become less reliable with increasing group levels for the same sample size. On the contrary, MCQR and MUQR utilize the whole sample and are not affected by the number of factor levels. More importantly, variance heterogeneity tests are limited to the analysis of factors, while MQR-based tests can be applied to both factors and continuous variables alike. Furthermore, both MCQR and MUQR were shown to have a higher power of detecting potential interactions

under skewed phenotype by a large margin (Figures 4.3). This is because the skewness increase sampling error of variance estimates. Furthermore, both MCQR and MUQR were found have a higher power of detecting antagonistic interactions. Note that the conditional variance given in equation 4.3 is a sum of weighted variance components given genotypes. The observed conditional variance is diluted by error variance where QR differentiates the change in QR estimates exclusively to change in quantiles of partial errors due to interactions from quantiles of the random error as in equation 4.5. Altogether, the proposed approach handles scale-shifts per genotypes more efficiently than tests of variance heterogeneity for skewed phenotypes and for antagonistic interactions effects.

This study extends the utility of QR by developing the formal connection between the density of interacting variables with QR estimates and providing tests of associations between QR estimates and percentiles with direction and magnitude. This approach, however, is not without limitations. The modeling of QR estimates using CQR, as previously proposed in Abadi et al. 2017, is computationally intensive [208]. Details on the computational challenges for CQR can be found in Chen et al [274]. Bootstrap methods are required to reliably estimate variance-covariance matrix of CQR estimate. CQR, as an optimization problem, is computationally infeasible in the context of genome-wide analysis. The recent development of UQR by Firpo et al 2009 provides a window of opportunity for a scalable approximation of univariable CQR estimates in the context of genome-wide analysis. A brief description of the differences between CQR and UQR are presented in the supplementary material. Both UQR parameters and their corresponding variance-covariance matrix have a closed-form solution and are, therefore, easily computed compared to CQR which requires bootstrap methods for proper estimation due to slow asymptotic convergence [252]. An assessment of CPU time required to for estimating QR parameters alone shows that UQR scales well with the number of SNPs, sample size, number of percentiles, and number of covariates compared to all optimization algorithms for CQR (Figure B2) [273, 275, 251]. However, UQR relies on kernel density estimates and may require careful considerations for highly sparse or bounded

distributions [276]. Hence, the computational advantages of UQR or CQR for sparse phenotypes are limited by proper kernel density estimations. Moreover, inference for potential interactions is based on the asymptotic results of CQR and UQR estimates. The asymptotic normality of our test statistic is jeopardized by the use of insufficient sample sizes and the modeling of extreme quantiles. Violations in the normality assumption lead to an increase in Type I error rates. The number and range of percentiles to choose from depends on the research's objective and sample size, and population representation. Nonetheless, a diagnostic plot is provided for assessing the normality assumption and overparameterization issues for estimating the covariance-variance matrix of CQR and UQR estimates given the range and number of percentiles to fit. We confirm that CQR under the naive i.i.d assumption, requires large sample sizes (i.e. $n \geq 1000$) to correctly estimate the asymptotic variance-covariance matrix. Hence, bootstrap methods are essential to appropriately compute the variance-covariance matrix of CQR estimates for proper inference.

In conclusion, MQR methods for detecting scale-shifts by genotypes are more purposeful compared to variance heterogeneity tests. While MCQR is computationally expensive, MUQR achieves similar desired power and Type I error rates without the computational overhead. Hence, MUQR could be utilized in large-scale genome-wide analysis for clinically relevant phenotypes to identify variants with potential interactions.

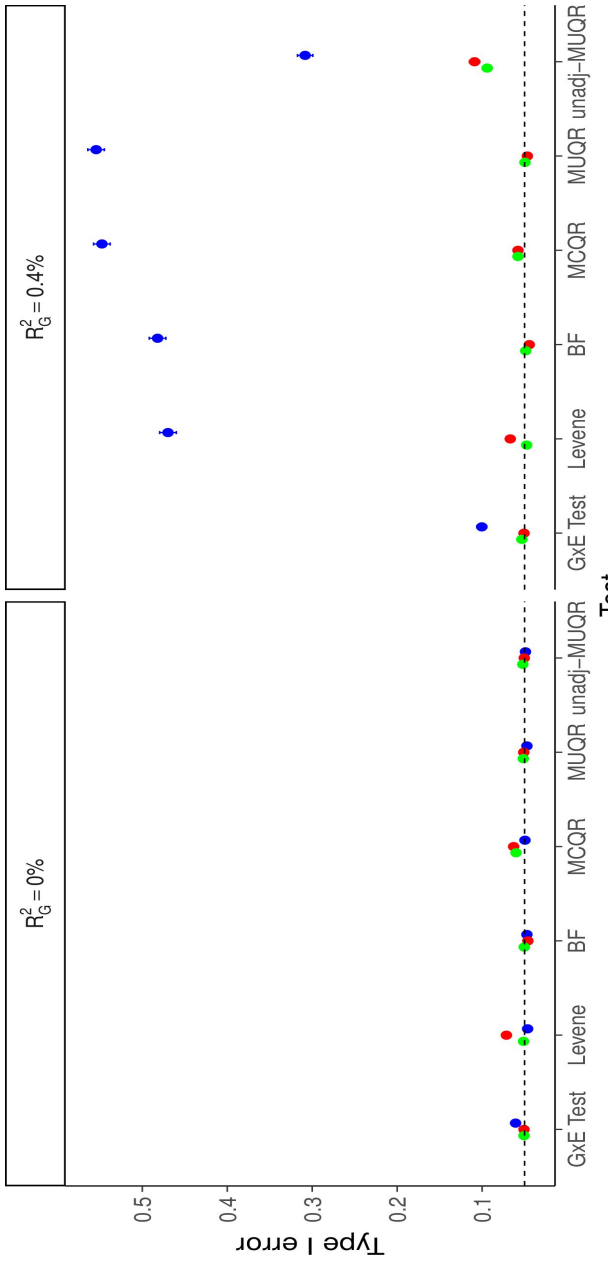


Figure 4.1: Type I error rates for test statistics of potential interactions. The error bars corresponds to the binomial confidence interval for $R = 10,000$ replications. R^2_G corresponds to the variance explained by the genetic variant's main effect. The variance explained by the interacting variable is fixed at 24%. The $G \times E$ test corresponds to the reference direct interaction using median CQR. Levene and BF correspond to the variance heterogeneity tests by Levene and the Brown-Forsythe F-tests respectively. Unadj-MUQR and MUQR corresponds to UQR models fitted on the raw and residuals scale of the response variable

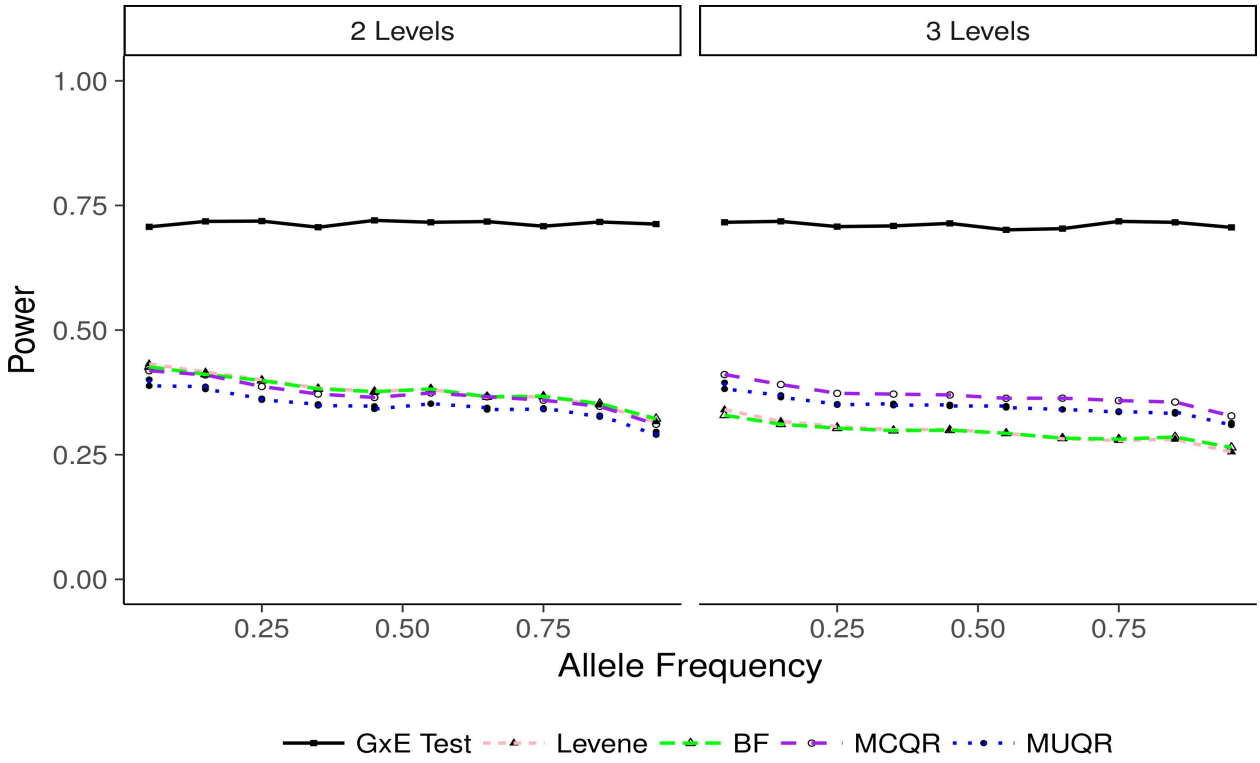


Figure 4.2: Power of detecting interaction effects for 2 and 3 genotype group levels

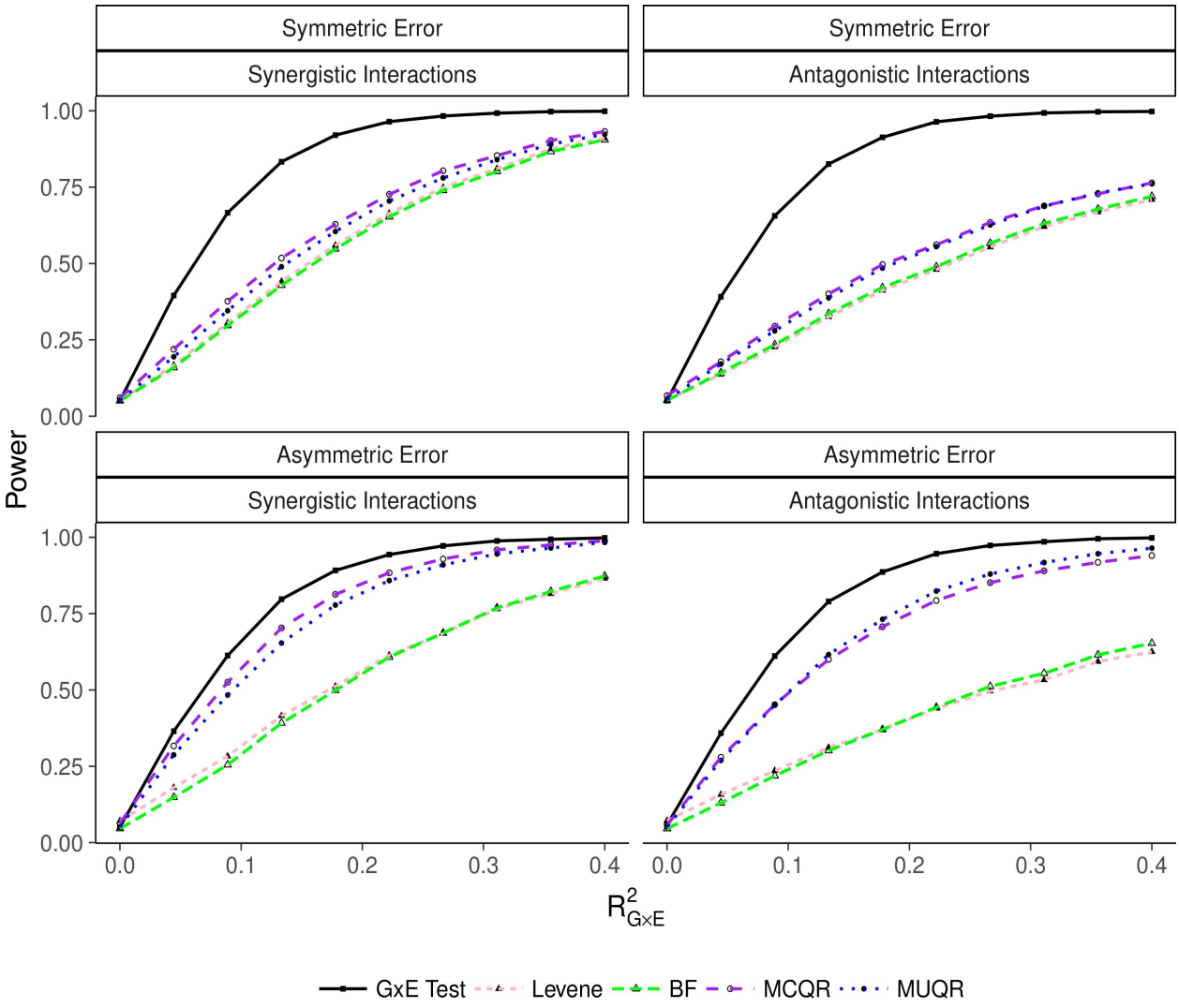


Figure 4.3: Power of detecting synergistic interaction effects. Power is presented for when the error distribution is symmetric ($\alpha_\epsilon = 0$) or asymmetric ($\alpha_\epsilon = 20$). The interacting variable is simulated from a standard normal relevance variance explained fixed at $R^2_X = 24\%$. MAF was fixed at 5% as there are no to minor differences in power due to re-adjustment of interaction effects to keep $R_{G\times E}$ fixed as intended.

Chapter 5

Conclusion

This thesis provides a review on the challenges and opportunities for personalized medicine. It tackles the growing gaps in healthcare quality between developed and undeveloped nations that jeopardize social stability and ethical standards. It tackles the growing gap in our ability to generate data compared to our ability to analyze and extract useful reproducible associations. There is a need for global investment in basic research that include 1) cost effect generation of high-quality high-throughput data, 2) hybrid education and multidisciplinary teams, 3) data storage and processing, 4) data integration and interpretation, and 5) individual and global economic relevance; to be followed by global investments into public health to adopt routine personalized medicine. This is, however, not the case given the recent funding cuts for the National Institutes of Health (NIH) budget. The review further identifies that the lack of robust statistical methods as the primary bottleneck and challenge in basic research towards personalized medicine. It highlights that unknown or unadjusted interactions result in true differences in the marginal effects which are only mitigated through adjustments or standardizations that limit generalizability. Hence, while increasing the sample size of observational studies may help reduce the effect of random error on the precision of estimates, systematic heterogeneity requires careful treatments in order to improve model accuracy and reliability. To this end, the heterogeneity of the marginal effects of known

BMI variants were investigated in the third chapter. A novel statistical framework based on Quantile regression, and Meta-regression was developed to identify variants with true heterogeneity that indicate the presence of potential interactions, termed, Meta-quantile Regression (MQR).

In chapter 3, the genetic architecture of genetic effects on BMI was shown to have a strong evidence of potential interactions. In addition, all *FTO*, *PCSK1*, *TCF7L2*, *MC4R*, *FANCL*, *GIPR*, *MAKP2K5*, and *NT5C2* were found to have potential interactions. On the contrary, height was found to have no evidence of a genetic architecture that includes environmental interactions. The analyses performed on height served as a control to BMI results given that height is a highly heritable trait that is fixed during adulthood. However, all SNP analyses were formed assuming an additive effect and no assessments on deviation from this assumption were performed. In addition, there is a notable relationship between the marginal and MQR effect sizes for BMI variants, which require further investigations.

In chapter 4, the computational cost and slow asymptotic convergence rates limitations of MQR based on conditional quantile regression were discussed, and overcome using unconditional quantile regression in the fourth chapter. MQR was also found to have higher power of detecting potential interactions compared to the variance heterogeneity tests for asymmetric population distributions and antagonistic interaction effects while maintaining nominal false positive rates. MQR is, however, not without limitations. It assumes a linear quantile-widening effect of variables, which is simple, but may not capture complex scale effects. There is also a need to emphasize that potential interactions may not reflect true interactions with other variables. Interaction effects simply reflect deviation from additive effects, and this can occur as a result of nonlinear effects. This also includes gene-gene interactions that may not involve environmental factors. Furthermore, it is not sufficient to simply identify potential interactions in the context of personalized medicine. Here is a need to show the nature of interactions. There is a need to further extend the utility of MQR to help identify and characterize interactions.

Although it is difficult to predict the future, it is helpful to consider emerging trends together with the current challenges to foresee the new opportunities. In the coming years, it is likely that the challenges of homogenous and heterogeneous sample population recruitment,deep phenotyping, and model developments will be key objectives towards personalized medicine. As such, it is likely that 1) individuals will be more precisely characterized with increasing precision, 2) study populations will grow to allow millions of samples for observational study designs, 3) new statistical method will be developed to discover and reproduce relationships from these data, 4) far more complex diagnostic and prognostic categories that are currently in use will arise using multidimensional characterization of patients, and 5) analytical and algorithmic models will be useful for clinical purposes even when they defy easy summary in language to most clinicians [277].

Appendix A

Supplemental Note

Analytical Description: Ordinary least squares (OLS) regression is the classic method to estimate mean effects of SNPs on a quantitative trait. OLS models are particularly useful when the assumptions of linearity, normality, and homoscedasticity are met, but otherwise, require proper corrections in order to allow unbiased parameter estimation and valid inference. These models are developed on the basis of true fixed effects and do not capture true variability in the effects of genetic risk factors in the presence of single and mixed gene-environment (G×E) and gene-gene (G×G) interactions. If such interactions are unadjusted, OLS models will produce estimates with limited reproducibility that depend on the context of the sample population and the degree of exposure to interacting variables (e.g. environmental exposure).[1] Reproducibility is a well-known problem in genetic epidemiology for complex phenotypes that involve interactions.[2] Alternatively, GWAS may use case/control designs to compare BMI categories, where binary logistic regression is used to estimate the fixed effects of SNPs on the probability of belonging to either of two factor levels (e.g. normal-weight vs. obesity subgroups). However, subgroup analysis not only reduces statistical power due to loss of sample size and uneven group levels, but also limits interpretation to pair-wise comparisons. In addition, logistic regression profiles pre-selected segments of the BMI distribution, which can be problematic to assign a priori.

Conditional quantile regression (CQR) is an alternative regression technique that permits

the assessment of associations at the full scope of the outcome distribution by examining the effects of regressors at a series of quantiles of the outcome distribution without dividing the sample into subgroups.[3,4] CQR models the effects of a change in one unit of a predictor on the position of a given quantile of the outcome. It also utilizes the entire data set for parameter estimation, confidence interval construction and hypothesis testing regardless of the specified quantiles and does not suffer the statistical limitations of subgroup analysis. This regression framework has recently gained traction in clinical epidemiology to generate fetal, childhood and adolescent growth curves.[5-7] Recent reports have highlighted the potential applications of CQR in genetic epidemiology.[8-10] To our knowledge, CQR has not been applied to model the variability in effect size estimates along the sample outcome distribution in the presence of single and mixed G×E and G×G interactions.

Variations in effect size estimates due to unadjusted interactions can be modelled using CQR as a re-formulation of heteroscedastic OLS models.[3,11,12] Lets consider a sample of n independent and identical distributed (i.i.d) variables $Y_1, ..., Y_n$ with *cdf* $F_Y(y)$, where $y_1, ..., y_n$ are their respective observed values. Lets also assume they follow a linear relationship with an interaction term given as

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 g_i + \beta_3 x_i g_i + \epsilon_i \tag{5.1}$$

where x_i corresponds to the unknown/unmeasured interacting variable, $X \sim F_X(\mu_x, \sigma_x^2)$; g_i is the observed genotype of the genetic variant G under Hardy-Weinberg equilibrium (HWE) with a population allele frequency, p, where $G \sim B(2, p)$; and ϵ_i is the random error with $\epsilon \sim F_\epsilon(0, \sigma_\epsilon)$. The coefficients β_0, β_1, β_2 and β_3 represent the intercept, the marginal effect of X, the marginal effect of G and the interaction effect of G and X, respectively. The conditional distribution of the response variable Y can be described as $F_{Y|X=x,G=g} \sim \left(\beta_0 + \beta_1 x + \beta_2 g + \beta_3 xg, \sigma_y^2\right)$. If the interacting variable, X, is not adjusted

then the conditional density of Y given G can be shown to have a mean and variance:

$$E\left[Y \mid G = g\right] = (\beta_0 + \beta_1\mu_x) + (\beta_2 + \beta_3\mu_x)g \tag{5.2}$$

$$Var(Y \mid G = g) = \sigma_x^2(\beta_1 + \beta_3 g)^2 + \sigma_\epsilon^2 \tag{5.3}$$

The resulting conditional distribution $F_{Y|G}(y \mid G = g)$ simply translates to a heteroscedastic linear model with partitioned residuals where $\sigma(g) = (\beta_1 + \beta_3 g)$. That is

$$y_i = \big(\beta_0 + \beta_1\mu_x\big) + \big(\beta_2 + \beta_3\mu_x\big)g_i + \big(\beta_1 + \beta_3 g_i\big)\epsilon_{i,1} + \epsilon_{i,2} \tag{5.4}$$

where $\epsilon_1 \sim F_{\epsilon_1}(0, \sigma_x^2)$ and $\epsilon_2 \sim F_{\epsilon_2}(0, \sigma_\epsilon^2)$. The conditional quantile function for the heteroscedastic model under i.i.d errors is

$$\begin{aligned} Q_Y(\tau \mid G = g) &= \big[\beta_0 + \beta_1\mu_x + \beta_1 Q_{\epsilon_1}(\tau) + Q_{\epsilon_2}(\tau)\big] + g\big[\beta_2 + \beta_3\mu_x + \beta_3 Q_{\epsilon_1}(\tau)\big] \\ &= \beta_0^*(\tau) + \beta_1^*(\tau)g \end{aligned} \tag{5.5}$$

which is a CQR model with the true fixed parameters $\beta_0^*(\tau)$ and $\beta_1^*(\tau)$. τ can be any quantile of the sample outcome distribution of Y. This formulation can be generalized further for a set of k independent interacting variables in matrix form as

$$Q_Y(\tau \mid G = g) = A\beta(\tau) \tag{5.6}$$

where $A \in \mathbb{R}^{n\times 2}$ is the design matrix $\left[1, G^{'}\right]$ and

$$\beta(\tau) = \begin{pmatrix} \beta_0 + \sum_{j=1}^{k} \beta_{x_j}\mu_j \\ \beta_g + \sum_{j=1}^{k} \beta_{int_j}\mu_j \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^{k} \beta_{x_j} Q_{\epsilon_j}(\tau) + Q_{\epsilon_{k+1}}(\tau) \\ \sum_{j=1}^{k} \beta_{int_j} Q_{\epsilon_j}(\tau) \end{pmatrix} \tag{5.7}$$

Here, β_g and β_{x_j} are the main effects of the genetic variant and $j \in 1, .., j$ are the

unknown/unmeasured variables with their respective interaction coefficients β_{int_j}. The cumulative two-way interactions of k variables results in a linear function with τ as a result of the symmetric heteroscedasticity function $\mathbf{1}'\boldsymbol{\gamma}$ where $\mathbf{1} \in \mathbb{R}^{k\times 1}$ and γ has elements $\sigma_j(g) = \beta_{x_j} + \beta_{int_j} g$. Under an additive genetic model, the main effect of the genetic variant $\beta(\tau)$ is a fixed constant for all $\boldsymbol{\tau} \in \tau_1, ..., \tau_m$ if and only if all interacting effects are zero, i.e. $\beta_{int_j} = 0$. It is possible to further break down the independence assumption between interacting variables using a variance-covariance matrix of partial errors, but the above formulation serves as a simple analytical demonstration for the use of CQR in modelling unadjusted interactions. A linear trend of estimates with τ corresponds to cumulative two-way interactions, while quadratic curves supports complex higher order interactions. Hence, the association of genetic variants under unadjusted interacting variables simply reduces to the modelling of CQR estimates along the distribution of the outcome at $\boldsymbol{\tau} \in \tau_1, ..., \tau_m$.

This is accomplished by using meta-regression (MR) to model the heterogeneity of CQR estimates across the sample outcome distribution and estimate the change in CQR estimates with $\boldsymbol{\tau}$.[13,14] That is, fitting the MR model

$$\beta(\tau) = \begin{pmatrix} \mathbf{1}' & \boldsymbol{\tau}' - 0.5 \end{pmatrix} \begin{pmatrix} \beta_m \\ \beta_\tau \end{pmatrix} + \epsilon \tag{5.8}$$

where β_m is the median effect of the genetic variant, β_τ is the slope coefficient for the change in the median effect with τ, and $\epsilon \in \mathbb{R}^{n\times 1}$ are random errors with the cross-quantile variance-covariance matrix of the estimates under i.i.d errors. This framework provides both location-shift and change in location-shift estimates to further decipher the nature of complex genetic associations.

Simulations: The power to detect unadjusted interactions using CQR and MR was explored using simulations. Equation 1 describes the effects of an interaction between a SNP, G, and a variable, X, on a quantitative trait, Y. Without loss of generality, G was assumed to be biallellic with a MAF, p, under HWE and an additive genetic effect on Y.

Moreover, G was encoded such that mean genotype was zero ($-2p$, $1-2p$, or $2-2p$).[11] The total variance of Y was assumed to be 1 and the variance of each component of equation 1 was partitioned accordingly. Specifically, the proportion of the variance (R^2) of Y that was explained by G, X and the interaction between G and X was $R^2_G = 2p(1-p)\beta_2^2$, $R^2_X = \beta_1^2$ and $R^2_{G\times X} = 2p(1-p)\beta_3^2$, respectively. The error term, ϵ, was assumed to have a normal distribution with a mean of 0 and a variance of $1 - R^2_G - R^2_X - R^2_{G\times X}$. Unless otherwise specified, the simulation conditions were MAF = 0.2, N = 10,000, $R^2_G = 0.004$, $R^2_X = 0.25$, and $R^2_{G\times X}$ was varied between 0 and 0.004. When more than one interaction was considered, R^2_X was divided equally between all interaction covariates, while each additional interaction was equal to $R^2_{G\times X}$. All regression models were fitted with Y as the dependent variable and G as the independent variable. CQR models were fitted at every 10^{th} percentile of the distribution of Y from the 5^{th} to the 95^{th} percentiles. A total of 1,000 Markov chain marginal bootstrap (MCMB) replicates were used to compute confidence intervals and the cross-percentile variance-covariance matrix for CQR estimates.[12,15,16] Variability in the CQR estimates of G at these percentiles was modelled using MR, assuming normality, to determine the effects of percentiles on mean CQR estimates. The power to detect interactions at a threshold of $p < 0.05$ was computed from 1,000 replicates of each simulation condition.

Sample Stratification and Interactions: The analysis of secondary traits (e.g. BMI) collected from case-control studies with disease status (e.g. T2D) as a primary outcome can be prone to artifacts if potential stratification of secondary traits is not addressed.[17] This stems from the fact that secondary traits are often strong risk factors for disease status and can thus be stratified in cases and controls. Since effect alleles of disease-associated SNPs are typically enriched in cases and depleted in controls, the stratification of allele frequencies and secondary traits can correspond. The coinciding stratification of secondary trait distributions and allele frequency distributions may result in spurious associations between these disease-associated SNPs and secondary traits. This phenomenon has also been observed in population-based designs when disease prevalence differs between the sample and general

populations.[18] Yaghootkar, et al., have recently developed an analytical model relating regression estimate bias to differences between disease prevalence in the sample and general populations.[18] This model described regression estimate bias in the main effects of SNPs as a function of the partitioning of allele frequencies by disease status as well as the partitioning of variance by genotype (i.e. heteroscedasticity). They also extended this description to include regression models fitted with adjustment for disease status and show that the bias persists even after this adjustment.[18] Importantly, when regression models are adjusted for disease status the bias in regression estimates is *not* a function of the partitioning partitioning of variance by genotype.[18] This is critical because it means that while estimates of the main effects of SNPs from CQR models may be affected by sample stratification in the same way as estimates from OLS models, the variation of CQR estimates across the sample distribution is not a function of differences in disease prevalence between sample and general populations. The analytical model presented here is not primarily concerned with main effects of SNPs on continuous outcomes, rather with modelling the *variation* of CQR estimates across the sample outcome distribution.

The effect of sample stratification on the power to detect of unadjusted gene interactions with CQR and MR was assessed in simulations. Consider the disease outcome (Z), the continuous risk factor (Y) and the SNP (G), whose relationship is described using a liability scale disease (probit) model.[18]

$$z_i = \beta_4 g_i + \beta_5 y_i + \varphi_i \tag{5.9}$$

where the coefficients β_4 and β_5 represent the respective marginal effects of G and Y on Z, φ_i is the random error with $\varphi \sim F_\varphi(0, \sigma_\varphi)$, and y_i is specified in equation 1. Disease status (D) is defined as follows;

$$\alpha = \Phi^{-1}(1 - \pi_0) \tag{5.10}$$

$$D = \begin{cases} 1 & \text{if } z_i > \alpha \\ 0 & \text{if } z_i \leq \alpha \end{cases} \tag{5.11}$$

where π_0 is the disease prevalence in the general population. Figure S2A shows a schematic representation of this model. A population of 100,000 individuals was simulated with the following conditions; $\pi_0 = 0.1$ (i.e. population disease prevalence of 10%), MAF = 0.2, $R^2_{G[Y]} = 0.004$, $R^2_X = 0.25$, $R^2_{G[Z]} = 0.01$ (equivalent to OR $\sim$ 1.4 for G on D), R^2_Y = 0.20 (equivalent to OR $\sim$ 2.5 for Y on D) and $R^2_{G\times X[Y]}$ was varied between 0 and 0.004. A random sample of N = 10,000 individuals was then drawn from this population with pre-specified proportion of cases (5, 10, 25 and 50%) and then disease adjusted CQR models ($y \sim g + D$) were fitted across the distribution of Y as in simulations above. Variability in the CQR estimates of G at these percentiles was modelled using MR to determine the effects of percentiles on mean CQR estimates. The power to detect interactions at a threshold of p < 0.05 was computed from 1,000 replicates of each simulation condition.

Supplemental References

1. Begum, F., Ghosh, D., Tseng, G.C., and Feingold, E. (2012). Comprehensive literature review and statistical considerations for GWAS meta-analysis. Nucleic Acids Res 40, 37773784.

2. Li, A., and Meyre, D. (2013). Challenges in reproducibility of genetic association studies: lessons learned from the obesity field. Int J Obes (Lond) 37, 559567.

3. Koenker, R. (2005). Quantile Regression (Cambridge University Press).

4. Koenker, R., and Hallock, K. (2001). Quantile regression: An introduction. Journal of Economic Perspectives.

5. Kiserud, T., Piaggio, G., Carroli, G., Widmer, M., Carvalho, J., Neerup Jensen, L., Giordano, D., Cecatti, J.G., Abdel Aleem, H., Talegawkar, S.A., et al. (2017). The World Health Organization Fetal Growth Charts: A Multinational Longitudinal Study of Ultrasound Biometric Measurements and Estimated Fetal Weight. PLoS Med 14, e1002220.

6. Fernndez, J.R., Redden, D.T., Pietrobelli, A., and Allison, D.B. (2004). Waist circumference percentiles in nationally representative samples of African-American, European-American, and Mexican-American children and adolescents. J Pediatr 145, 439444.

7. Wei, Y., Pere, A., Koenker, R., and He, X. (2006). Quantile regression methods for reference growth charts. Statist. Med. 25, 13691382.

8. Beyerlein, A. (2014). Quantile Regression–Opportunities and Challenges From a User's Perspective. American Journal of Epidemiology 180, 330331.

9. Beyerlein, A., Kries, von, R., Ness, A.R., and Ong, K.K. (2011). Genetic markers of obesity risk: stronger associations with body composition in overweight compared to normal-weight children. PLoS ONE 6, e19057.

10. Mitchell, J.A., Hakonarson, H., Rebbeck, T.R., and Grant, S.F.A. (2013). Obesity-susceptibility loci and the tails of the pediatric BMI distribution. Obesity 21, 12561260.

11. Par, G., Cook, N.R., Ridker, P.M., and Chasman, D.I. (2010). On the use of variance per genotype as a tool to identify quantitative trait interaction effects: a report from the

Women's Genome Health Study. PLoS Genet 6, e1000981.

12. Koenker, R., and Bassett, G., Jr (1982). Robust tests for heteroscedasticity based on regression quantiles. Econometrica 50, 4361.

13. Thompson, S.G., and Higgins, J.P.T. (2002). How should meta-regression analyses be undertaken and interpreted? Statist. Med. 21, 15591573.

14. Borenstein, M., Hedges, L.V., Higgins, J.P.T., and Rothstein, H.R. (2009). Meta-Regression, in Introduction to Meta-Analysis (Chichester, UK: John Wiley & Sons, Ltd).

15. He, X., and Hu, F. (2002). Markov chain marginal bootstrap. Journal of the American Statistical Association 97, 783795.

16. Kocherginsky, M., He, X., and Mu, Y. (2005). Practical Confidence Intervals for Regression Quantiles. Journal of Computational and Graphical Statistics 14, 4155.

17. Monsees, G.M., Tamimi, R.M., and Kraft, P. (2009). Genome-wide association scans for secondary traits using case-control samples. Genet. Epidemiol. 33, 717728.

18. Yaghootkar, H., Bancks, M.P., Jones, S.E., McDaid, A., Beaumont, R., Donnelly, L., Wood, A.R., Campbell, A., Tyrrell, J., Hocking, L.J., et al. (2017). Quantifying the extent to which index event biases influence large genetic association studies. Hum Mol Genet 26, 10181030.

Supplemental Acknowledgements

ARIC (phs000280.v3.p1): The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C). Funding for CARe genotyping was provided by NHLBI Contract N01-HC-65226. Funding for GENEVA was provided by National Human Genome Research Institute grant U01HG004402 (E. Boerwinkle). Individual-participant phenotypic and genotypic data was extracted from the following dbGaP datasets (phs000280 .v3.p1, pht004062 .v1.p1, pht004063 .v1.p1, pht004064 .v1.p1, pht004065 .v1.p1, pht004032 .v1.p1, pht004033 .v1.p1, pht004034 .v1.p1, pht004035 .v1.p1, pht003266 .v1.p1, pht000114 .v2.p1, phg000079.v1). The authors would like to thank the participants, investigators and staff of the ARIC study for their important contributions.

CARDIA (phs000285.v3.p2): The Coronary Artery Risk Development in Young Adults Study (CARDIA) is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with the University of Alabama at Birmingham (N01-HC95095 and N01-HC48047), University of Minnesota (N01-HC48048), Northwestern University (N01-HC48049), and Kaiser Foundation Research Institute (N01-HC48050). This manuscript was not approved by CARDIA. The opinions and conclusions contained in this publication are solely those of the authors, and are not endorsed by CARDIA or the NHLBI and should not be assumed to reflect the opinions or conclusions of either. Funding for CARe genotyping was provided by NHLBI Contract N01-HC-65226. Individual-participant phenotypic and genotypic data was extracted from the following dbGaP datasets (phs000285 .v3.p2, pht001583.v2.p2, pht001632 .v2.p2, pht001671 .v2.p2, pht001720 .v2.p2, pht001785.v2.p2, pht001857 .v2.p2, pht001589 .v2.p2, pht001645 .v2.p2, pht001697.v2.p2, pht001744 .v2.p2, pht001804 .v2.p2, pht001851 .v2.p2, pht001557.v2.p2, pht001737 .v2.p2, pht001799 .v2.p2, pht001845 .v2.p2, pht001569.v2.p2, pht001601 .v2.p2, pht001656 .v2.p2, pht001715 .v2.p2, pht001761.v2.p2, pht001811 .v2.p2, pht001555 .v3.p2, phg000091 .v2.p1,

phg000092.v2.p1). The authors would like to thank the participants, investigators and staff of the CARDIA study for their important contributions.

CHS (phs000287.v5.p1): The research reported in this article was supported by contract numbers N01-HC-85079, N01-HC-85080, N01-HC-85081, N01-HC-85082, N01-HC-85083, N01-HC-85084, N01-HC-85085, N01-HC-85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133, N01-HC-85239 and HHSN268201200036C; grant numbers U01 HL080295 from the National Heart, Lung, and Blood Institute and R01 AG-023629 from the National Institute on Aging, with additional contribution from the National Institute of Neurological Disorders and Stroke. A full list of principal CHS investigators and institutions can be found at http://www.chs-nhlbi.org/pi.html. This manuscript was not prepared in collaboration with CHS investigators and does not necessarily reflect the opinions or views of CHS, or the NHLBI. Support for the genotyping through the CARe Study was provided by NHLBI Contract N01-HC-65226. Individual-participant phenotypic and genotypic data was extracted from the following dbGaP datasets (phs000287 .v5.p1, pht001452 .v1.p1, pht001488 .v1.p1, pht001489 .v1.p1, pht001490 .v1.p1, pht001491 .v2.p1, pht001492 .v1.p1, pht001493 .v1.p1, pht001494 .v1.p1, pht001495 .v1.p1, pht001474 .v1.p1, pht001475 .v1.p1, phg000077 .v1). The authors would like to thank the participants, investigators and staff of the CHS for their important contributions.

Framingham (phs000007.v28.p1): The Framingham Heart Study is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with Boston University (Contract No. N01-HC-25195 and HHSN268201500001I). This manuscript was not prepared in collaboration with investigators of the Framingham Heart Study and does not necessarily reflect the opinions or views of the Framingham Heart Study, Boston University, or NHLBI. Funding for CARe genotyping was provided by NHLBI Contract N01-HC-65226. Funding support for the Framingham Metabolomics - Risk Factor Study: Gas Chromatography/Mass Spec - BMI/Lipids/Gluc dataset was provided by FHS contract supplement, NHLBI Intramural Funding. Funding support for the Framingham Central

Metabolomics - HILIC - Installment 1 dataset was provided by NIH grant R01DK081572. Funding support for the Framingham Central Metabolomics - HILIC - Installment 2 dataset was provided by NIH grant R01 DK081572. Funding support for the Framingham Metabolomics (HILIC - Installment 1) dataset was provided by NIH grant R01 DK081572. Funding support for the Framingham Metabolomics (HILIC - Installment 2) dataset was provided by NIH grant R01 DK081572. Funding support for the Framingham Metabolomics (Hilic - installment 3) dataset was provided by NIH grant R01 DK081572. Individual-participant phenotypic and genotypic data was extracted from the following dbGaP datasets (phs000007.v28.p1, pht000009.v2.p10, pht000010.v3.p10, pht000011.v3.p10, pht000012.v3.p10, pht000013.v3.p10, pht000014.v3.p10, pht000015.v3.p10, pht000016.v3.p10, pht000017.v3.p10, pht000018.v4.p10, pht000019.v3.p10, pht000020.v3.p10, pht000021.v3.p10, pht000022.v4.p10, pht000023.v4.p10, pht000024.v4.p10, pht000025.v4.p10, pht000026.v3.p10, pht000027.v3.p10, pht000028.v3.p10, pht000029.v3.p10, pht000040.v4.p10, pht003099.v4.p10, pht000030.v7.p10, pht000031.v7.p10, pht000032.v6.p10, pht000033.v8.p10, pht000034.v7.p10, pht000035.v8.p10, pht000036.v8.p10, pht000747.v5.p10, pht000041.v6.p10, pht000074.v9.p10, pht000182.v12.p10, pht001415.v16.p10, pht000183.v12.p10, phg000076.v5). The authors would like to thank the participants, investigators and staff of the Framingham study for their important contributions.

MESA (phs000209.v13.p3): MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-RR-025005, and UL1-TR-000040. The MESA CARe data used for the analyses described in this manuscript were obtained through dbGaP (phs000209.v13.p3, pht001116.v10.p3, pht001118.v8.p3, pht001119.v8.p3, pht001120.v8.p3, pht003091.v3.p3, phg000081.v2). Funding for CARe genotyping was provided by NHLBI Contract N01-HC-65226. The authors would like to thank the participants, investigators and staff of the MESA study for their important contributions.

COPD (phs000179.v4.p2): This research used data generated by the COPDGene study, which was supported by NIH grants U01HL089856 and U01HL089897. The COPDGene project is also supported by the COPD Foundation through contributions made by an Industry Advisory Board comprised of Pfizer, AstraZeneca, Boehringer Ingelheim, Novartis, and Sunovion. Individual-participant phenotypic and genotypic data was extracted from the following dbGaP datasets (phs000179.v4.p2, pht002239.v3.p2, pht002237.v2.p2, pht002238.v4.p2, phg000490.v1). The authors would like to thank the participants, investigators and staff of the COPD study for their important contributions.

eMERGE (phs000888.v1.p1): Group Health Cooperative/University of Washington Funding support for Alzheimer's Disease Patient Registry (ADPR) and Adult Changes in Thought (ACT) study was provided by a U01 from the National Institute on Aging (Eric B. Larson, PI, U01AG006781). A gift from the 3M Corporation was used to expand the ACT cohort. DNA aliquots sufficient for GWAS from ADPR Probable AD cases, who had been enrolled in Genetic Differences in Alzheimer's Cases and Controls (Walter Kukull, PI, R01 AG007584) and obtained under that grant, were made available to eMERGE without charge. Funding support for genotyping, which was performed at Johns Hopkins University, was provided by the NIH (U01HG004438). Genome-wide association analyses were supported through a Cooperative Agreement from the National Human Genome Research Institute, U01HG004610 (Eric B. Larson, PI). Assistance with phenotype harmonization and genotype data cleaning was provided by the eMERGE Administrative Coordinating Center (U01HG004603) and the National Center for Biotechnology Information (NCBI). The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000234.v1.p1. ***Mayo Clinic*** samples and associated genotype and phenotype data used in this study were provided by the Mayo Clinic. Funding support for the Mayo Clinic was provided through a cooperative agreement with the National Human Genome Research Institute (NHGRI), Grant #: UOIHG004599; and by grant HL75794 from the National Heart Lung and Blood

Institute (NHLBI). Funding support for genotyping, which was performed at The Broad Institute, was provided by the NIH (U01HG004424). Assistance with phenotype harmonization and genotype data cleaning was provided by the eMERGE Administrative Coordinating Center (U01HG004603) and the National Center for Biotechnology Information (NCBI). The datasets used for analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000203.v1.p1. ***Marshfield Clinic Research Foundation*** funding support for the Personalized Medicine Research Project (PMRP) was provided through a cooperative agreement (U01HG004608) with the National Human Genome Research Institute (NHGRI), with additional funding from the National Institute for General Medical Sciences (NIGMS) The samples used for PMRP analyses were obtained with funding from Marshfield Clinic, Health Resources Service Administration Office of Rural Health Policy grant number D1A RH00025, and Wisconsin Department of Commerce Technology Development Fund contract number TDF FYO10718. Funding support for genotyping, which was performed at Johns Hopkins University, was provided by the NIH (U01HG004438). Assistance with phenotype harmonization and genotype data cleaning was provided by the eMERGE Administrative Coordinating Center (U01HG004603) and the National Center for Biotechnology Information (NCBI). The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000170.v1.p1. ***Northwestern University*** samples and data used in this study were provided by the NUgene Project (www.nugene.org). Funding support for the NUgene Project was provided by the Northwestern Universitys Center for Genetic Medicine, Northwestern University, and Northwestern Memorial Hospital. Assistance with phenotype harmonization was provided by the eMERGE Coordinating Center (Grant number U01HG04603). This study was funded through the NIH, NHGRI eMERGE Network (U01HG004609). Funding support for genotyping, which was performed at The Broad Institute, was provided by the NIH (U01HG004424). Assistance with phenotype harmonization and genotype data cleaning was provided by the

eMERGE Administrative Coordinating Center (U01HG004603) and the National Center for Biotechnology Information (NCBI). The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000237.v1.p1. ***Vanderbilt University*** funding support for the Vanderbilt Genome-Electronic Records (VGER) project was provided through a cooperative agreement (U01HG004603) with the National Human Genome Research Institute (NHGRI) with additional funding from the National Institute of General Medical Sciences (NIGMS). The dataset and samples used for the VGER analyses were obtained from Vanderbilt University Medical Center's BioVU, which is supported by institutional funding and by the Vanderbilt CTSA grant UL1RR024975 from NCRR/NIH. Funding support for genotyping, which was performed at The Broad Institute, was provided by the NIH (U01HG004424). Assistance with phenotype harmonization and genotype data cleaning was provided by the eMERGE Administrative Coordinating Center (U01HG004603) and the National Center for Biotechnology Information (NCBI). The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000188.v1.p1. ***Geisinger Health System*** samples and data in this obesity study were provided by the non-alcoholic steatohepatitis (NASH) project. Funding for the NASH project was provided by a grant from the Clinic Research Fund of Geisinger Clinic. Funding support for the genotyping of the NASH cohort was provided by a Geisinger Clinic operating funds and an award from the Clinic Research Fund. The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000380.v1.p1. Samples and data in this study were provided by the abdominal aortic aneurysm (AAA) project. Funding for the AAA project was provided by a grant from the Clinic Research Fund of Geisinger Clinic. Funding support for the genotyping of the AAA cohort was provided by a Geisinger Clinic operating funds and an award from the Clinic Research Fund. The datasets used for the analyses described in this manuscript were obtained from dbGaP at

http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000387.v1.p1. Samples and data in this study were provided by the Geisinger MyCode Project. Funding for the MyCode Project was provided by a grant from Commonwealth of Pennsylvania and the Clinic Research Fund of Geisinger Clinic. Funding support for the genotyping of the MyCode cohort was provided by a Geisinger Clinic operating funds and an award from the Clinic Research Fund. The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000381.v1.p1. ***Mount Sinai School of Medicine*** samples and data used in this study were provided by the Mount Sinai School of Medicine (MSSM) Biobank Project funded by The Charles R. Bronfman Institute for Personalized Medicine (IPM) at Mount Sinai School of Medicine. The Coronary Artery Disease study (IPM BioBank GWAS) is a genome-wide association study funded by the Charles R. Bronfman Institute for Personalized Medicine. The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000388.v1.p1. ***The Childrens Hospital of Philadelphia (CHOP)*** samples and associated genotype and phenotype data used in this study were provided by the Center for Applied Genomics at the Childrens Hospital of Philadelphia. Genotyping for this project was performed at the Center for Applied Genomics and supported by an Institutional DevelopmentAwardfromTheChildrensHospitalofPhiladelphia. Wegratefullythankallthe children and their families who enrolled in this study, and all individuals who donated blood samples for research purposes. The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000490.v1.p1. ***Boston Children's Hospital (BCH)*** samples and data used in this study are provided by The Gene Partnership (TGP) (http://www.genepartnership.org/) a prospective longitudinal study to study the genetic and environmental contributions to childhood health and diseases, collect genetic information on a large number of children who

have been phenotyped, and implement the Informed Cohort and the Informed Cohort Oversight Board (ICOB). Children's Hospital Boston (CHB) has committed $10 million for the start-up of the TGP. The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000495.v1.p1. ***Cincinnati Childrens Hospital Medical Center (CCHMC)*** CCHMC is a participating Pediatric Institution for Phase II of the eMERGE network, a national consortium formed for the purpose of integrating electronic medical records with DNA and sera repositories for large scale, high throughput genetic research. Multiple CCHMC PIs have contributed genome wide association data with various funding support mechanisms. These support mechanisms can be categorized into two groups: disease specific awards (PI initiatives) which focus on particular samples and phenotypes and non-specific awards which contributed to a clinical service. Disease specific awards: 1. Juvenile idiopathic arthritis (JIA): Samples were collected and genotyping was performed by Dr. David Glass with funding support from N01AR42272 and P01AR048929 (PI: Glass). Additional support and genotyping for systemic JIA has been provided by Dr. Dan Kastners laboratory at the NIH. As of the date of submission, the JIA GWAS data have not been published. 2. Absence seizures: Samples were collected by Dr. Tracy Glauser and genotyping was performed with the support of 5 U01 NS045911 (PI: Glauser) from the National Institute of Neurological Disorders and Stroke. 3. Autism Spectrum Disorder (ASD): Samples were collected by Drs. Cynthia Molloy and Patricia Manning-Courtney and genotyping was performed with the support of Award 1984, Genome-wide Association Study of Autism Characterized by Developmental Regression (PIs: Molloy & Manning), from Autism Speaks Inc. 4. Eosinophilic Esophagitis: Samples were collected and genotyping was performed by Dr. Marc Rothenberg with funding support of 5 U19 AI066738 Project 3, Eosinophilic esophagitis and food allergy (PI: Sampson, Co-PI & Project 3 PI: Rothenberg). As of the date of the submission, the eosinophilic esophagitis data have not been published. 5. Bicuspid Aortic Valve: Samples

were collected and genotyping was performed by Dr. Woodrow Benson with funding support from NIH/NHLBI award HL69712, Genetic mechanisms of cardiac disease in the young (PI: Benson), and NIH/NHLBI award HL74728, SCCOR in Pediatric Heart Development and Disease titled Molecular mechanisms of valve development and disease (PI: Benson). Non-specific awards: 1. The Cincinnati Control Cohort is a collection of biological samples that have been collected and genotyped through a multidisciplinary approach and with collaboration of more than twenty divisions within CCHMC, supported by the Cincinnati Childrens Research Foundation. Lead PIs responsible for this collection are Drs. David Glass and Ardythe Morrow. 2. Clinical cytogenetics samples. Since 2007, more than 2000 samples, enriched for developmental delay, autism and various rare or common genetic diseases as well as specific chromosomal abnormalities such as deletions and duplications, have been genotyped for the purpose of uncovering chromosomal abnormalities. The extraction of data from the EPIC electronic medical record into the de-identified data warehouse, i2b2, was made possible by institutional resources and 1UL1RR026314, Cincinnati Center for Clinical and Translational Sciences and Training Grant (PI: Heubi). The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000494.v1.p1. Assistance with phenotype harmonization and genotype data cleaning was provided by the eMERGE Administrative Coordinating Center (U01HG004603) and the National Center for Biotechnology Information (NCBI). The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number (phs000888.v1.p1,pht004678.v1.p1, pht004677.v1.p1, pht004680.v1.p1, pht005581.v1.p1, pht005587.v1.p1, phg000569.v1, phg000896.v1).

WHI (phs000200.v10.p3): The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C. This manuscript

was not prepared in collaboration with investigators of the WHI, has not been reviewed and/or approved by the Women?s Health Initiative (WHI), and does not necessarily reflect the opinions of the WHI investigators or the NHLBI. ***WHI PAGE*** is funded through the NHGRI Population Architecture Using Genomics and Epidemiology (PAGE) network (Grant Number U01 HG004790). Assistance with phenotype harmonization, SNP selection, data cleaning, meta-analyses, data management and dissemination, and general study coordination, was provided by the PAGE Coordinating Center (U01HG004801-01). ***GARNET*** funding support for WHI GARNET was provided through the NHGRI Genomics and Randomized Trials Network (GARNET) (Grant Number U01 HG005152). Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the GARNET Coordinating Center (U01 HG005157). Assistance with data cleaning was provided by the National Center for Biotechnology Information. Funding support for genotyping, which was performed at the Broad Institute of MIT and Harvard, was provided by the NIH Genes, Environment and Health Initiative [GEI] (U01 HG004424). ***WHISP*** the Women?s Health Initiative Sequencing Project (WHISP) was funded by Grant Number RC2 HL102924. This study was part of the NHLBI Grand Opportunity Exome Sequencing Project (GOESP). Funding for GO-ESP was provided by NHLBI grants RC2 HL103010 (HeartGO), RC2 HL102923 (LungGO) and RC2 HL102924 (WHISP). The exome sequencing was performed through NHLBI grants RC2 HL102925 (BroadGO) and RC2 HL102926 (SeattleGO). ***SHARe*** funding for WHI SHARe genotyping was provided by NHLBI Contract N02-HL-64278. ***WHISE*** the WHI Sight Exam and the Memory Study was funded in part by Wyeth Pharmaceuticals, Inc, St. Davids, PA. The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap through dbGaP accession (phs000200.v10.p3, pht000998.v5.p3, pht001019.v5.p3, pht000987.v5.p3, pht000998.v5.p3, phg000592.v1).The authors would like to thank the participants, investigators and staff of the WHI study for their important contributions.

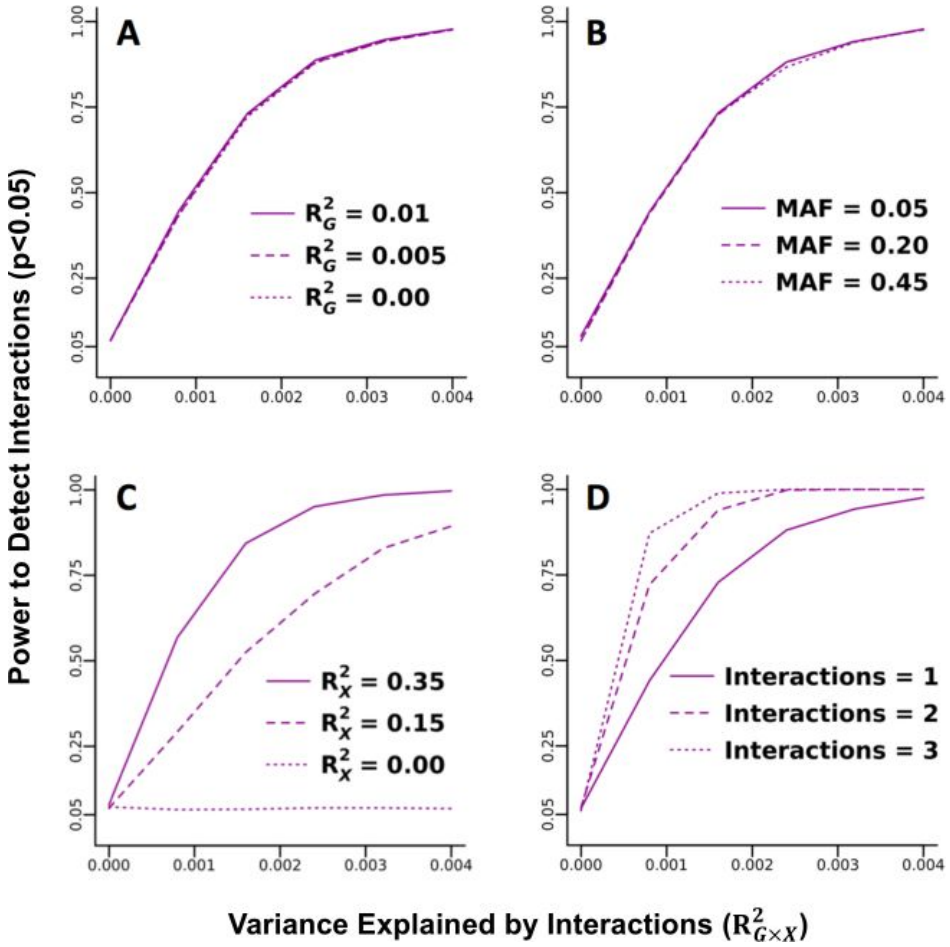


Figure A1: Simulation study of the power to detect unadjusted interactions using conditional quantile regression (CQR) and meta-regression (MR). The power to detect unadjusted interactions between a SNP (G) and a continuous variable (X) was simulated in a sample of 10,000 individuals. Unless otherwise indicated, the simulation conditions were minor allele frequency (MAF) = 0.2, variance explained by G (R2G) = 0.004, variance explained by X (R^2_X) = 0.25, and the variance explained by the interaction between G and X ($R^2_{G\times X}$) was varied between 0 and 0.004. CQR models were fitted at every 10^{th} percentile of the distribution of Y from the 5^{th} to the 95^{th} percentiles and MR was used to model the relationship between variation in CQR estimates and the Y percentiles. The power to detect unadjusted interactions at a threshold of $p < 0.05$ was computed from 1,000 replicates of each simulation condition and plotted against the value of $R^2_{G\times X}$. The power to detect interactions at different values of R^2_G MAF, R^2_X and the number of interactions was investigated (A, B, C and D, respectively). When more than one interaction was considered, R^2_X was divided equally between all interaction covariates, while each additional interaction was equal to $R^2_{G\times X}$. Overall the power to detect unadjusted interactions was not affected by the main effects of G or the MAF, but was enhanced by the main effects of X and the number of interactions.

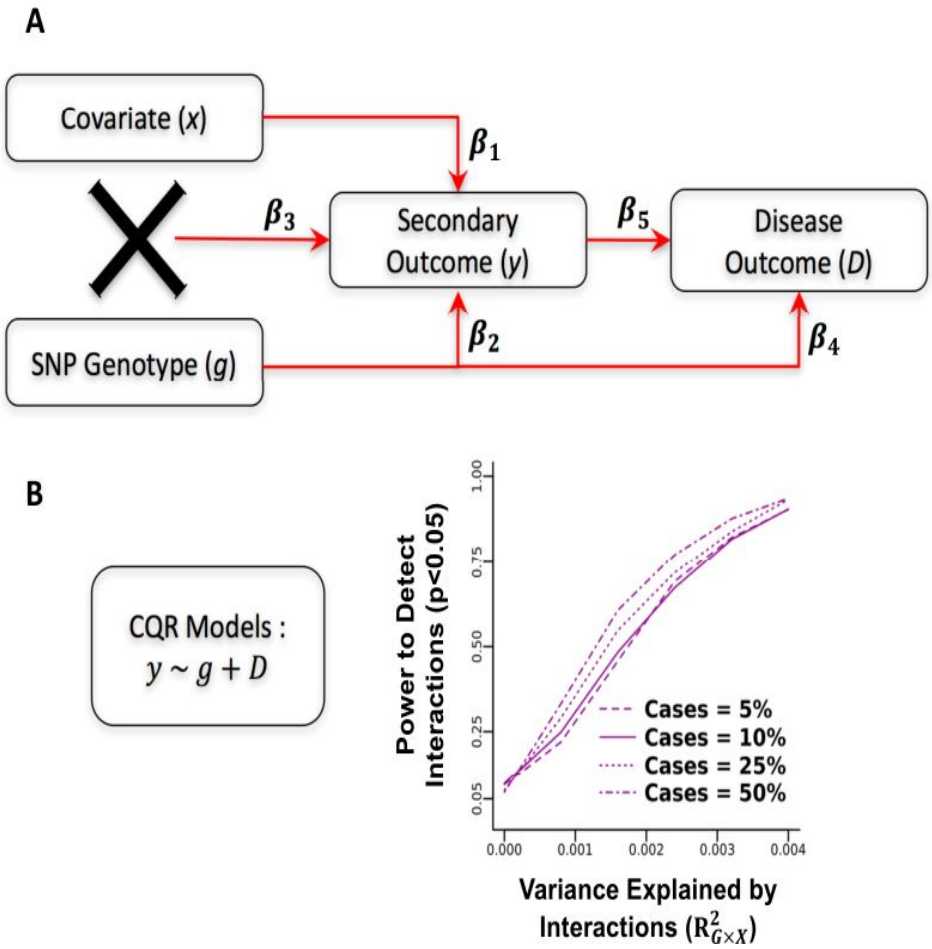
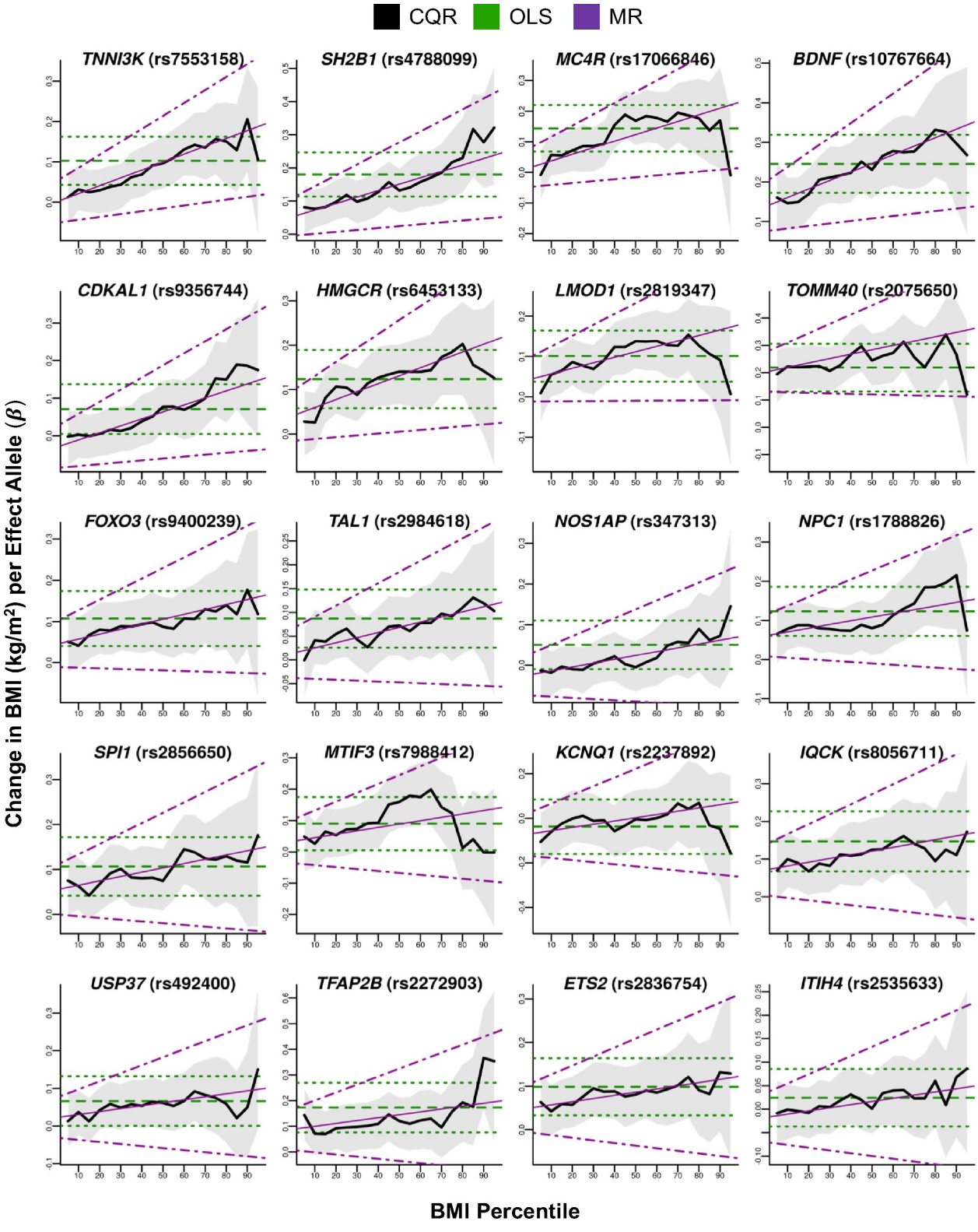


Figure A2: Sample stratification and the detection of unadjusted interactions in simulations. (A) A schematic representation of the model described by equations 1, 9, 10 and 11 in Appendix A. (B) Investigating the effects of sample stratification on the power to detect unadjusted interactions using conditional quantile regression (CQR) and meta-regression (MR) in a simulation study. The simulation conditions were minor allele frequency (MAF) = 0.2, variance of Z explained by G ($R^2_{G[Z]}$) = 0.01 (equivalent to OR $\sim$ 1.4 of G on D), variance Z explained by Y (R^2_Z) = 0.2 (equivalent to OR 2.5 of Y on D), variance of Y explained by G ($R^2_{G[Y]}$) = 0.004, variance of Y explained by X (R^2_X) = 0.25 and the variance of Y explained by the interaction between G and X ($R^2_{G\times X}$) was varied between 0 and 0.004. A population of 100,000 individuals was generated with disease prevalence (n_0) = 10%. A sample population of 10,000 individuals with pre-specified proportion of cases was randomly selected from this population. The power to detect unadjusted interactions between the SNP (G) and the continuous variable (X) in this sample was computed and plotted as in Figure A1, except that CQR models were adjusted for disease status (D). Overall the power to detect unadjusted interactions was not affected by sample stratification when CQR models were adjusted for disease status.


CQR
OLS
MR
TNNI3K (rs7553158)
SH2B1 (rs4788099)
MC4R (rs17066846)
BDNF (rs10767664)
CDKAL1 (rs9356744)
HMGCR (rs6453133)
LMOD1 (rs2819347)
TOMM40 (rs2075650)
FOXO3 (rs9400239)
TAL1 (rs2984618)
NOS1AP (rs347313)
NPC1 (rs1788826)
SPI1 (rs2856650)
MTIF3 (rs7988412)
KCNQ1 (rs2237892)
IQCK (rs8056711)
USP37 (rs492400)
TFAP2B (rs2272903)
ETS2 (rs2836754)
ITIH4 (rs2535633)
Change in BMI (kg/m²) per Effect Allele (β)
BMI Percentile

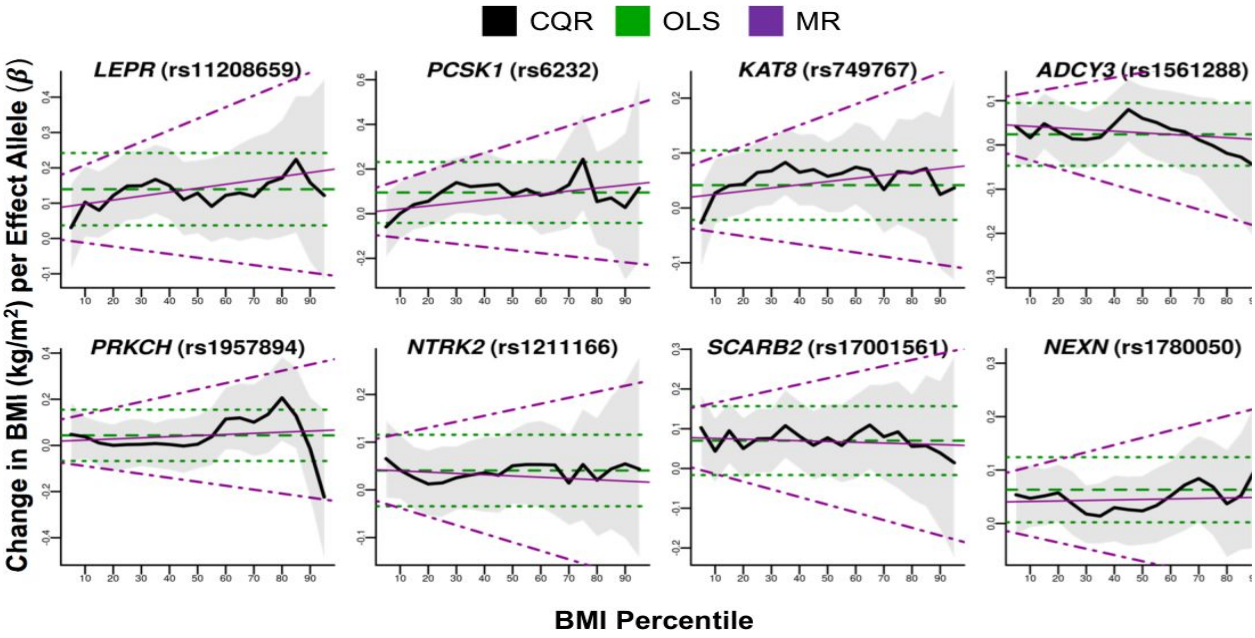
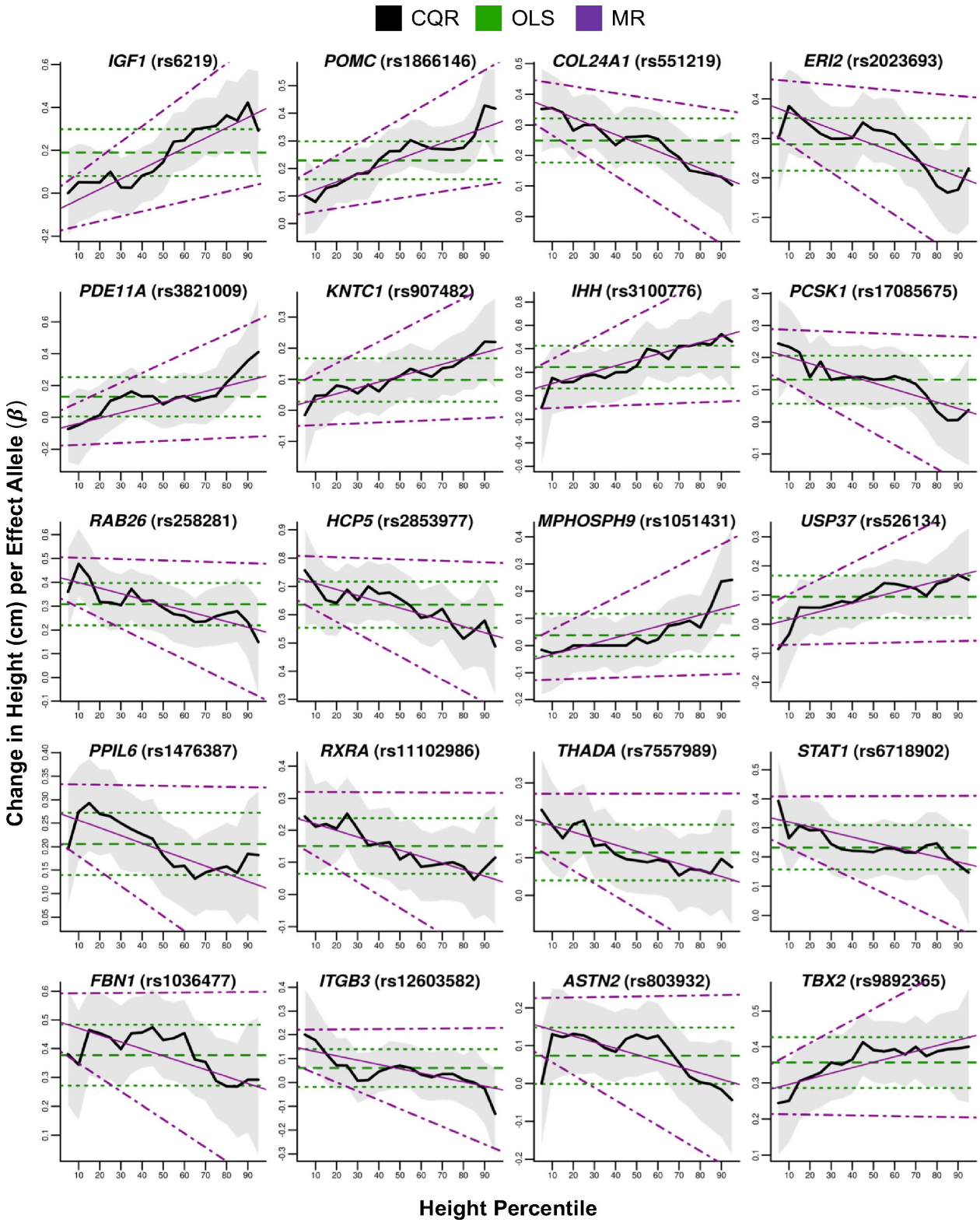
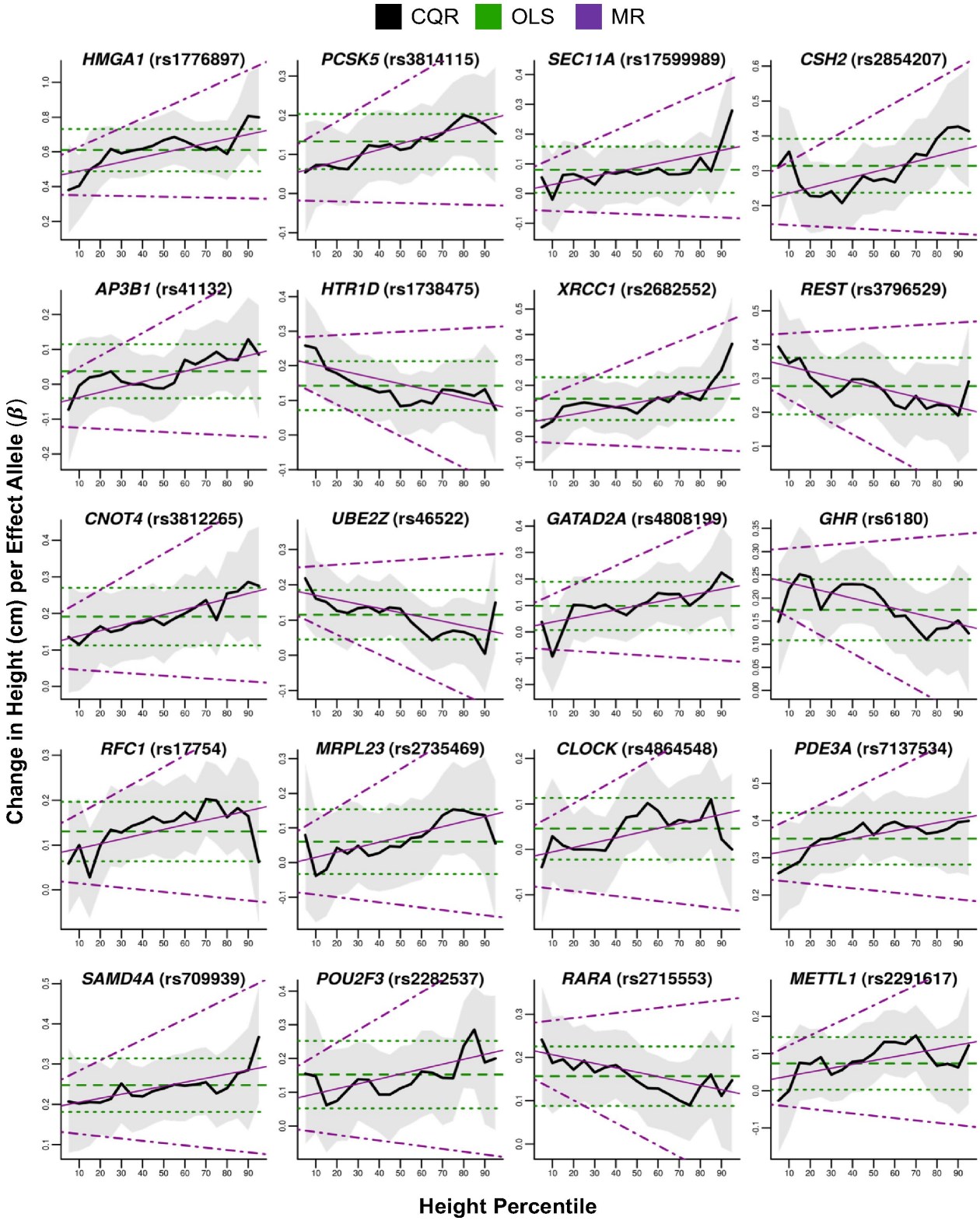
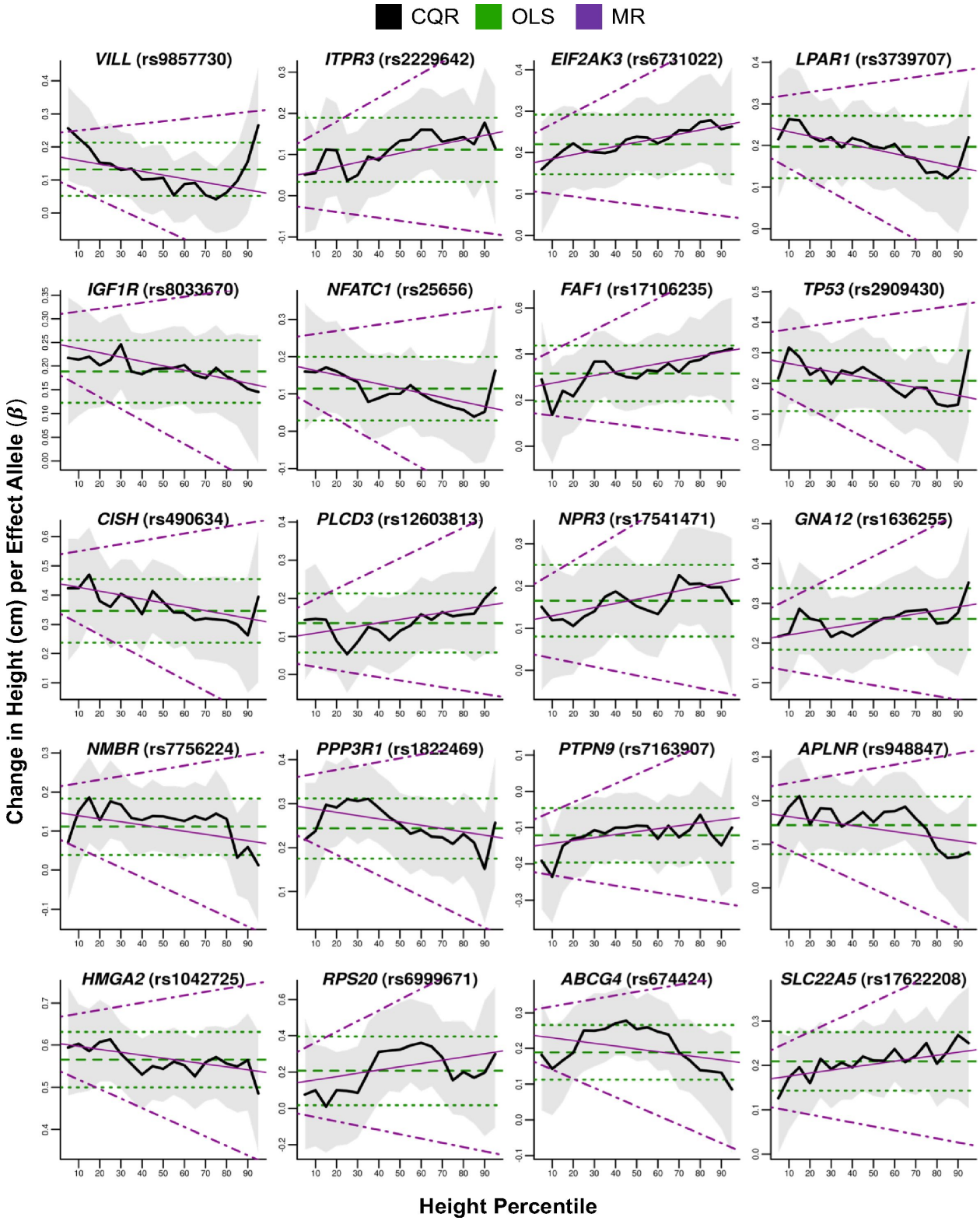
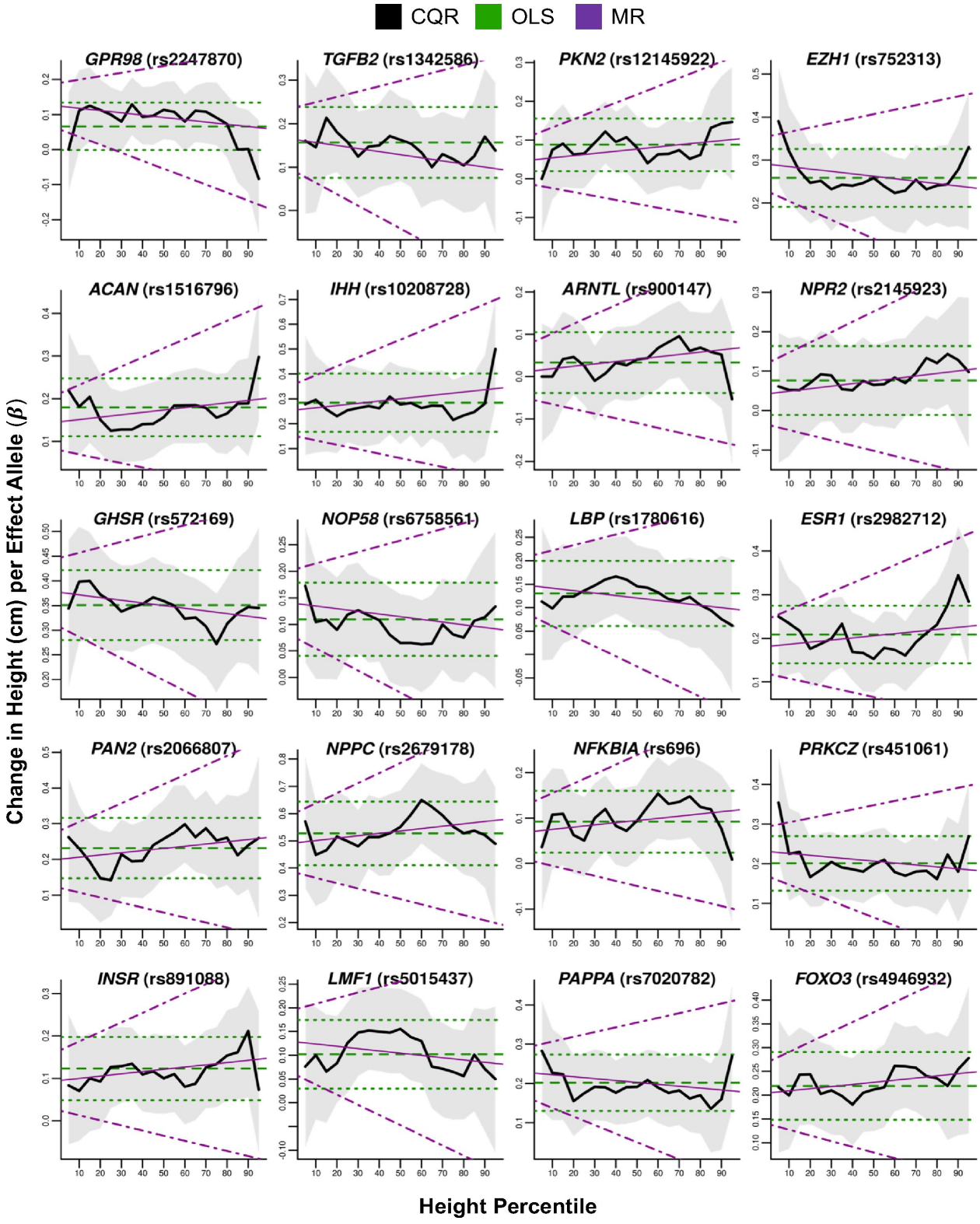


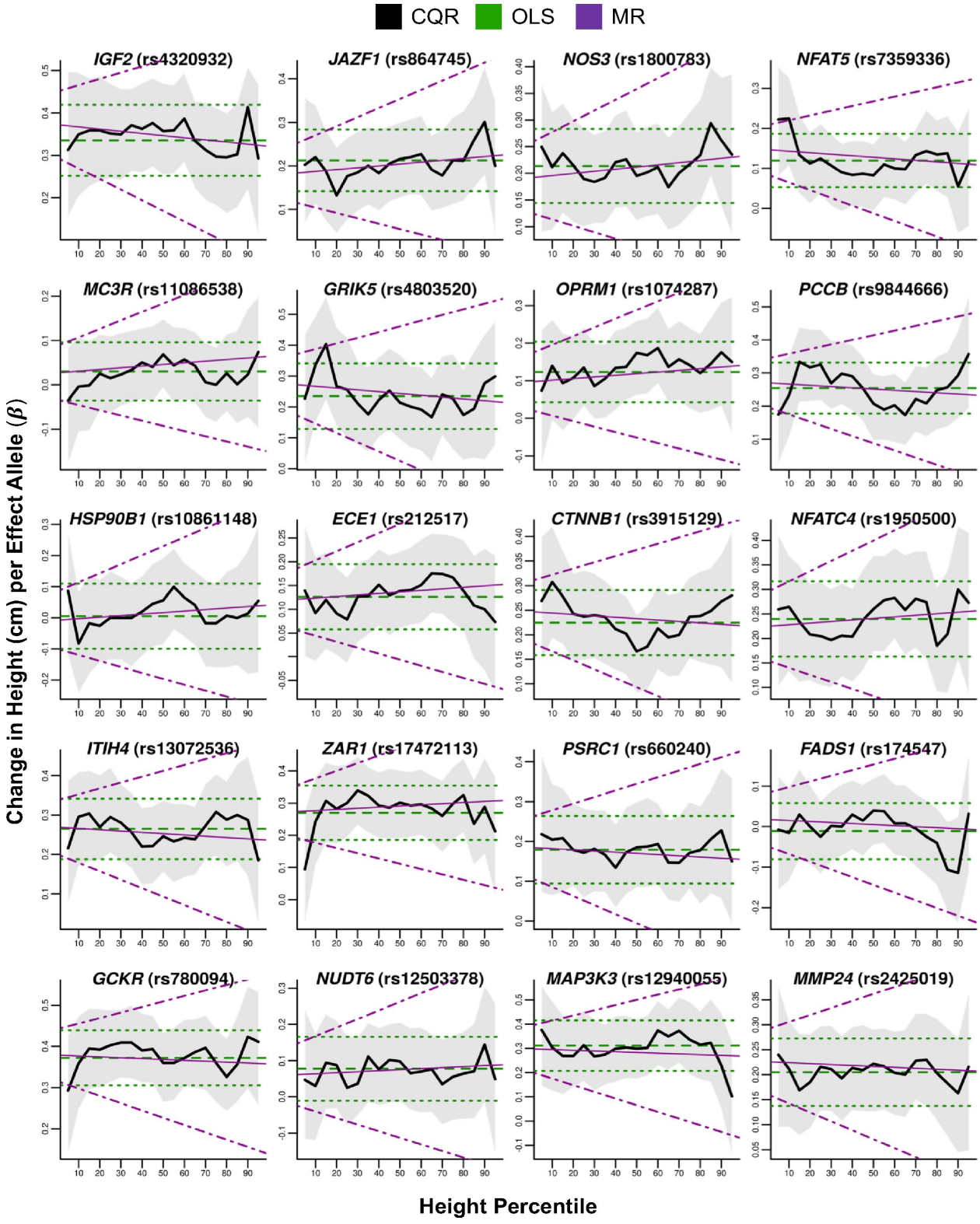
Figure A3: The effects of BMIl obesity-associated SNPs across the sample BMI distribution (continued). As in Figure 2, estimates of the change in BMI per effect allele (β_{CQR}, kg/m^2 per Effect Allele) from conditional quantile regression (CQR) models of BMI/obesity-associated SNPs was plotted against the BMI percentile (thick-black line) along with the 95% confidence intervals (shaded-grey region). The results from ordinary least square (OLS) (β_{OLS} , kg/m^2 per Effect Allele, horizontal-dashed-green line) and the 95% confidence intervals (horizontal-dashed-green lines) were also plotted for comparison. The change in CQR estimates across BMI percentiles was modelled using meta-regression (MR) and estimates from MR (β_{MR} , kg/m^2 per Effect Allele per BMI Percentile, thin-magenta line) and the 95% confidence intervals (dotdashed-magenta lines) were plotted. MR analysis did not detect significant ($p < 1.32 \times 10^{-3}$) increases in the effects of these SNPs across the sample BMI distribution.

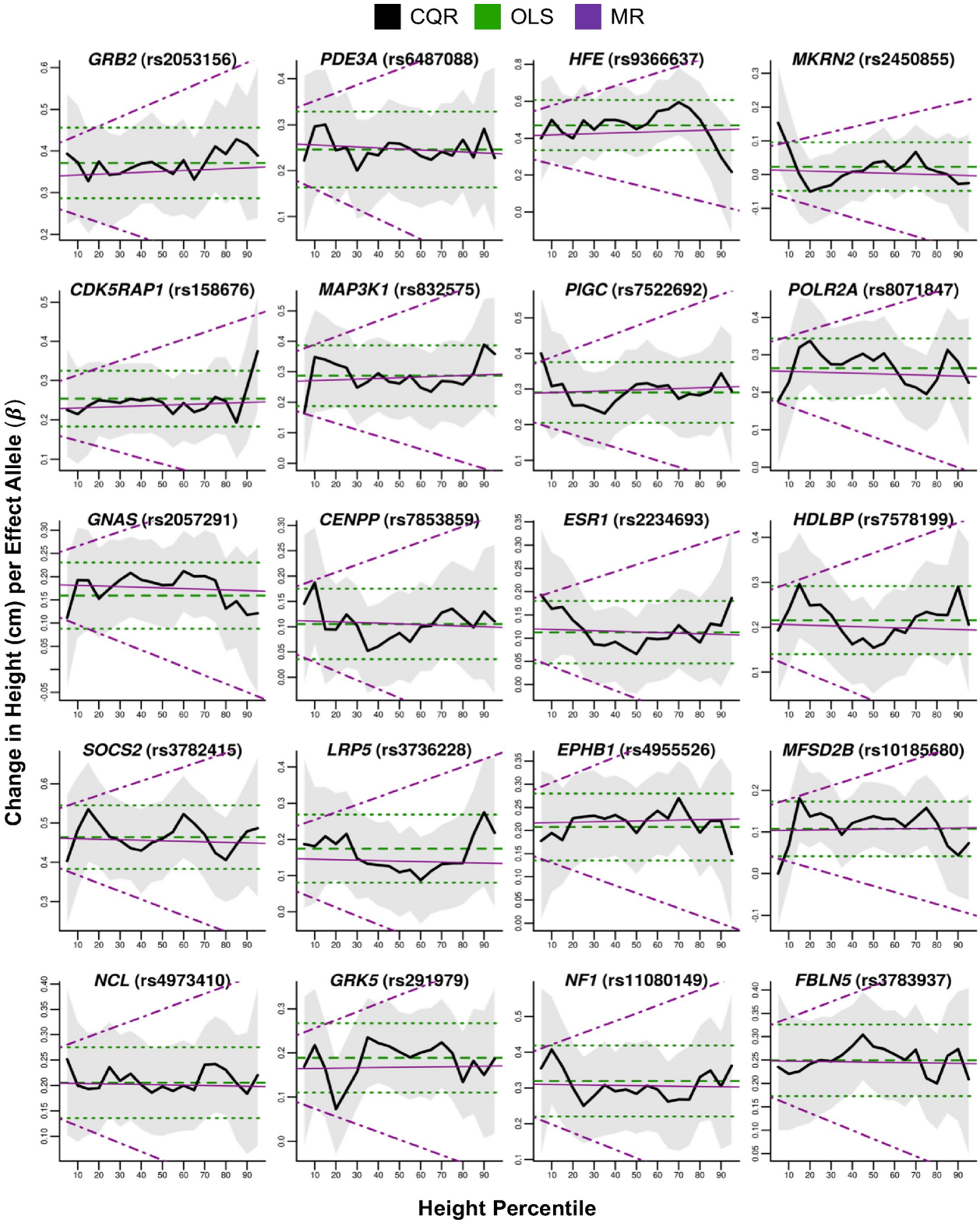

CQR
OLS
MR
IGF1 (rs6219)
POMC (rs1866146)
COL24A1 (rs551219)
ERI2 (rs2023693)
PDE11A (rs3821009)
KNTC1 (rs907482)
IHH (rs3100776)
PCSK1 (rs17085675)
RAB26 (rs258281)
HCP5 (rs2853977)
MPHOSPH9 (rs1051431)
USP37 (rs526134)
PPIL6 (rs1476387)
RXRA (rs11102986)
THADA (rs7557989)
STAT1 (rs6718902)
FBN1 (rs1036477)
ITGB3 (rs12603582)
ASTN2 (rs803932)
TBX2 (rs9892365)
Change in Height (cm) per Effect Allele (β)
Height Percentile


CQR
OLS
MR
HMGA1 (rs1776897)
PCSK5 (rs3814115)
SEC11A (rs17599989)
CSH2 (rs2854207)
AP3B1 (rs41132)
HTR1D (rs1738475)
XRCC1 (rs2682552)
REST (rs3796529)
CNOT4 (rs3812265)
UBE2Z (rs46522)
GATAD2A (rs4808199)
GHR (rs6180)
RFC1 (rs17754)
MRPL23 (rs2735469)
CLOCK (rs4864548)
PDE3A (rs7137534)
SAMD4A (rs709939)
POU2F3 (rs2282537)
RARA (rs2715553)
METTL1 (rs2291617)
Change in Height (cm) per Effect Allele (β)
Height Percentile


CQR
OLS
MR
VILL (rs9857730)
ITPR3 (rs2229642)
EIF2AK3 (rs6731022)
LPAR1 (rs3739707)
IGF1R (rs8033670)
NFATC1 (rs25656)
FAF1 (rs17106235)
TP53 (rs2909430)
CISH (rs490634)
PLCD3 (rs12603813)
NPR3 (rs17541471)
GNA12 (rs1636255)
NMBR (rs7756224)
PPP3R1 (rs1822469)
PTPN9 (rs7163907)
APLNR (rs948847)
HMGA2 (rs1042725)
RPS20 (rs6999671)
ABCG4 (rs674424)
SLC22A5 (rs17622208)
Change in Height (cm) per Effect Allele (β)
Height Percentile


CQR
OLS
MR
GPR98 (rs2247870)
TGFB2 (rs1342586)
PKN2 (rs12145922)
EZH1 (rs752313)
ACAN (rs1516796)
IHH (rs10208728)
ARNTL (rs900147)
NPR2 (rs2145923)
GHSR (rs572169)
NOP58 (rs6758561)
LBP (rs1780616)
ESR1 (rs2982712)
PAN2 (rs2066807)
NPPC (rs2679178)
NFKBIA (rs696)
PRKCZ (rs451061)
INSR (rs891088)
LMF1 (rs5015437)
PAPPA (rs7020782)
FOXO3 (rs4946932)
Change in Height (cm) per Effect Allele (β)
Height Percentile


CQR
OLS
MR
IGF2 (rs4320932)
JAZF1 (rs864745)
NOS3 (rs1800783)
NFAT5 (rs7359336)
MC3R (rs11086538)
GRIK5 (rs4803520)
OPRM1 (rs1074287)
PCCB (rs9844666)
HSP90B1 (rs10861148)
ECE1 (rs212517)
CTNNB1 (rs3915129)
NFATC4 (rs1950500)
ITIH4 (rs13072536)
ZAR1 (rs17472113)
PSRC1 (rs660240)
FADS1 (rs174547)
GCKR (rs780094)
NUDT6 (rs12503378)
MAP3K3 (rs12940055)
MMP24 (rs2425019)
Change in Height (cm) per Effect Allele (β)
Height Percentile


CQR
OLS
MR
GRB2 (rs2053156)
PDE3A (rs6487088)
HFE (rs9366637)
MKRN2 (rs2450855)
CDK5RAP1 (rs158676)
MAP3K1 (rs832575)
PIGC (rs7522692)
POLR2A (rs8071847)
GNAS (rs2057291)
CENPP (rs7853859)
ESR1 (rs2234693)
HDLBP (rs7578199)
SOCS2 (rs3782415)
LRP5 (rs3736228)
EPHB1 (rs4955526)
MFSD2B (rs10185680)
NCL (rs4973410)
GRK5 (rs291979)
NF1 (rs11080149)
FBLN5 (rs3783937)
Change in Height (cm) per Effect Allele (β)
Height Percentile

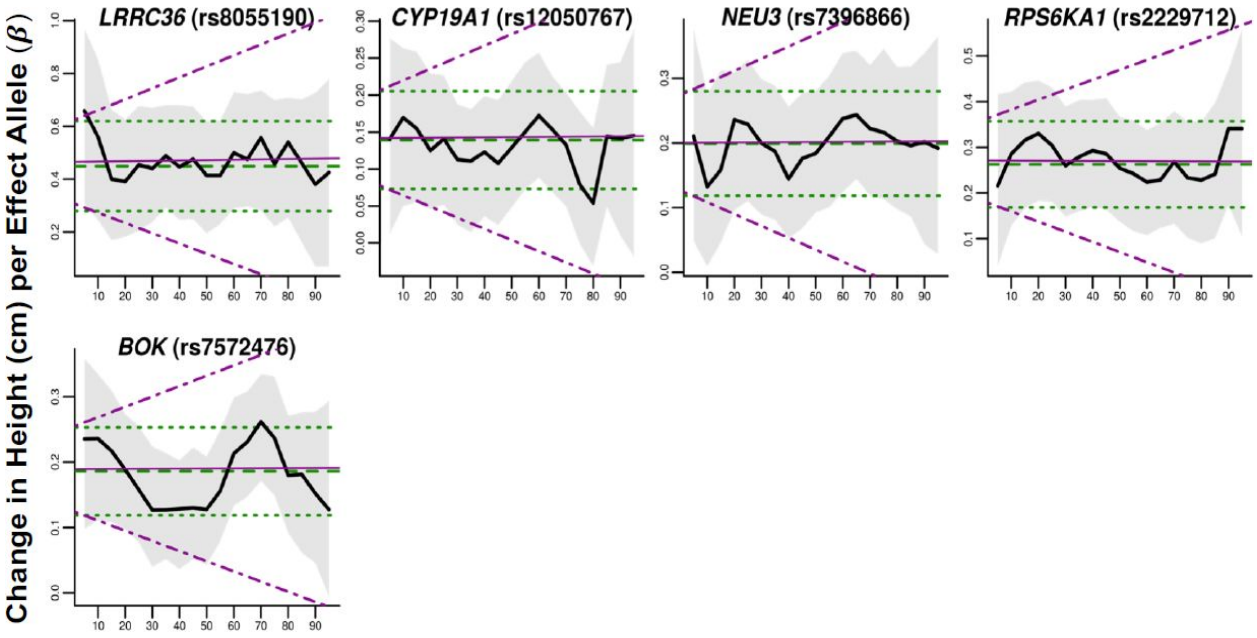
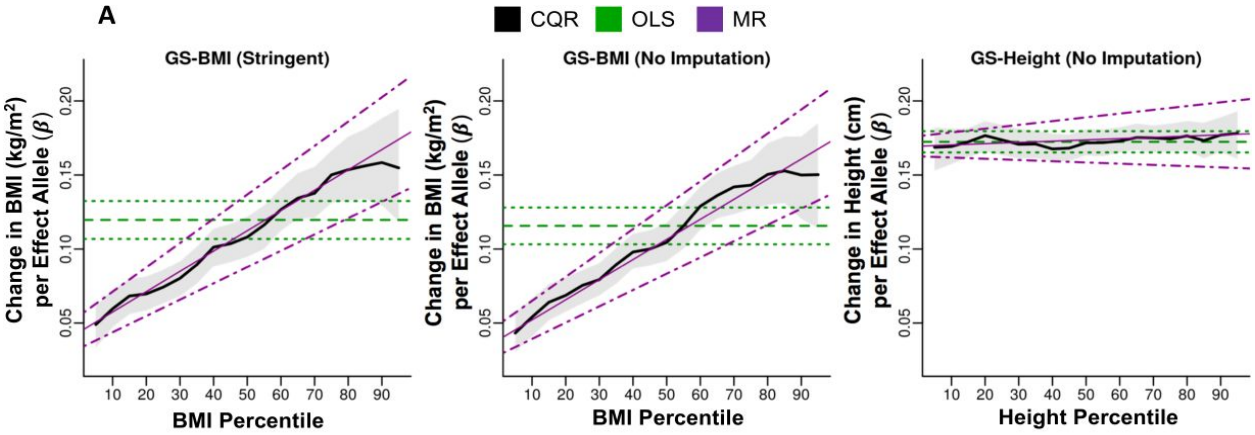


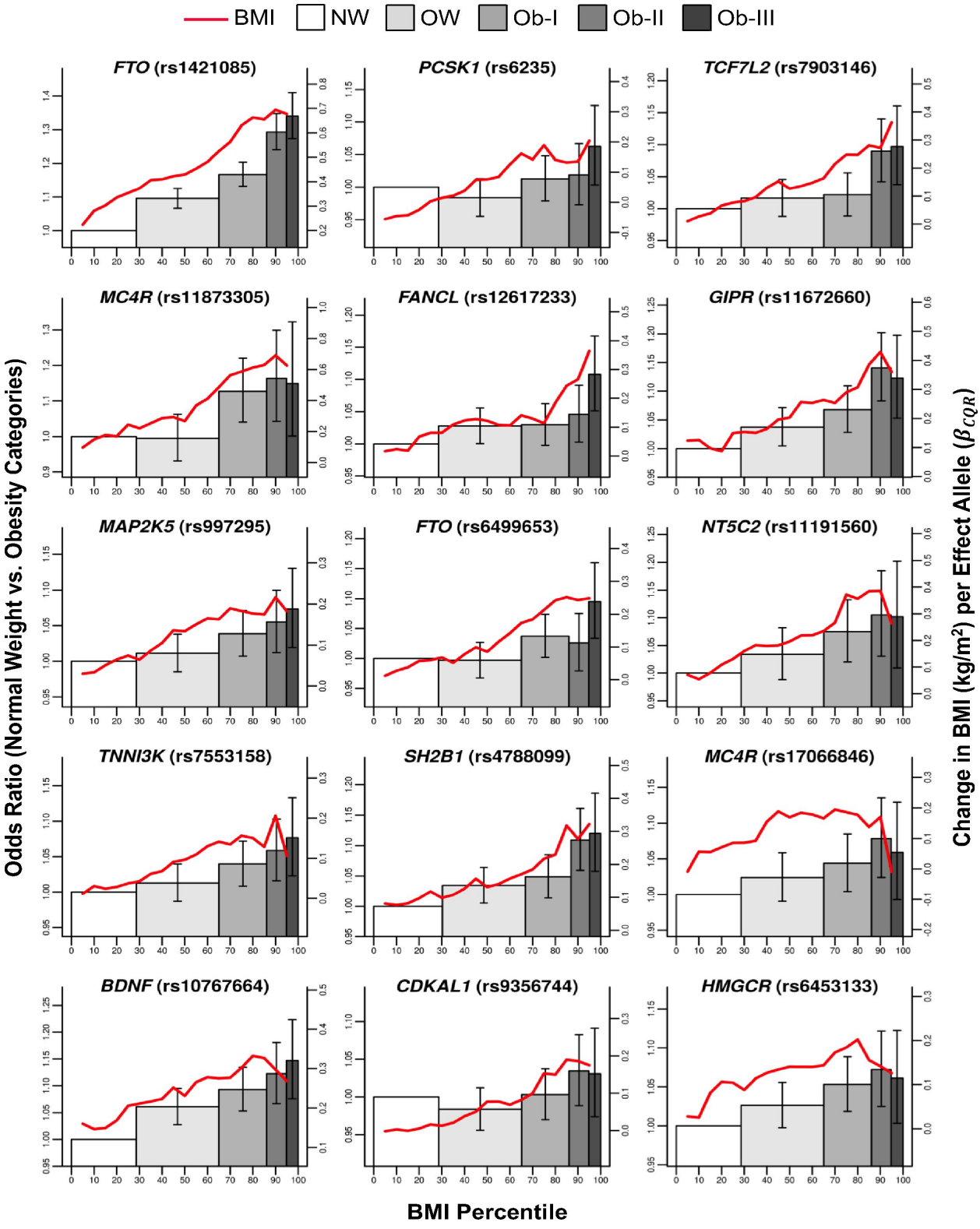
Figure A4: The effects of height-associated SNPs across the distribution of height. Conditional quantile regression (CQR) models of height-associated SNPs were fitted every 5^{th} percentile of height and adjusted for age, sex and study. Estimates of the change in height per effect allele (β_{CQR} , cm per Effect Allele) from these models was plotted against the height percentile (thick-black line) along with the 95% confidence intervals (shaded-grey region). The results from ordinary least square (OLS) models (β_{OLS} , cm per Effect Allele, horizontal-dashed- green line) and the 95% confidence intervals (horizontal-dotted-green lines) were also plotted for comparison. The change in CQR estimates across height percentiles was modelled using meta-regression (MR) and estimates from MR (β_{MR} , cm per Effect Allele per Height Percentile, thin-magenta line) and the 95% confidence intervals (dotdashed-magenta lines) were plotted.

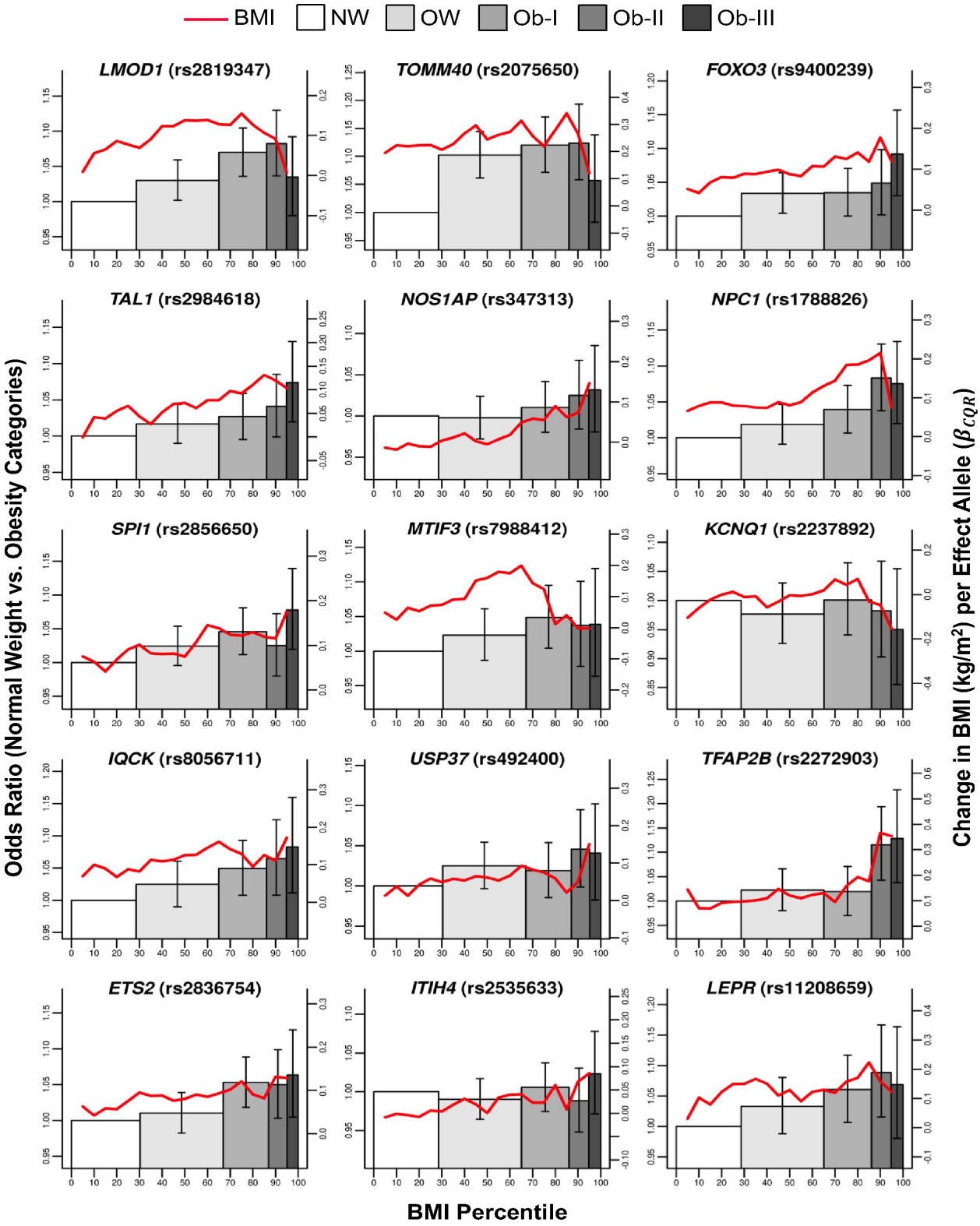


B

SNP	OLS Models		Meta-Regression of CQR Models		
	β_{OLS} [95%CI]	p-value	RI_{50}	β_{MR} [95%CI]	p-value
GS-BMI (Stringent)	0.120 [0.107, 0.132]	2.02E-74	0.112	0.137 [0.110, 0.164]	2.18E-23 *
GS-BMI (No Imputation)	0.116 [0.103, 0.128]	5.80E-74	0.106	0.136 [0.110, 0.161]	3.93E-25 *
GS-Height (No Imputation)	0.172 [0.165, 0.180]	<2.2E-308	0.174	0.009 [-0.008, 0.025]	0.323

Figure A5: Sensitivity analysis of GS Results. (A) CQR models of GS-BMI (Stringent), GS-BMI (No Imputation) and GS-Height (No Imputation) fitted as in Figure 2 and plotted against respective outcome percentiles. The thick-black line is the estimated change in each trait per effect allele (BMI, β_{CQR}, kg/m^2 per Effect Allele; Height, β_{CQR}, cm per Effect Allele) and shaded-grey region represents the 95% confidence intervals. Also plotted are the OLS regression estimates (BMI, β_{OLS} in kg/m^2 per Effect Allele; Height, β_{OLS}, cm per Effect Allele, horizontal-dashed-green line) and 95% confidence intervals (horizontal-dotted-green lines). The change in CQR estimates across outcome percentiles was modeled using meta-regression (MR). Estimates from MR (BMI,β_{MR}, kg/m^2 per Effect Allele per BMI Percentile; Height, β_{MR}, cm per Effect Allele per Height Percentile; thin-magenta line) and the 95% confidence intervals (dotdashed-magenta lines) were also plotted. (B) The results from OLS and MR modelling of GS-BMI (Stringent), GS-BMI (No Imputation) and GS-Height (No Imputation). $(*)$ denotes statistical significance, RI_{50} is the re-centered intercept of the MR models and 95% CI are the 95% confidence intervals.


BMI
NW
OW
Ob-I
Ob-II
Ob-III
FTO (rs1421085)
PCSK1 (rs6235)
TCF7L2 (rs7903146)
MC4R (rs11873305)
FANCL (rs12617233)
GIPR (rs11672660)
MAP2K5 (rs997295)
FTO (rs6499653)
NT5C2 (rs11191560)
TNNI3K (rs7553158)
SH2B1 (rs4788099)
MC4R (rs17066846)
BDNF (rs10767664)
CDKAL1 (rs9356744)
HMGCR (rs6453133)
Odds Ratio (Normal Weight vs. Obesity Categories)
Change in BMI (kg/m2) per Effect Allele (βCQR)
BMI Percentile


BMI
NW
OW
Ob-I
Ob-II
Ob-III
LMOD1 (rs2819347)
TOMM40 (rs2075650)
FOXO3 (rs9400239)
TAL1 (rs2984618)
NOS1AP (rs347313)
NPC1 (rs1788826)
SPI1 (rs2856650)
MTIF3 (rs7988412)
KCNQ1 (rs2237892)
IQCK (rs8056711)
USP37 (rs492400)
TFAP2B (rs2272903)
ETS2 (rs2836754)
ITIH4 (rs2535633)
LEPR (rs11208659)
Odds Ratio (Normal Weight vs. Obesity Categories)
Change in BMI (kg/m2) per Effect Allele (βCQR)
BMI Percentile

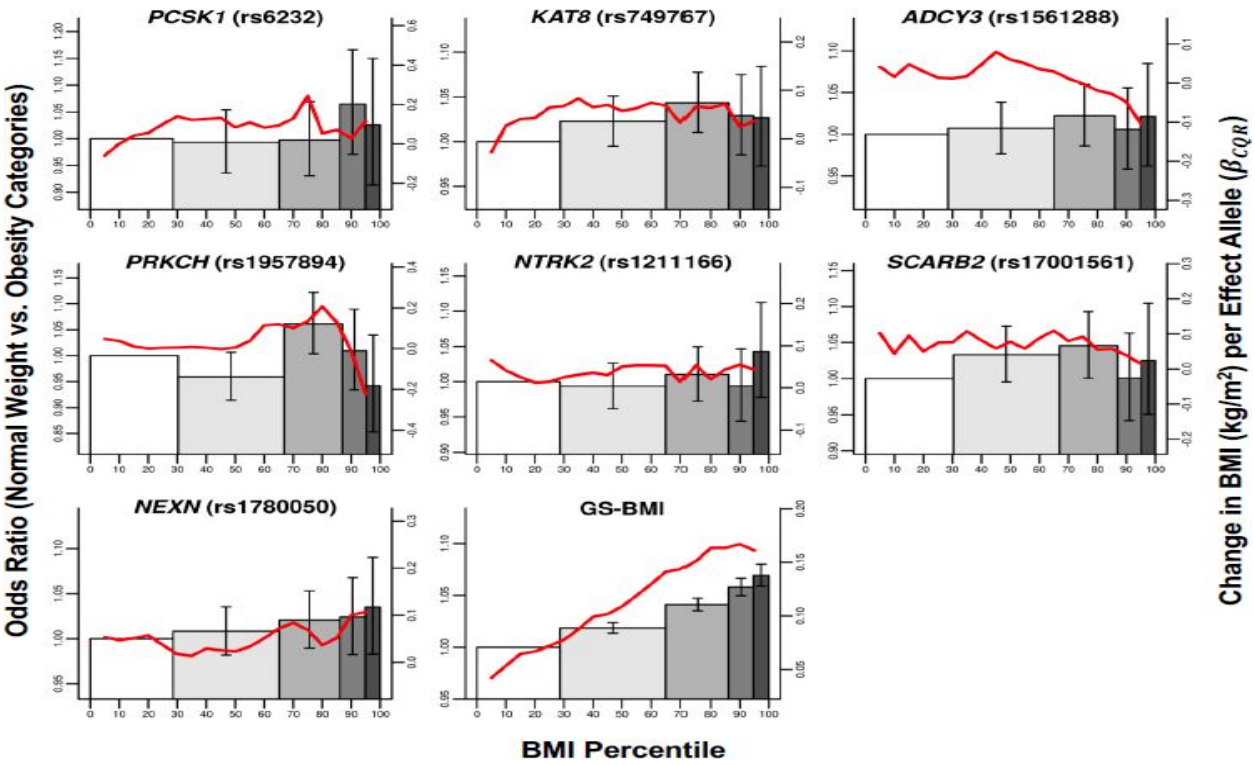


Figure A6: Comparing patterns from subgroup analysis and conditional quantile regression (CQR). BMI was divided into BMI categories, and the effects of each SNP on the risk of overweight (OW), obesity class I (Ob-I), class II (Ob-II) and class III (Ob-III) relative to normal weight (NW) were tested using logistic regression. Models were adjusted for age, age- squared, sex and study. Bar plots of the odds ratio (OR, left axis) for these categories were plotted and bar widths were defined by the percentile cut-offs of each category. Error bars correspond to the 95% confidence intervals. These bar plots were then overlaid with the results from similarly adjusted CQR models (thick-red line, right axis). The patterns from subgroup analysis correspond closely to those from CQR.

Table A1 Subject characteristics. Subject characteristics of the studies included in the analysis of BMI and height; sample size (N), height (mean ± sd), BMI (mean ± sd), age (mean ± sd), the proportion of women, the proportion with diabetes, and the proportions of BMI categories including normal weight (NW), overweight (OW), and obesity classes I (Ob-I), II (Ob-II) and III (Ob-III); within each study and overall are presented.

Study	N	Age (yrs)	Women	Height (cm)	BMI (kg/m^2)	Diabetes	BMI Categories (NW/OW/Ob-I/Ob-II/Ob-III)
ARIC CARe	10094	57.02± 6.25	53.40%	168.24 ± 9.47	27.50 ± 5.00	20.50%	33.6/40.8/17.9/5.4/2.3%
CARDIA CARe	1513	40.06 ± 4.20	53.20%	171.45 ± 9.20	26.98 ± 5.83	7.90%	43.1/33.8/13.7/5.9/3.4%
CHS CARe	4194	75.71 ± 5.73	56.40%	164.82 ± 9.50	26.44 ± 4.51	25.60%	40.2/41.6/14.1/3.1/1.1%
Framingham CARe	2222	48.00 ± 13.06	66.00%	166.70 ± 9.43	26.54 ± 5.11	14.90%	43.6/36.4/13.5/4.5/2.0%
MESA CARe	2435	65.47 ± 10.18	52.30%	168.18 ± 9.65	27.87 ± 5.08	14.20%	30.6/40.9/19.5/6.7/2.3%
COPDGene	3793	60.24 ± 8.82	50.20%	169.73 ± 9.51	29.12 ± 5.98	12.30%	26.1/36.5/22.3/9.3/5.8%
eMERGE II	24584	62.12 ± 15.14	54.90%	169.50 ± 10.06	29.51 ± 6.82	17.00%	26.2/34.8/21.4/10.1/7.5%
EpiDREAM	9148	55.10 ± 10.79	60.70%	166.93 ± 9.44	30.67 ± 6.15	14.70%	16.1/35.7/28.0/12.4/7.8%
MOPMAP (WHI)	2197	66.23 ± 6.87	100%	162.19 ± 5.81	28.93 ± 5.76	5.30%	28.7/33.5/22.6/10.8/4.4%
GARNET (WHI)	4244	68.35 ± 7.07	100%	161.04 ± 6.01	29.79 ± 5.88	8.50%	22.2/33.9/25.7/12.4/5.8%
GECCO (WHI)	2147	67.30 ± 6.66	100%	161.93 ± 6.10	28.21 ± 5.65	6.00%	32.0/36.2/20.0/7.9/3.9%
HIPFX (WHI)	2995	70.34 ± 6.60	100%	161.42 ± 6.33	26.74 ± 5.31	6.80%	42.6/34.1/16.5/4.7/2.2%
WHIMS+ (WHI)	5666	70.61 ± 6.07	100%	160.65 ± 5.96	28.50 ± 5.46	7.10%	28.7/36.9/22.2/8.5/3.7%
Overall	75232	62.08 ± 12.89	66%	166.72 ± 9.58	28.77 ± 6.10	14.80%	28.6/36.5/21.0/8.7/5.2%

Table A2 Study and quality control information. Additional details on the studies that were included in this report including, funding sources, study design, which populations were retained for analysis, datasets from which phenotypes and genotypes were extracted, citations detailing these studies (PMID), pre-QC sample size (N), QC criteria and number of samples that did not pass these thresholds, post-QC sample size (post-QC N), and average sample call rate (Mean CR).

Panel A: Individual Study Information

Study	Full name	Funding	Design	Study Population	Collection Time	PMID	dbGaP Datasets
ARIC CARe (phs000280.v3.p1, phs000557.v2.p1)	Atherosclerosis Risk in Communities (ARIC) Candidate Gene Association Resource (CARe)	National Heart, Lung, and Blood Institute (NHLBI)	Population-based	All	1987-2013	1342297, 20400780	pht004062.v1.p1, pht004063.v1.p1, pht004064.v1.p1, pht004065.v1.p1, pht004032.v1.p1, pht004033.v1.p1, pht004034.v1.p1, pht004035.v1.p1, pht003266.v1.p1, pht000114.v2.p1, phg000079.v1
CARDIA CARe (phs000285.v3.p2, phs000613.v1.p2)	Coronary Artery Risk Development in Young Adults (CARDIA), Candidate Gene Association Resource (CARe)	National Heart, Lung, and Blood Institute (NHLBI)	Population-based	All	1985-ongoing	3440391, 3204420, 20400780	pht001583.v2.p2, pht001632.v2.p2, pht001671.v2.p2, pht001720.v2.p2, pht001785.v2.p2, pht001857.v2.p2, pht001589.v2.p2, pht001645.v2.p2, pht001697.v2.p2, pht001744.v2.p2, pht001804.v2.p2, pht001851.v2.p2, pht001557.v2.p2, pht001737.v2.p2, pht001799.v2.p2, pht001845.v2.p2, pht001569.v2.p2, pht001601.v2.p2, pht001656.v2.p2, pht001715.v2.p2, pht001761.v2.p2, pht001811.v2.p2, pht001555.v3.p2, phg000091.v2.p1, phg000092.v2.p1
CHS CARe (phs000287.v6.p1, phs000377.v5.p1)	Cardiovascular Health Study (CHS) Candidate Gene Association Resource (CARe)	National Heart, Lung, and Blood Institute (NHLBI)	Population-based	All	1988-1999	1669507, 8520709, 20400780	pht001452.v1.p1, pht001488.v1.p1, pht001489.v1.p1, pht001490.v1.p1, pht001491.v2.p1, pht001492.v1.p1, pht001493.v1.p1, pht001494.v1.p1, pht001495.v1.p1, pht001474.v1.p1, pht001475.v1.p1, phg000077.v1
EpiDREAM	Epidemiological follow-up study (Epi) of the Diabetes Reduction Assessment with Ramipril and Rosiglitazone Medication (DREAM) trial	Canadian Institutes of Health Research (CIHR), Aventis Pharma, GlaxoSmithKline, King Pharmaceuticals and Wyeth Ayerst	Randomized Clinical Trial and Observational Study	Controls and observational participants	2001-2006	15322749	
Framingham CARe (phs000007.v29.p10, phs000282.v18.p10)	Framingham Candidate Gene Association Resource (CARe)	National Heart, Lung, and Blood Institute (NHLBI)	Population-based, multi-generational	All	1948-ongoing	14025561, 1208363, 17372189	pht000009.v2.p10, pht000010.v3.p10, pht000011.v3.p10, pht000012.v3.p10, pht000013.v3.p10, pht000014.v3.p10, pht000015.v3.p10, pht000016.v3.p10, pht000017.v3.p10, pht000018.v4.p10, pht000019.v3.p10, pht000020.v3.p10, pht000021.v3.p10, pht000022.v4.p10, pht000023.v4.p10, pht000024.v4.p10, pht000025.v4.p10, pht000026.v3.p10, pht000027.v3.p10, pht000028.v3.p10, pht000029.v3.p10, pht000040.v4.p10, pht003099.v4.p10, pht000030.v7.p10, pht000031.v7.p10, pht000032.v6.p10, pht000033.v8.p10, pht000034.v7.p10, pht000035.v8.p10, pht000036.v8.p10, pht000747.v5.p10, pht000041.v6.p10, pht000074.v9.p10, pht000182.v12.p10, pht001415.v16.p10, pht000183.v12.p10, phg000076.v5
MESA CARe (phs000209.v13.p3, phs000283.v7.p3)	Multi-Ethnic Study of Atherosclerosis (MESA) Candidate Gene Association Resource (CARe)	National Heart, Lung, and Blood Institute (NHLBI)	Population-based	All	1999-2009	12397006, 20400780	pht001116.v10.p3, pht001118.v8.p3, pht001119.v8.p3, pht001120.v8.p3, pht003091.v3.p3, phg000081.v2
COPDGene (phs000179.v5.p2, phs000765.v2.p2)	Genetic Epidemiology of Chronic Obstructive Pulmonary Disease (COPDGene)	National Heart, Lung, and Blood Institute (NHLBI)	Case-Control	Controls only	2008-2011	19300482, 17446335	pht002239.v3.p2, pht002237.v2.p2, pht002238.v4.p2, phg000490.v1
eMERGE II (phs000888.v1.p1)	electronic Medical Records and Genomics (eMERGE) Network Imputed GWAS for 41 Phenotypes	National Human Genome Research Institute (NHGRI)	Population-based databank with electronic medical records	All		21269473, 21508311, 25566314	pht004678.v1.p1, pht004677.v1.p1, pht004680.v1.p1, pht005581.v1.p1, pht005587.v1.p1, phg000569.v1, phg000896.v1
WHI (phs000200.v10.p3, phs000746.v1.p3)	Women's Health Initiative (WHI) Harmonized and Imputed GWAS Data		Randomized Clinical Trial and Observational Study		1992-ongoing	9492970, 14575938, 7722207	pht000998.v5.p3, pht001019.v5.p3, pht000987.v5.p3, pht000998.v5.p3, phg000592.v1
WHIMS+ (WHI)	Women's Health Initiative Memory Study (WHIMS)	National Heart, Lung, and Blood Institute (NHLBI)	Cohort	All		9875839	
GARNET (WHI)	Genomics and Randomized Trials Network	National Human Genome Research Institute (NHGRI)	Case-Control	Controls only		22829776	
HIPFX (WHI)	Hip Fracture GWAS	National Heart, Lung, and Blood Institute (NHLBI)	Case-Control	Controls only		12859159	
MOPMAP (WHI)	Genetic Modification of PM-Mediated Arrhythmogenesis in Populations (MOPMAP)	National Institute of Environmental Health Sciences (NIEHS)	Case-Control	All		18979352	
GECCO (WHI)	Genome-wide Association Study for Non-synonymous SNPs in Colon Cancer	National Cancer Institute (NCI)	Case-Control	Controls only		20560206	

Panel B: Individual QC Information

Study	N	QC Notes	post-QC N	Mean CR
ARIC CARe (phs000280.v3.p1, phs000557.v2.p1)	14098	Sex Discordant (0), Call Rate < 97% (399), -3 < Het < 3sd (56), Duplicates/Relatives (476), Total (891)	13207	99.81
CARDIA CARe (phs000285.v3.p2, phs000613.v1.p2)	3282	Sex Discordant (0), Call Rate < 95% (486),-3 < Het < 3sd (10), Duplicates/Relatives (17), Total (508)	2774	97.07
CHS CARe (phs000287.v6.p1, phs000377.v5.p1)	5279	Sex Discordant (0), Call Rate < 95% (252),-3 < Het < 3sd (46), Duplicates/Relatives (0), Total (291)	4988	97.22
EpiDREAM	17423	Sex Discordant (9), Call Rate < 97% (67), -3 < Het < 3sd (378), Duplicates/Relatives (791), Total (1242)	15803	98.05
Framingham CARe (phs000007.v29.p10, phs000282.v18.p10)	7816	Sex Discordant (0), Call Rate < 95% (160), -3 < Het < 3sd (15), Duplicates/Relatives (5385), Total (5547)	2269	97.37
MESA CARe (phs000209.v13.p3, phs000283.v7.p3)	6351	Sex Discordant (3), Call Rate < 97% (72), -3 < Het < 3sd (19), Duplicates/Relatives (211), Total (291)	6060	99.79
COPDGene (phs000179.v5.p2, phs000765.v2.p2)	9716	Sex Discordant (0), Call Rate < 98% (0), -3 < Het < 3sd (55), Duplicates/Relatives (0), Cases (3578), Total (3619)	6097	99.86
eMERGE II (phs000888.v1.p1)	52572	Duplicates/Relatives (2709), Total (2709)	49863	98.42
WHI (phs000200.v10.p3, phs000746.v1.p3)	31806	Duplicates (1997), Relatives (309), Total (2306)	29500	
WHIMS+ (WHI)	5687		5667	99.17
GARNET (WHI)	4880		4869	99.17
HIPFX (WHI)	3688		3169	99.17
MOPMAP (WHI)	3068		2198	97.52
GECCO (WHI)	2491		2150	96.12

Table A3 BMI/Obesity-associated SNP information. (A) Detailed information on the BMI/obesity-associated SNPs from CARe studies including, effect alleles / other alleles (E/O), minor alleles (MA), minor allele frequency (MAF), call rate (CR) and Hardy-Weinberg Fisher's Exact p-value (HWE). Where Proxy SNP is indicated, the R^2 correlation to the original SNP is presented and all remaining details pertain to the proxy SNP. E/O for proxies were determined from phasing with the original SNP. (B) Same as (A) except for non-CARe studies.

Part A - Panel 1: QC of BMI SNPs in CARe Studies

SNP	Gene	Proxy SNP	R^2	Chr:Position	E/O	MA
rs2984618	TAL1	NA	NA	1:47690438	T/G	T
rs11208659	LEPR	rs11208662	0.97	1:65987164	C/G	C
rs7553158	TNNI3K	NA	NA	1:75005238	G/A	G
rs1780050	NEXN	NA	NA	1:78400540	A/C	C
rs347313	NOS1AP	rs980828	1	1:162306415	G/T	T
rs2819347	LMOD1	NA	NA	1:201884288	G/C	G
rs1561288	ADCY3	NA	NA	2:25369002	C/T	T
rs12617233	FANCL	NA	NA	2:59039998	C/T	T
rs492400	USP37	rs526134	0.99	2:219402371	G/A	G
rs2535633	ITIH4	NA	NA	3:52859630	G/C	G
rs17001561	SCARB2	NA	NA	4:77096118	A/G	A
rs6453133	HMGCR	NA	NA	5:74692776	A/G	G
rs6235	PCSK1	NA	NA	5:95728898	G/C	G
rs6232	PCSK1	NA	NA	5:95751785	C/T	C
rs9356744	CDKAL1	NA	NA	6:20685486	T/C	C
rs2272903	TFAP2B	NA	NA	6:50786571	G/A	A
rs9400239	FOXO3	rs4946932	1	6:108974746	C/A	A
rs1211166	NTRK2	NA	NA	9:87285992	A/G	G
rs11191560	NT5C2	rs3824755	0.92	10:104595849	C/G	C
rs7903146	TCF7L2	NA	NA	10:114758349	C/T	T
rs2237892	KCNQ1	rs2283228	0.92	11:2849530	C/A	C
rs10767664	BDNF	NA	NA	11:27725986	A/T	T
rs2856650	SPI1	rs11570094	0.97	11:47359706	A/C	A
rs7988412	MTIF3	NA	NA	13:28000282	T/C	T
rs1957894	PRKCH	rs10144353	1	14:61911157	T/C	T
rs997295	MAP2K5	NA	NA	15:68016343	T/G	G
rs8056711	IQCK	rs739564	0.95	16:19740237	A/G	G
rs4788099	SH2B1	NA	NA	16:28855727	G/A	G
rs749767	KAT8	NA	NA	16:31124407	A/G	G
rs1421085	FTO	NA	NA	16:53800954	C/T	C
rs6499653	FTO	NA	NA	16:53877592	T/C	T
rs1788826	NPC1	NA	NA	18:21154024	G/A	G
rs17066846	MC4R	NA	NA	18:58044818	G/T	G
rs11873305	MC4R	NA	NA	18:58049192	A/C	C
rs2075650	TOMM40	NA	NA	19:45395619	A/G	G
rs11672660	GIPR	NA	NA	19:46180184	C/T	T
rs2836754	ETS2	NA	NA	21:40291740	C/T	T

Part A - Panel 2: ARIC CARe

SNP	MAF	CR	HWE
rs2984618	0.392	98.8	0.871
rs11208659	0.095	100	0.764
rs7553158	0.438	99.9	0.433
rs1780050	0.428	100	0.6
rs347313	0.445	99.9	0.755
rs2819347	0.338	99.9	0.251
rs1561288	0.229	100	0.126
rs12617233	0.407	100	0.873
rs492400	0.414	100	0.711
rs2535633	0.382	99.6	0.744
rs17001561	0.166	100	0.063
rs6453133	0.306	99.8	0.762
rs6235	0.28	99.9	0.308
rs6232	0.055	100	0.02
rs9356744	0.308	97.7	1
rs2272903	0.103	100	0.266
rs9400239	0.306	100	0.809
rs1211166	0.205	100	0.875
rs11191560	0.106	99.9	0.222
rs7903146	0.299	100	0.501
rs2237892	0.078	97.6	1
rs10767664	0.213	99.9	0.759
rs2856650	0.297	99.9	0.538
rs7988412	0.2	99.9	0.936
rs1957894	0.092	100	1
rs997295	0.41	99.9	0.367
rs8056711	0.188	100	0.933
rs4788099	0.357	99.7	0.287
rs749767	0.37	100	0.956
rs1421085	0.405	100	0.012
rs6499653	0.27	97.8	0.018
rs1788826	0.353	100	0.311
rs17066846	0.192	99.7	0.135
rs11873305	0.042	100	1
rs2075650	0.143	99.7	0.833
rs11672660	0.218	100	0.327
rs2836754	0.364	99.1	0.084

Part A - Panel 3: CARDIA CARe

SNP	MAF	CR	HWE
rs2984618	0.392	98.8	0.871
rs11208659	0.095	100	0.764
rs7553158	0.438	99.9	0.433
rs1780050	0.428	100	0.6
rs347313	0.445	99.9	0.755
rs2819347	0.338	99.9	0.251
rs1561288	0.229	100	0.126
rs12617233	0.407	100	0.873
rs492400	0.414	100	0.711
rs2535633	0.382	99.6	0.744
rs17001561	0.166	100	0.063
rs6453133	0.306	99.8	0.762
rs6235	0.28	99.9	0.308
rs6232	0.055	100	0.02
rs9356744	0.308	97.7	1
rs2272903	0.103	100	0.266
rs9400239	0.306	100	0.809
rs1211166	0.205	100	0.875
rs11191560	0.106	99.9	0.222
rs7903146	0.299	100	0.501
rs2237892	0.078	97.6	1
rs10767664	0.213	99.9	0.759
rs2856650	0.297	99.9	0.538
rs7988412	0.2	99.9	0.936
rs1957894	0.092	100	1
rs997295	0.41	99.9	0.367
rs8056711	0.188	100	0.933
rs4788099	0.357	99.7	0.287
rs749767	0.37	100	0.956
rs1421085	0.405	100	0.012
rs6499653	0.27	97.8	0.018
rs1788826	0.353	100	0.311
rs17066846	0.192	99.7	0.135
rs11873305	0.042	100	1
rs2075650	0.143	99.7	0.833
rs11672660	0.218	100	0.327
rs2836754	0.364	99.1	0.084

Part A - Panel 4: CHS CARe

SNP	MAF	CR	HWE
rs2984618	0.394	99.7	0.674
rs11208659	0.098	100	0.335
rs7553158	0.44	100	0.684
rs1780050	0.427	100	0.156
rs347313	0.445	100	0.552
rs2819347	0.315	100	0.198
rs1561288	0.219	100	0.786
rs12617233	0.394	100	0.771
rs492400	0.434	100	0.051
rs2535633	0.394	99.5	0.746
rs17001561	0.164	100	0.31
rs6453133	0.304	99.3	0.187
rs6235	0.275	100	0.296
rs6232	0.052	100	0.428
rs9356744	0.323	97.2	0.032
rs2272903	0.109	100	0.342
rs9400239	0.298	100	0.825
rs1211166	0.203	100	0.536
rs11191560	0.103	100	0.559
rs7903146	0.305	100	0.512
rs2237892	0.068	99.4	0.389
rs10767664	0.226	99.9	0.309
rs2856650	0.315	99.8	0.352
rs7988412	0.177	100	0.525
rs1957894	0.088	100	0.848
rs997295	0.428	100	0.9
rs8056711	0.184	100	0.918
rs4788099	0.377	99.9	0.599
rs749767	0.398	100	0.384
rs1421085	0.413	100	0.065
rs6499653	0.259	97.3	0.087
rs1788826	0.355	100	0.266
rs17066846	0.189	99.3	0.39
rs11873305	0.04	100	0.22
rs2075650	0.122	100	0.249
rs11672660	0.213	100	0.712
rs2836754	0.368	99.8	0.921

Part A - Panel 5: EpiDREAM CARe

SNP	MAF	CR	HWE
rs2984618	0.398	100	0.222
rs11208659	0.098	100	0.812
rs7553158	0.44	100	0.949
rs1780050	0.419	100	0.492
rs347313	0.455	100	0.302
rs2819347	0.33	99.9	0.394
rs1561288	NA	NA	NA
rs12617233	0.403	100	0.587
rs492400	NA	NA	NA
rs2535633	0.407	100	0.745
rs17001561	NA	NA	NA
rs6453133	0.302	100	0.234
rs6235	0.265	100	0.284
rs6232	0.049	100	1
rs9356744	0.331	100	0.795
rs2272903	0.107	100	0.382
rs9400239	0.319	100	0.597
rs1211166	0.193	100	0.347
rs11191560	0.1	100	0.223
rs7903146	0.31	100	0.017
rs2237892	0.073	100	0.087
rs10767664	0.216	99.9	0.241
rs2856650	0.309	100	0.864
rs7988412	NA	NA	NA
rs1957894	NA	NA	NA
rs997295	0.409	100	0.352
rs8056711	0.182	99.9	0.361
rs4788099	NA	NA	NA
rs749767	0.391	100	0.843
rs1421085	0.432	100	0.014
rs6499653	0.26	99.8	0.264
rs1788826	0.349	100	0.102
rs17066846	0.183	99.9	0.575
rs11873305	0.039	100	0.329
rs2075650	0.139	100	0.896
rs11672660	0.209	100	0.681
rs2836754	NA	NA	NA

Part A - Panel 6: Framingham CARe

SNP	MAF	CR	HWE
rs2984618	0.399	100	0.096
rs11208659	0.08	100	0.772
rs7553158	0.443	100	0.172
rs1780050	0.413	100	0.434
rs347313	0.454	100	0.799
rs2819347	0.331	100	0.703
rs1561288	0.22	100	0.328
rs12617233	0.399	100	0.793
rs492400	0.378	100	0.72
rs2535633	0.364	100	0.856
rs17001561	0.176	100	0.885
rs6453133	0.286	99.9	0.836
rs6235	0.278	100	1
rs6232	0.051	100	1
rs9356744	0.313	99.2	0.277
rs2272903	0.112	100	0.832
rs9400239	0.328	100	0.392
rs1211166	0.226	100	0.81
rs11191560	0.09	100	0.434
rs7903146	0.319	100	0.846
rs2237892	0.069	100	0.505
rs10767664	0.225	100	1
rs2856650	0.324	100	0.386
rs7988412	0.18	100	0.668
rs1957894	0.085	100	1
rs997295	0.434	100	0.04
rs8056711	0.191	99.8	0.338
rs4788099	0.388	100	0.929
rs749767	0.408	100	0.727
rs1421085	0.434	100	0.346
rs6499653	0.252	100	0.371
rs1788826	0.372	100	0.719
rs17066846	0.188	99.7	0.891
rs11873305	0.042	100	0.616
rs2075650	0.111	100	0.667
rs11672660	0.198	100	0.185
rs2836754	0.379	100	0.421

Part A - Panel 7: MESA CARe

SNP	MAF	CR	HWE
rs2984618	0.38	99.8	1
rs11208659	0.094	100	0.72
rs7553158	0.448	100	0.367
rs1780050	0.426	100	0.199
rs347313	0.458	100	0.165
rs2819347	0.316	100	0.925
rs1561288	0.217	100	0.72
rs12617233	0.4	100	1
rs492400	0.425	100	0.384
rs2535633	0.396	100	0.082
rs17001561	0.167	100	0.189
rs6453133	0.306	99.8	0.063
rs6235	0.261	100	0.834
rs6232	0.042	100	0.798
rs9356744	0.335	96.9	0.963
rs2272903	0.124	100	0.926
rs9400239	0.29	100	3.67E-03
rs1211166	0.194	100	0.218
rs11191560	0.106	100	0.669
rs7903146	0.296	100	0.355
rs2237892	0.079	99.6	0.889
rs10767664	0.224	99.2	1.88E-03
rs2856650	0.305	99.7	0.114
rs7988412	0.172	99.9	0.522
rs1957894	0.094	100	0.905
rs997295	0.442	99.9	0.118
rs8056711	0.193	100	0.697
rs4788099	0.371	100	0.318
rs749767	0.392	100	0.865
rs1421085	0.417	100	0.134
rs6499653	0.254	99.1	0.668
rs1788826	0.358	100	0.217
rs17066846	0.184	99.5	1
rs11873305	0.041	100	0.186
rs2075650	0.136	100	0.389
rs11672660	0.214	100	0.904
rs2836754	0.38	100	0.112

Part B - COPDGene

SNP	Gene	Proxy SNP	R^2	Chr:Position	E/O	MA	MAF	CR	HWE
rs2984618	TAL1	rs977747	0.99	1:47684677	T/G	T	0.395	99.7	0.76
rs11208659	LEPR	NA	NA	1:65979280	C/T	C	0.095	98.8	0.569
rs7553158	TNNI3K	NA	NA	1:75005238	G/A	G	0.448	100	0.844
rs1780050	NEXN	rs11162405	0.99	1:78469660	A/G	G	0.41	99.9	0.638
rs347313	NOS1AP	rs347306	1	1:162302635	C/T	T	0.438	100	0.322
rs2819347	LMOD1	rs2820312	0.94	1:201869257	A/G	A	0.318	100	0.525
rs1561288	ADCY3	rs6718083	1	2:25362194	G/A	A	0.225	100	0.163
rs12617233	FANCL	NA	NA	2:59039998	C/T	T	0.406	100	0.686
rs492400	USP37	NA	NA	2:219349752	C/T	C	0.425	99.3	0.894
rs2535633	ITIH4	rs2256332	0.94	3:52855865	A/G	A	0.392	100	0.233
rs17001561	SCARB2	NA	NA	4:77096118	A/G	A	0.165	99.9	0.376
rs6453133	HMGCR	NA	NA	5:74692776	A/G	G	0.295	99.8	0.369
rs6235	PCSK1	rs6234	0.99	5:95728974	C/G	C	0.273	99	0.538
rs6232	PCSK1	NA	NA	5:95751785	C/T	C	0.049	99.8	0.727
rs9356744	CDKAL1	rs9350271	1	6:20683164	G/A	NA	NA	NA	NA
rs2272903	TFAP2B	NA	NA	6:50786571	G/A	A	0.105	100	0.729
rs9400239	FOXO3	NA	NA	6:108977663	C/T	NA	NA	NA	NA
rs1211166	NTRK2	rs1147199	0.97	9:87275895	G/A	A	0.197	99.6	0.027
rs11191560	NT5C2	NA	NA	10:104869038	C/T	C	0.084	100	0.045
rs7903146	TCF7L2	NA	NA	10:114758349	C/T	T	0.293	99.9	0.273
rs2237892	KCNQ1	NA	NA	11:2839751	T/C	T	0.066	99.8	0.792
rs10767664	BDNF	rs11030104	0.9	11:27684517	A/G	G	0.202	100	0.451
rs2856650	SPI1	rs11570094	0.97	11:47359706	A/C	A	0.293	99.8	0.256
rs7988412	MTIF3	rs10220056	0.92	13:28003781	G/T	NA	NA	NA	NA
rs1957894	PRKCH	rs1957895	0.97	14:61908332	G/T	G	0.088	99.8	0.687
rs997295	MAP2K5	NA	NA	15:68016343	T/G	G	0.4	99.7	0.839
rs8056711	IQCK	rs950928	1	16:19824638	T/C	C	0.176	99.9	0.955
rs4788099	SH2B1	NA	NA	16:28855727	G/A	G	0.386	100	0.81
rs749767	KAT8	rs9925964	0.95	16:31129895	A/G	NA	NA	NA	NA
rs1421085	FTO	NA	NA	16:53800954	C/T	C	0.414	99.9	0.547
rs6499653	FTO	NA	NA	16:53877592	T/C	T	0.257	98.5	0.113
rs1788826	NPC1	rs1429934	0.95	18:21162288	C/T	C	0.346	99.7	0.078
rs17066846	MC4R	rs17773774	0.94	18:58060126	A/C	A	0.196	99.9	0.027
rs11873305	MC4R	NA	NA	18:58049192	A/C	C	0.037	100	0.174
rs2075650	TOMM40	NA	NA	19:45395619	A/G	G	0.133	100	0.833
rs11672660	GIPR	NA	NA	19:46180184	C/T	T	0.206	100	0.655
rs2836754	ETS2	NA	NA	21:40291740	C/T	T	0.375	100	0.225

Table A4: Height-associated SNP information. (A) Detailed information on the height-associated SNPs from CARe studies including, effect alleles / other alleles (E/O), minor alleles (MA), minor allele frequency (MAF), call rate (CR) and Hardy-Weinberg Fisher's Exact p-value (HWE). Where Proxy SNP is indicated, the R2 correlation to the original SNP is presented and all remaining details pertain to the proxy SNP. E/O for proxies were determined from phasing with the original SNP. (B) Same as (A) except for non-CARe studies.

Part A - Panel 1: QC of Height SNPs in CARe Studies

SNP	Gene	Proxy SNP	R^2	Chr:Position	E/O	MA
rs451061	PRKCZ	NA	NA	1:2075068	C/G	C
rs212517	ECE1	NA	NA	1:21577159	A/T	T
rs1738475	HTR1D	NA	NA	1:23536891	C/G	G
rs2229712	RPS6KA1	NA	NA	1:26883511	A/C	NA
rs17106235	FAF1	NA	NA	1:50943370	G/C	G
rs551219	COL24A1	NA	NA	1:86519721	T/C	T
rs12145922	PKN2	NA	NA	1:89146234	A/C	C
rs660240	PSRC1	rs602633	0.94	1:109821511	T/G	T
rs7522692	PIGC	NA	NA	1:172436835	G/A	G
rs1342586	TGFB2	NA	NA	1:218597859	T/C	T
rs10185680	MFSD2B	NA	NA	2:24275306	G/A	A
rs1866146	POMC	NA	NA	2:25380573	G/A	G
rs780094	GCKR	NA	NA	2:27741237	C/T	T
rs7557989	THADA	NA	NA	2:43630657	T/C	T
rs1822469	PPP3R1	NA	NA	2:68454685	C/T	T
rs6731022	EIF2AK3	NA	NA	2:88917035	C/G	C
rs3821009	PDE11A	NA	NA	2:178682471	T/C	T
rs6718902	STAT1	NA	NA	2:191838204	T/C	T
rs6758561	NOP58	rs2176167	0.98	2:203100918	C/T	C
rs526134	USP37	NA	NA	2:219402371	A/G	G
rs10208728	IHH	NA	NA	2:219917303	A/G	G
rs3100776	IHH	NA	NA	2:219921200	C/T	C
rs4973410	NCL	NA	NA	2:232331734	C/T	T
rs2679178	NPPC	NA	NA	2:232797861	C/T	T
rs7578199	HDLBP	NA	NA	2:242192848	T/C	C
rs7572476	BOK	NA	NA	2:242496325	C/T	T
rs2450855	MKRN2	rs2633442	0.91	3:12609937	G/A	A
rs9857730	VILL	NA	NA	3:38051941	C/T	C
rs3915129	CTNNB1	rs13076290	0.99	3:41260369	T/C	T
rs490634	CISH	NA	NA	3:50640830	C/T	NA
rs13072536	ITIH4	NA	NA	3:52861211	A/T	T
rs4955526	EPHB1	NA	NA	3:134317337	C/T	T
rs9844666	PCCB	NA	NA	3:135974216	G/A	A

Continued on next page

Table A4 – *Continued from previous page*

SNP	Gene	Proxy SNP	R^2	Chr:Position	E/O	MA
rs572169	GHSR	NA	NA	3:172165727	T/C	T
rs17754	RFC1	NA	NA	4:39289308	C/G	G
rs17472113	ZAR1	NA	NA	4:48495662	A/T	NA
rs4864548	CLOCK	NA	NA	4:56413803	A/G	A
rs3796529	REST	NA	NA	4:57797414	T/C	T
rs12503378	NUDT6	NA	NA	4:123810734	C/G	G
rs17541471	NPR3	NA	NA	5:32755589	C/T	C
rs6180	GHR	NA	NA	5:42719239	A/C	C
rs832575	MAP3K1	NA	NA	5:56161787	A/G	G
rs41132	AP3B1	NA	NA	5:77408842	A/C	C
rs2247870	GPR98	NA	NA	5:90151589	A/G	G
rs17085675	PCSK1	NA	NA	5:95727664	T/A	T
rs17622208	SLC22A5	NA	NA	5:131717050	A/G	A
rs9366637	HFE	NA	NA	6:26089098	C/T	T
rs2853977	HCP5	NA	NA	6:31379304	A/T	A
rs2229642	ITPR3	NA	NA	6:33659472	G/C	NA
rs1776897	HMGA1	NA	NA	6:34195011	G/T	G
rs4946932	FOXO3	NA	NA	6:108974746	C/A	A
rs1476387	PPIL6	NA	NA	6:109764535	G/T	T
rs7756224	NMBR	NA	NA	6:142406840	C/T	T
rs2234693	ESR1	NA	NA	6:152163335	C/T	C
rs2982712	ESR1	NA	NA	6:152358179	C/T	C
rs1074287	OPRM1	rs510769	0.97	6:154362019	T/C	T
rs1636255	GNA12	NA	NA	7:2892804	C/A	A
rs864745	JAZF1	NA	NA	7:28180556	T/C	C
rs3812265	CNOT4	NA	NA	7:135048804	T/C	T
rs1800783	NOS3	NA	NA	7:150689397	T/A	A
rs6999671	RPS20	rs7004280	0.9	8:56894350	C/G	C
rs2145923	NPR2	NA	NA	9:35788239	C/T	C
rs3814115	PCSK5	NA	NA	9:78504729	C/T	C
rs7853859	CENPP	NA	NA	9:95151377	T/C	C
rs3739707	LPAR1	NA	NA	9:113792706	C/A	A
rs7020782	PAPPA	NA	NA	9:119106881	A/C	C
rs803932	ASTN2	NA	NA	9:119458020	C/T	C
rs11102986	RXRA	NA	NA	9:137285503	G/A	A
rs291979	GRK5	NA	NA	10:121129797	A/G	A
rs2735469	MRPL23	NA	NA	11:2022804	A/G	A
rs4320932	IGF2	NA	NA	11:2171601	T/C	C
rs900147	ARNTL	rs1481892	0.98	11:13301921	G/C	G
rs948847	APLNR	rs10736682	0.91	11:57008536	G/A	G
rs174547	FADS1	rs1535	0.97	11:61597972	A/G	G
rs3736228	LRP5	NA	NA	11:68201295	C/T	T

Continued on next page

Table A4 – *Continued from previous page*

SNP	Gene	Proxy SNP	R^2	Chr:Position	E/O	MA
rs7396866	NEU3	rs12225387	0.9	11:74604767	G/A	G
rs674424	ABCG4	NA	NA	11:119030752	T/C	T
rs2282537	POU2F3	NA	NA	11:120187971	G/A	A
rs6487088	PDE3A	NA	NA	12:20588382	T/C	C
rs7137534	PDE3A	NA	NA	12:20831777	T/C	T
rs2066807	PAN2	NA	NA	12:56740682	G/C	G
rs2291617	METTL1	NA	NA	12:58166403	T/G	G
rs1042725	HMGA2	NA	NA	12:66358347	C/T	T
rs3782415	SOCS2	NA	NA	12:93967755	C/T	C
rs6219	IGF1	NA	NA	12:102790192	T/C	T
rs10861148	HSP90B1	NA	NA	12:104340080	A/C	A
rs907482	KNTC1	rs7963565	0.94	12:122703014	T/C	T
rs1051431	MPHOSPH9	NA	NA	12:123645803	G/A	G
rs1950500	NFATC4	NA	NA	14:24830850	T/C	T
rs696	NFKBIA	NA	NA	14:35871093	C/T	T
rs709939	SAMD4A	NA	NA	14:55249345	T/C	C
rs3783937	FBLN5	NA	NA	14:92407693	C/T	T
rs1036477	FBN1	NA	NA	15:48914926	A/G	G
rs12050767	CYP19A1	NA	NA	15:51557257	C/T	C
rs7163907	PTPN9	NA	NA	15:75845097	C/T	C
rs17599989	SEC11A	rs1051168	0.91	15:85200520	T/G	T
rs1516796	ACAN	NA	NA	15:89353798	A/C	A
rs8033670	IGF1R	rs8038415	0.96	15:99499434	C/T	C
rs5015437	LMF1	NA	NA	16:987371	A/G	A
rs258281	RAB26	NA	NA	16:2191734	G/A	A
rs2023693	ERI2	rs9930741	0.99	16:20695486	T/C	C
rs8055190	LRRC36	NA	NA	16:67391618	C/T	T
rs7359336	NFAT5	NA	NA	16:69733460	G/A	G
rs8071847	POLR2A	NA	NA	17:7407327	G/A	G
rs2909430	TP53	NA	NA	17:7578645	T/C	C
rs11080149	NF1	NA	NA	17:29623288	T/C	T
rs2715553	RARA	NA	NA	17:38496320	A/G	G
rs752313	EZH1	NA	NA	17:40901824	C/T	T
rs12603813	PLCD3	NA	NA	17:43196584	T/C	C
rs12603582	ITGB3	NA	NA	17:45377577	G/T	T
rs46522	UBE2Z	NA	NA	17:46988597	C/T	C
rs9892365	TBX2	NA	NA	17:59491384	A/G	A
rs12940055	MAP3K3	NA	NA	17:61722142	C/T	T
rs2854207	CSH2	NA	NA	17:61947107	G/C	G
rs2053156	GRB2	NA	NA	17:73378440	T/G	G
rs25656	NFATC1	NA	NA	18:77227476	A/G	NA
rs891088	INSR	NA	NA	19:7184762	G/A	G

Continued on next page

Table A4 – *Continued from previous page*

SNP	Gene	Proxy SNP	R^2	Chr:Position	E/O	MA
rs4808199	GATAD2A	NA	NA	19:19545099	G/A	A
rs4803520	GRIK5	NA	NA	19:42500373	G/A	A
rs2682552	XRCC1	NA	NA	19:44069741	A/T	A
rs158676	CDK5RAP1	NA	NA	20:31974395	A/G	G
rs2425019	MMP24	NA	NA	20:33819415	G/A	G
rs1780616	LBP	NA	NA	20:36972942	C/T	T
rs11086538	MC3R	rs6127698	0.95	20:54823416	G/T	T
rs2057291	GNAS	NA	NA	20:57472043	A/G	A

Part A - Panel 2: ARIC CARe

SNP	MAF	CR	HWE
rs451061	0.385	100	0.377
rs212517	0.402	99.9	0.351
rs1738475	0.412	100	0.622
rs2229712	NA	NA	NA
rs17106235	0.093	99.3	0.174
rs551219	0.293	98.1	5.76E-04
rs12145922	0.43	98.3	0.346
rs660240	0.219	99.9	1
rs7522692	0.216	99.9	0.597
rs1342586	0.217	99.9	0.953
rs10185680	0.465	99.4	0.547
rs1866146	0.339	99.9	0.578
rs780094	0.4	100	0.3
rs7557989	0.33	100	0.115
rs1822469	0.4	99.8	0.081
rs6731022	0.343	100	0.233
rs3821009	0.08	99.8	0.685
rs6718902	0.244	100	0.117
rs6758561	0.348	100	0.228
rs526134	0.427	100	0.173
rs10208728	0.101	100	0.35
rs3100776	0.042	99.8	1
rs4973410	0.47	99.9	0.646
rs2679178	0.082	100	0.693
rs7578199	0.242	100	0.175
rs7572476	0.456	100	0.779
rs2450855	0.437	100	0.642

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs9857730	0.212	99.5	0.742
rs3915129	0.452	98.8	0.258
rs490634	NA	NA	NA
rs13072536	0.233	99.9	0.759
rs4955526	0.354	100	0.811
rs9844666	0.238	99.8	0.891
rs572169	0.312	99.9	1
rs17754	0.429	100	0.382
rs17472113	NA	NA	NA
rs4864548	0.364	100	0.683
rs3796529	0.195	98.2	0.823
rs12503378	0.168	100	0.212
rs17541471	0.196	100	0.393
rs6180	0.464	100	0.459
rs832575	0.119	99.8	0.037
rs41132	0.241	99.3	0.623
rs2247870	0.458	100	0.378
rs17085675	0.29	100	0.79
rs17622208	0.462	100	0.888
rs9366637	0.064	100	0.507
rs2853977	0.454	99.8	0.047
rs2229642	NA	NA	NA
rs1776897	0.085	100	0.18
rs4946932	0.305	100	0.1
rs1476387	0.417	100	0.566
rs7756224	0.426	100	0.166
rs2234693	0.456	100	0.092
rs2982712	0.434	100	0.037
rs1074287	0.26	100	0.127
rs1636255	0.289	99.8	1
rs864745	0.499	100	0.349
rs3812265	0.248	100	0.04
rs1800783	0.379	100	0.409
rs6999671	0.034	100	0.049
rs2145923	0.178	100	0.76
rs3814115	0.317	100	0.301
rs7853859	0.371	99.4	0.638
rs3739707	0.247	100	0.284
rs7020782	0.303	100	0.724
rs803932	0.333	99.3	0.15
rs11102986	0.178	99.7	0.919
rs291979	0.229	100	0.978

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs2735469	0.156	100	0.162
rs4320932	0.2	99.6	0.366
rs900147	0.292	99.9	0.289
rs948847	0.432	100	6.08E-04
rs174547	0.334	100	0.42
rs3736228	0.147	99.6	0.131
rs7396866	0.281	100	0.961
rs674424	0.256	100	1
rs2282537	0.141	100	0.15
rs6487088	0.196	100	0.507
rs7137534	0.328	100	0.223
rs2066807	0.067	99.8	0.751
rs2291617	0.335	97.7	0.542
rs1042725	0.492	100	0.437
rs3782415	0.203	100	0.622
rs6219	0.101	100	0.87
rs10861148	0.105	99.9	0.312
rs907482	0.343	100	0.691
rs1051431	0.216	100	0.207
rs1950500	0.293	100	0.414
rs696	0.372	99.5	0.624
rs709939	0.443	100	0.84
rs3783937	0.239	100	0.477
rs1036477	0.105	100	0.09
rs12050767	0.44	100	0.984
rs7163907	0.242	98.8	0.785
rs17599989	0.273	99.9	0.209
rs1516796	0.48	100	0.106
rs8033670	0.498	99.9	0.705
rs5015437	0.382	98.4	0.718
rs258281	0.183	99.8	0.594
rs2023693	0.412	99.3	0.55
rs8055190	0.044	100	0.474
rs7359336	0.425	100	0.555
rs8071847	0.213	100	0.423
rs2909430	0.131	99.4	0.19
rs11080149	0.132	100	0.761
rs2715553	0.445	100	0.242
rs752313	0.479	100	0.14
rs12603813	0.247	98.8	0.346
rs12603582	0.223	99.9	0.841
rs46522	0.466	99.4	0.065

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs9892365	0.33	100	0.544
rs12940055	0.113	100	0.082
rs2854207	0.273	100	0.16
rs2053156	0.171	99.9	9.22E-03
rs25656	NA	NA	NA
rs891088	0.262	100	0.625
rs4808199	0.183	100	0.258
rs4803520	0.116	100	0.559
rs2682552	0.19	100	0.155
rs158676	0.313	100	0.594
rs2425019	0.463	100	0.873
rs1780616	0.343	100	0.27
rs11086538	0.492	99.7	0.216
rs2057291	0.336	100	0.894

Part A - Panel 3: CARDIA CARe

SNP	MAF	CR	HWE
rs451061	0.391	100	0.106
rs212517	0.397	99.9	0.788
rs1738475	0.426	100	0.752
rs2229712	NA	NA	NA
rs17106235	0.108	98.7	0.789
rs551219	0.29	98.8	0.66
rs12145922	0.412	97.6	0.334
rs660240	0.227	99.8	0.341
rs7522692	0.205	99.9	0.156
rs1342586	0.219	99.8	1
rs10185680	0.442	98.5	0.227
rs1866146	0.349	100	0.611
rs780094	0.403	99.6	0.218
rs7557989	0.343	100	0.531
rs1822469	0.4	99.9	0.362
rs6731022	0.33	99.9	0.352
rs3821009	0.073	99.9	0.701
rs6718902	0.246	100	1
rs6758561	0.347	100	0.461
rs526134	0.414	100	0.711
rs10208728	0.101	99.9	0.015

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs3100776	0.039	97.5	0.021
rs4973410	0.469	99.9	0.034
rs2679178	0.087	100	0.254
rs7578199	0.251	100	0.452
rs7572476	0.452	100	0.049
rs2450855	0.418	99.9	0.751
rs9857730	0.2	99.6	0.421
rs3915129	0.437	96.6	0.202
rs490634	0.132	99.4	0.262
rs13072536	0.239	99.9	0.832
rs4955526	0.35	100	0.337
rs9844666	0.241	99.1	0.138
rs572169	0.287	99.9	0.802
rs17754	0.448	99.9	0.35
rs17472113	0.195	86	0.659
rs4864548	0.379	100	1
rs3796529	0.182	98.2	0.163
rs12503378	0.171	100	0.928
rs17541471	0.201	100	0.936
rs6180	0.451	99.9	0.835
rs832575	0.125	99.9	1
rs41132	0.242	99.6	0.779
rs2247870	0.446	100	0.835
rs17085675	0.297	99.9	0.296
rs17622208	0.472	99.7	0.679
rs9366637	0.063	99.7	0.829
rs2853977	0.424	99.7	0.833
rs2229642	NA	0	1
rs1776897	0.097	99.2	0.46
rs4946932	0.306	100	0.809
rs1476387	0.433	100	0.753
rs7756224	0.42	100	0.246
rs2234693	0.467	100	0.109
rs2982712	0.428	100	0.528
rs1074287	0.246	100	0.406
rs1636255	0.285	100	0.45
rs864745	0.495	100	0.472
rs3812265	0.239	100	0.525
rs1800783	0.369	99.9	1
rs6999671	0.036	100	0.719
rs2145923	0.189	100	0.867
rs3814115	0.331	100	0.523

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs7853859	0.377	99.3	0.51
rs3739707	0.254	100	0.309
rs7020782	0.299	99.9	0.806
rs803932	0.308	99.1	0.364
rs11102986	0.191	99.1	0.055
rs291979	0.211	100	0.644
rs2735469	0.161	100	0.634
rs4320932	0.21	95	0.874
rs900147	0.295	100	1
rs948847	0.422	100	0.792
rs174547	0.319	99.9	0.315
rs3736228	0.135	99.7	0.66
rs7396866	0.27	100	0.031
rs674424	0.251	100	0.413
rs2282537	0.143	100	0.296
rs6487088	0.189	100	0.315
rs7137534	0.311	100	0.904
rs2066807	0.06	99.7	0.111
rs2291617	0.339	96.7	0.382
rs1042725	0.495	100	0.644
rs3782415	0.196	100	0.253
rs6219	0.099	99.9	1
rs10861148	0.112	99.9	0.897
rs907482	0.336	100	0.454
rs1051431	0.228	100	0.381
rs1950500	0.277	99.8	0.653
rs696	0.374	99.8	0.442
rs709939	0.482	100	1
rs3783937	0.254	100	0.839
rs1036477	0.112	100	0.699
rs12050767	0.449	100	0.64
rs7163907	0.246	96.6	0.888
rs17599989	0.26	99.9	0.23
rs1516796	0.482	100	0.837
rs8033670	0.484	100	0.757
rs5015437	0.385	98.2	1
rs258281	0.181	99.7	0.862
rs2023693	0.436	99.1	0.529
rs8055190	0.043	100	5.66E-03
rs7359336	0.419	99.9	0.874
rs8071847	0.219	100	0.822
rs2909430	0.125	97.8	0.189

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs11080149	0.125	100	0.814
rs2715553	0.454	100	0.12
rs752313	0.488	100	0.607
rs12603813	0.235	99.7	1
rs12603582	0.228	100	0.011
rs46522	0.499	94.2	0.072
rs9892365	0.325	100	0.953
rs12940055	0.1	100	0.474
rs2854207	0.293	99.7	8.97E-03
rs2053156	0.194	99.9	0.138
rs25656	NA	NA	NA
rs891088	0.262	100	0.073
rs4808199	0.168	100	0.067
rs4803520	0.121	100	0.228
rs2682552	0.198	100	0.935
rs158676	0.338	100	1
rs2425019	0.441	99.9	0.23
rs1780616	0.355	100	0.955
rs11086538	0.476	99.8	0.502
rs2057291	0.345	100	0.955

Part A - Panel 4: CHS CARe

SNP	MAF	CR	HWE
rs451061	0.373	100	1
rs212517	0.412	99.8	0.252
rs1738475	0.414	99.9	0.427
rs2229712	NA	NA	NA
rs17106235	0.102	99	0.126
rs551219	0.278	98.7	0.877
rs12145922	0.427	96.2	0.034
rs660240	0.21	99.7	0.429
rs7522692	0.203	99.8	0.418

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs1342586	0.218	99.9	0.821
rs10185680	0.475	99.2	0.78
rs1866146	0.329	99.6	0.442
rs780094	0.42	99.9	0.825
rs7557989	0.33	100	0.531
rs1822469	0.404	99.7	0.387
rs6731022	0.328	99.9	0.133
rs3821009	0.081	99.9	0.214
rs6718902	0.251	100	0.049
rs6758561	0.356	100	0.662
rs526134	0.434	100	0.041
rs10208728	0.103	99.9	0.803
rs3100776	0.037	99.5	0.661
rs4973410	0.474	100	0.017
rs2679178	0.09	100	0.509
rs7578199	0.248	100	0.934
rs7572476	0.462	100	0.926
rs2450855	0.427	100	0.469
rs9857730	0.21	100	0.609
rs3915129	0.438	98.1	0.975
rs490634	0.116	98.6	1
rs13072536	0.236	99.4	0.764
rs4955526	0.357	100	0.867
rs9844666	0.23	99.6	0.24

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs572169	0.305	100	0.586
rs17754	0.443	100	0.876
rs17472113	0.207	88.2	9.21E-05
rs4864548	0.366	100	0.666
rs3796529	0.186	97.1	0.098
rs12503378	0.174	100	0.519
rs17541471	0.198	100	0.012
rs6180	0.462	100	0.369
rs832575	0.129	100	1
rs41132	0.236	99.3	0.898
rs2247870	0.476	100	0.805
rs17085675	0.291	100	0.126
rs17622208	0.467	99.6	0.226
rs9366637	0.069	99.9	0.185
rs2853977	0.42	99.9	5.84E-05
rs2229642	NA	0	1
rs1776897	0.093	99.7	0.36
rs4946932	0.298	100	0.796
rs1476387	0.429	100	0.208
rs7756224	0.428	100	0.22
rs2234693	0.458	100	0.71
rs2982712	0.415	100	0.172
rs1074287	0.248	100	0.967
rs1636255	0.287	99.9	0.547

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs864745	0.496	99.9	0.926
rs3812265	0.246	100	0.051
rs1800783	0.39	98	0.432
rs6999671	0.036	100	0.27
rs2145923	0.19	100	0.134
rs3814115	0.336	100	0.836
rs7853859	0.388	99.3	0.871
rs3739707	0.257	100	0.628
rs7020782	0.302	100	0.144
rs803932	0.333	97.8	0.551
rs11102986	0.185	99.3	0.041
rs291979	0.239	100	0.641
rs2735469	0.157	99.6	0.77
rs4320932	0.212	98.8	0.61
rs900147	0.292	100	0.823
rs948847	0.418	100	0.975
rs174547	0.325	100	0.276
rs3736228	0.141	99.9	0.408
rs7396866	0.286	100	0.029
rs674424	0.249	100	0.174
rs2282537	0.143	100	0.38
rs6487088	0.204	100	0.184
rs7137534	0.314	100	0.198
rs2066807	0.061	99.6	1

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs2291617	0.312	99.6	0.943
rs1042725	0.476	100	0.194
rs3782415	0.204	100	0.506
rs6219	0.104	100	0.62
rs10861148	0.117	100	0.551
rs907482	0.336	100	0.604
rs1051431	0.226	99.8	0.454
rs1950500	0.277	99.9	0.281
rs696	0.38	99.4	0.768
rs709939	0.443	100	0.779
rs3783937	0.262	100	0.402
rs1036477	0.111	100	0.815
rs12050767	0.446	100	0.803
rs7163907	0.27	98.6	0.107
rs17599989	0.261	100	0.472
rs1516796	0.468	100	0.337
rs8033670	0.498	100	0.498
rs5015437	0.373	99.4	0.741
rs258281	0.198	99.8	0.12
rs2023693	0.419	99.5	0.241
rs8055190	0.048	100	0.178
rs7359336	0.416	99.9	0.657
rs8071847	0.21	100	0.745
rs2909430	0.137	99.1	0.896

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs11080149	0.121	100	0.31
rs2715553	0.451	100	0.152
rs752313	0.491	100	0.805
rs12603813	0.248	99.8	0.773
rs12603582	0.215	99.9	0.235
rs46522	0.476	98.6	0.455
rs9892365	0.331	100	0.676
rs12940055	0.124	100	0.832
rs2854207	0.274	100	0.907
rs2053156	0.195	99.9	0.018
rs25656	NA	NA	NA
rs891088	0.26	100	0.016
rs4808199	0.172	100	0.871
rs4803520	0.124	100	0.943
rs2682552	0.194	100	0.623
rs158676	0.308	100	0.587
rs2425019	0.461	99.9	0.264
rs1780616	0.355	99.9	0.788
rs11086538	0.474	99.9	0.951
rs2057291	0.342	100	0.811

Part A - Panel 5: EpiDREAM

SNP	MAF	CR	HWE
rs451061	0.384	100	0.192
rs212517	0.408	99.9	0.88
rs1738475	NA	NA	NA
rs2229712	NA	NA	NA
rs17106235	NA	NA	NA
rs551219	0.292	97.8	0.026
rs12145922	0.417	99.6	0.401
rs660240	NA	NA	NA
rs7522692	NA	NA	NA
rs1342586	0.229	99.9	0.459
rs10185680	0.454	97.8	0.418
rs1866146	0.349	100	0.322
rs780094	0.415	100	0.667
rs7557989	NA	NA	NA
rs1822469	0.416	99.7	0.203
rs6731022	0.318	99.9	0.682
rs3821009	0.086	100	0.083
rs6718902	0.244	99.9	0.977
rs6758561	0.339	99.4	0.574
rs526134	NA	NA	NA
rs10208728	NA	NA	NA
rs3100776	NA	NA	NA
rs4973410	NA	NA	NA
rs2679178	0.087	100	0.646
rs7578199	0.255	100	0.458
rs7572476	0.469	100	0.231
rs2450855	NA	NA	NA
rs9857730	0.208	98.4	8.74E-03
rs3915129	0.449	98.3	0.101
rs490634	0.13	99.6	0.816
rs13072536	0.254	99.8	0.082
rs4955526	NA	NA	NA
rs9844666	0.244	100	0.887
rs572169	0.293	100	0.143
rs17754	0.426	99.9	0.305
rs17472113	0.263	99	0.356
rs4864548	0.368	100	0.088
rs3796529	0.191	99.7	0.378
rs12503378	0.171	100	0.172
rs17541471	0.193	100	1
rs6180	0.457	99.9	0.689

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs832575	0.128	100	0.348
rs41132	0.236	99.9	0.908
rs2247870	0.456	100	0.146
rs17085675	0.28	100	0.392
rs17622208	0.462	100	0.366
rs9366637	0.067	100	0.315
rs2853977	0.402	100	0.048
rs2229642	0.48	95.5	0.019
rs1776897	NA	NA	NA
rs4946932	0.319	100	0.597
rs1476387	0.416	100	0.426
rs7756224	NA	NA	NA
rs2234693	0.46	100	0.933
rs2982712	0.422	100	0.48
rs1074287	NA	NA	NA
rs1636255	NA	NA	NA
rs864745	NA	NA	NA
rs3812265	0.241	100	0.841
rs1800783	0.386	99.9	0.643
rs6999671	NA	NA	NA
rs2145923	0.174	100	2.74E-05
rs3814115	0.325	100	0.886
rs7853859	0.369	99.9	0.637
rs3739707	0.26	100	0.957
rs7020782	0.303	100	0.729
rs803932	NA	NA	NA
rs11102986	0.181	99.3	0.75
rs291979	0.22	100	0.903
rs2735469	0.162	100	0.644
rs4320932	0.202	100	0.184
rs900147	0.298	100	9.30E-03
rs948847	0.413	99.9	0.504
rs174547	0.333	99.9	0.073
rs3736228	0.13	99.9	0.488
rs7396866	NA	NA	NA
rs674424	0.263	100	0.808
rs2282537	NA	NA	NA
rs6487088	0.197	100	0.197
rs7137534	0.323	100	0.261
rs2066807	0.067	99.8	0.314
rs2291617	0.316	98.5	0.942
rs1042725	0.487	100	0.818

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs3782415	0.214	100	0.153
rs6219	0.096	100	0.336
rs10861148	0.11	100	0.79
rs907482	0.331	100	0.741
rs1051431	0.22	100	0.738
rs1950500	NA	NA	NA
rs696	0.387	99.7	0.612
rs709939	0.472	100	0.785
rs3783937	0.24	100	1
rs1036477	0.11	100	0.748
rs12050767	0.444	100	0.597
rs7163907	0.253	100	0.782
rs17599989	NA	NA	NA
rs1516796	0.48	100	0.706
rs8033670	0.485	100	0.933
rs5015437	NA	NA	NA
rs258281	0.182	98.5	0.916
rs2023693	0.419	96.6	0.407
rs8055190	NA	NA	NA
rs7359336	0.42	100	0.035
rs8071847	0.207	100	0.022
rs2909430	0.138	100	0.174
rs11080149	0.119	100	0.842
rs2715553	0.455	99.8	0.752
rs752313	0.495	100	0.268
rs12603813	0.252	100	0.471
rs12603582	0.22	100	0.903
rs46522	NA	NA	NA
rs9892365	0.33	100	0.018
rs12940055	0.12	100	0.692
rs2854207	0.275	100	0.753
rs2053156	0.184	100	0.095
rs25656	NA	NA	NA
rs891088	0.262	100	0.978
rs4808199	NA	NA	NA
rs4803520	0.116	99.7	1
rs2682552	0.184	100	0.25
rs158676	0.323	100	0.038
rs2425019	0.474	100	0.295
rs1780616	0.355	100	0.891
rs11086538	0.485	99.8	9.41E-03
rs2057291	0.328	99.9	0.776

Part A - Panel 6: Framingham CARe

SNP	MAF	CR	HWE
rs451061	0.358	100	1
rs212517	0.386	100	1
rs1738475	0.412	100	1
rs2229712	NA	NA	NA
rs17106235	0.102	99.8	0.495
rs551219	0.276	99	0.915
rs12145922	0.448	99.1	0.798
rs660240	0.226	99.8	0.47
rs7522692	0.195	100	0.348
rs1342586	0.189	100	0.129
rs10185680	0.464	99.5	0.235
rs1866146	0.316	99.2	0.625
rs780094	0.427	100	0.731
rs7557989	0.32	100	0.699
rs1822469	0.407	100	0.257
rs6731022	0.302	100	0.549
rs3821009	0.068	100	0.738
rs6718902	0.23	100	0.475
rs6758561	0.356	100	0.714
rs526134	0.378	100	0.72
rs10208728	0.102	100	0.643
rs3100776	0.038	100	0.565
rs4973410	0.481	99.8	0.354
rs2679178	0.093	100	0.614
rs7578199	0.235	100	0.725
rs7572476	0.457	100	0.308
rs2450855	0.437	100	0.392
rs9857730	0.224	99.9	0.04
rs3915129	0.448	100	0.932
rs490634	0.13	98.6	0.344
rs13072536	0.22	99.9	0.269
rs4955526	0.361	100	0.171
rs9844666	0.244	99.9	0.306
rs572169	0.297	99.9	0.02
rs17754	0.45	100	0.552
rs17472113	0.265	98.7	1
rs4864548	0.342	100	0.049
rs3796529	0.188	99.3	0.403
rs12503378	0.177	100	0.385
rs17541471	0.217	100	0.457
rs6180	0.456	99.9	0.31

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs832575	0.14	99.9	0.727
rs41132	0.227	99.6	1
rs2247870	0.463	100	0.866
rs17085675	0.288	100	0.681
rs17622208	0.495	100	0.354
rs9366637	0.068	100	0.738
rs2853977	0.442	100	3.69E-03
rs2229642	NA	0	1
rs1776897	0.089	99.9	0.606
rs4946932	0.328	100	0.392
rs1476387	0.41	99.9	0.434
rs7756224	0.438	100	0.266
rs2234693	0.464	100	0.799
rs2982712	0.426	100	0.932
rs1074287	0.225	100	0.118
rs1636255	0.296	99.8	0.686
rs864745	0.478	100	0.129
rs3812265	0.229	100	0.031
rs1800783	0.398	100	0.482
rs6999671	0.03	100	1
rs2145923	0.194	100	0.179
rs3814115	0.295	100	0.42
rs7853859	0.388	99.9	0.214
rs3739707	0.262	100	0.128
rs7020782	0.296	100	0.92
rs803932	0.348	97.3	0.302
rs11102986	0.205	100	0.796
rs291979	0.246	100	1
rs2735469	0.167	100	1
rs4320932	0.209	100	0.799
rs900147	0.291	100	0.262
rs948847	0.408	100	0.663
rs174547	0.344	100	0.455
rs3736228	0.147	100	1
rs7396866	0.277	100	0.295
rs674424	0.252	100	0.502
rs2282537	0.139	100	1
rs6487088	0.211	100	0.011
rs7137534	0.308	100	0.114
rs2066807	0.066	99.9	1.69E-03
rs2291617	0.324	99.3	0.382
rs1042725	0.47	100	0.076

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs3782415	0.188	100	1
rs6219	0.085	100	0.278
rs10861148	0.106	100	0.662
rs907482	0.334	100	1
rs1051431	0.229	100	0.721
rs1950500	0.291	100	0.152
rs696	0.398	99.8	0.726
rs709939	0.428	100	0.731
rs3783937	0.234	100	0.814
rs1036477	0.118	100	0.415
rs12050767	0.424	99.9	0.301
rs7163907	0.241	99.1	0.489
rs17599989	0.272	99.9	0.033
rs1516796	0.464	99.9	0.672
rs8033670	0.488	98	0.932
rs5015437	0.338	99.6	0.22
rs258281	0.173	100	1
rs2023693	0.456	100	0.397
rs8055190	0.053	100	1
rs7359336	0.4	99.9	0.792
rs8071847	0.226	100	0.335
rs2909430	0.13	99.9	0.024
rs11080149	0.111	100	0.524
rs2715553	0.442	100	0.201
rs752313	0.477	100	0.866
rs12603813	0.239	95.1	0.057
rs12603582	0.252	99.8	0.911
rs46522	0.457	100	0.932
rs9892365	0.331	100	0.341
rs12940055	0.134	100	1
rs2854207	0.273	100	0.29
rs2053156	0.198	100	0.791
rs25656	NA	NA	NA
rs891088	0.289	100	0.307
rs4808199	0.163	100	0.537
rs4803520	0.114	100	1
rs2682552	0.207	100	0.442
rs158676	0.338	100	0.707
rs2425019	0.426	100	0.121
rs1780616	0.37	100	0.471
rs11086538	0.464	99.9	0.352
rs2057291	0.342	100	0.709

Part A - Panel 6: MESA CARe

SNP	MAF	CR	HWE
rs451061	0.373	100	0.436
rs212517	0.415	100	0.113
rs1738475	0.439	100	0.249
rs2229712	NA	NA	NA
rs17106235	0.104	99.9	0.512
rs551219	0.292	99.4	0.589
rs12145922	0.422	99.6	0.261
rs660240	0.204	99.8	0.416
rs7522692	0.19	100	0.742
rs1342586	0.223	99.9	0.907
rs10185680	0.456	99.7	0.252
rs1866146	0.334	99.8	0.964
rs780094	0.424	99.8	0.967
rs7557989	0.325	100	0.309
rs1822469	0.416	100	0.113
rs6731022	0.321	100	0.403
rs3821009	0.083	100	0.284
rs6718902	0.247	100	0.05
rs6758561	0.351	100	0.563
rs526134	0.425	100	0.384
rs10208728	0.098	100	1
rs3100776	0.031	99.7	1
rs4973410	0.493	100	0.57
rs2679178	0.1	100	0.501
rs7578199	0.259	100	0.399
rs7572476	0.472	100	0.329
rs2450855	0.428	100	0.482
rs9857730	0.2	99.8	0.31
rs3915129	0.44	99.6	0.773
rs490634	0.12	99.5	0.924
rs13072536	0.251	100	0.281
rs4955526	0.373	100	0.762
rs9844666	0.244	99.9	0.226
rs572169	0.283	100	3.21E-03
rs17754	0.433	100	0.71
rs17472113	0.238	95.2	0.059
rs4864548	0.371	100	0.224
rs3796529	0.192	97.1	0.098

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs12503378	0.177	100	0.835
rs17541471	0.209	100	0.22
rs6180	0.467	99.9	0.807
rs832575	0.126	99.9	0.408
rs41132	0.222	99.8	0.379
rs2247870	0.462	100	0.684
rs17085675	0.276	100	0.723
rs17622208	0.472	99.9	0.569
rs9366637	0.06	99.9	0.371
rs2853977	0.406	99.7	7.05E-03
rs2229642	NA	NA	NA
rs1776897	0.088	99.8	0.705
rs4946932	0.29	100	3.67E-03
rs1476387	0.425	100	0.804
rs7756224	0.418	100	0.505
rs2234693	0.47	100	0.44
rs2982712	0.402	100	0.833
rs1074287	0.251	100	0.161
rs1636255	0.274	99.4	0.168
rs864745	0.48	100	0.019
rs3812265	0.236	100	0.736
rs1800783	0.38	100	0.699
rs6999671	0.034	100	0.759
rs2145923	0.192	100	0.434
rs3814115	0.319	100	0.889
rs7853859	0.396	99.8	0.498
rs3739707	0.264	100	0.322
rs7020782	0.306	100	0.924
rs803932	0.322	97	0.066
rs11102986	0.19	100	0.742
rs291979	0.223	100	0.907
rs2735469	0.172	100	0.476
rs4320932	0.21	99.8	0.222
rs900147	0.315	100	0.639
rs948847	0.406	100	0.706
rs174547	0.33	100	0.855
rs3736228	0.133	99.9	0.66
rs7396866	0.271	100	0.918
rs674424	0.256	100	0.671
rs2282537	0.14	100	1
rs6487088	0.201	100	0.9
rs7137534	0.31	100	0.107

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs2066807	0.063	99.9	0.49
rs2291617	0.32	100	0.852
rs1042725	0.484	100	0.746
rs3782415	0.2	100	0.376
rs6219	0.11	100	0.146
rs10861148	0.12	100	0.502
rs907482	0.326	100	0.712
rs1051431	0.24	100	0.375
rs1950500	0.284	100	0.398
rs696	0.386	100	0.932
rs709939	0.472	100	0.113
rs3783937	0.265	100	0.212
rs1036477	0.11	100	0.918
rs12050767	0.438	100	0.967
rs7163907	0.262	96.5	0.67
rs17599989	0.277	100	0.117
rs1516796	0.461	100	0.935
rs8033670	0.488	100	0.24
rs5015437	0.366	99.9	0.965
rs258281	0.192	100	0.514
rs2023693	0.444	99.8	0.084
rs8055190	0.046	100	0.642
rs7359336	0.431	100	0.679
rs8071847	0.206	100	0.62
rs2909430	0.13	99.3	1
rs11080149	0.116	100	0.323
rs2715553	0.444	100	0.742
rs752313	0.5	100	0.598
rs12603813	0.241	99.8	0.618
rs12603582	0.241	99.9	0.868
rs46522	0.484	99.9	0.062
rs9892365	0.327	100	0.782
rs12940055	0.112	100	0.36
rs2854207	0.285	99.9	0.655
rs2053156	0.199	100	0.227
rs25656	NA	NA	NA
rs891088	0.254	100	0.24
rs4808199	0.166	100	0.77
rs4803520	0.116	100	0.692
rs2682552	0.182	100	0.786
rs158676	0.321	100	0.125
rs2425019	0.454	99.9	9.23E-04

Continued on next page

Table A4 – *Continued from previous page*

SNP	MAF	CR	HWE
rs1780616	0.346	100	0.395
rs11086538	0.473	99.2	0.142
rs2057291	0.341	100	0.528

Part B - COPDGene

SNP	Gene	Proxy SNP	R^2	Chr:Position	E/O	MA	MAF	CR	HWE
rs451061	PRKCZ	rs424079	1	1:2071340	C/A	C	0.389	99.9	0.142
rs212517	ECE1	rs212524	0.98	1:21583311	C/T	T	0.403	100	0.12
rs1738475	HTR1D	rs627304	1	1:23537555	T/C	C	0.413	100	0.315
rs2229712	RPS6KA1	NA	NA	1:26883511	A/C	NA	NA	NA	NA
rs17106235	FAF1	rs17106184	1	1:50909985	A/G	NA	NA	NA	NA
rs551219	COL24A1	rs618555	0.98	1:86481084	T/C	T	0.288	100	1
rs12145922	PKN2	rs1002436	0.96	1:89146852	G/A	NA	NA	NA	NA
rs660240	PSRC1	NA	NA	1:109817838	T/C	T	0.215	99.8	0.47
rs7522692	PIGC	rs1129942	1	1:172437592	G/A	NA	NA	NA	NA
rs1342586	TGFB2	rs10482796	0.99	1:218605635	C/T	NA	NA	NA	NA
rs10185680	MFSD2B	NA	NA	2:24275306	G/A	A	0.456	99.9	0.28
rs1866146	POMC	NA	NA	2:25380573	G/A	G	0.323	100	0.393
rs780094	GCKR	NA	NA	2:27741237	C/T	T	0.416	100	0.92
rs7557989	THADA	NA	NA	2:43630657	T/C	T	0.332	100	0.942
rs1822469	PPP3R1	rs687	0.99	2:68415767	G/A	NA	NA	NA	NA
rs6731022	EIF2AK3	rs11684404	1	2:88924622	C/T	NA	NA	NA	NA
rs3821009	PDE11A	rs1946812	1	2:178674935	A/G	NA	NA	NA	NA
rs6718902	STAT1	rs2066804	0.98	2:191841759	A/G	A	0.251	100	0.3
rs6758561	NOP58	NA	NA	2:203126559	A/G	A	0.345	99.9	0.35

Continued on next page

Table A4 – *Continued from previous page*

SNP	Gene	Proxy SNP	R^2	Chr:Position	E/O	MA	MAF	CR	HWE
rs526134	USP37	rs1516086	0.99	2:219431535	A/G	NA	NA	NA	NA
rs10208728	IHH	rs10445823	0.99	2:219910164	T/C	C	0.099	100	0.316
rs3100776	IHH	rs6436122	0.96	2:220036390	A/G	A	0.041	100	0.408
rs4973410	NCL	NA	NA	2:232331734	C/T	T	0.468	100	1
rs2679178	NPPC	rs2580821	1	2:232804155	C/A	A	0.086	99.9	0.606
rs7578199	HDLBP	NA	NA	2:242192848	T/C	C	0.247	100	0.206
rs7572476	BOK	NA	NA	2:242496325	C/T	T	0.46	100	0.493
rs2450855	MKRN2	NA	NA	3:12602494	G/A	NA	NA	NA	NA
rs9857730	VILL	NA	NA	3:38051941	C/T	C	0.209	99.9	0.588
rs3915129	CTNNB1	NA	NA	3:41243742	G/T	G	0.461	99.9	0.414
rs490634	CISH	rs201194	0.99	3:50642975	C/T	C	0.127	100	0.509
rs13072536	ITIH4	rs2276817	1	3:52860936	C/T	T	0.245	100	0.38
rs4955526	EPHB1	NA	NA	3:134317337	C/T	T	0.357	99.9	0.548
rs9844666	PCCB	NA	NA	3:135974216	G/A	A	0.241	100	0.79
rs572169	GHSR	NA	NA	3:172165727	T/C	T	0.314	99.9	0.105
rs17754	RFC1	rs1057807	0.99	4:39289473	A/G	G	0.436	99.9	0.121
rs17472113	ZAR1	rs9993088	0.99	4:48494546	A/C	A	0.275	99.9	0.903
rs4864548	CLOCK	NA	NA	4:56413803	A/G	A	0.366	100	0.162
rs3796529	REST	NA	NA	4:57797414	T/C	T	0.192	100	0.272
rs12503378	NUDT6	rs1048201	0.98	4:123814308	C/T	T	0.169	100	0.728
rs17541471	NPR3	rs10053636	1	5:32760375	C/T	NA	NA	NA	NA
rs6180	GHR	NA	NA	5:42719239	A/C	C	0.459	99.9	0.974
rs832575	MAP3K1	NA	NA	5:56161787	A/G	G	0.124	99.9	0.1
rs41132	AP3B1	rs252749	1	5:77389973	G/A	A	0.246	100	0.57

Continued on next page

Table A4 – *Continued from previous page*

SNP	Gene	Proxy SNP	R^2	Chr:Position	E/O	MA	MAF	CR	HWE
rs2247870	GPR98	NA	NA	5:90151589	A/G	NA	NA	NA	NA
rs17085675	PCSK1	NA	NA	5:95727664	T/A	NA	NA	NA	NA
rs17622208	SLC22A5	NA	NA	5:131717050	A/G	A	0.462	99.9	0.492
rs9366637	HFE	NA	NA	6:26089098	C/T	T	0.061	100	0.396
rs2853977	HCP5	NA	NA	6:31379304	A/T	NA	NA	NA	NA
rs2229642	ITPR3	rs2296742	1	6:33659793	A/G	A	0.468	100	0.948
rs1776897	HMGA1	NA	NA	6:34195011	G/T	G	0.089	99.9	0.088
rs4946932	FOXO3	rs2153960	0.92	6:108988184	A/G	G	0.306	99.9	0.939
rs1476387	PPIL6	NA	NA	6:109764535	G/T	T	0.415	100	0.593
rs7756224	NMBR	rs4577816	1	6:142423810	T/C	NA	NA	NA	NA
rs2234693	ESR1	NA	NA	6:152163335	C/T	NA	NA	NA	NA
rs2982712	ESR1	NA	NA	6:152358179	C/T	C	0.415	100	0.525
rs1074287	OPRM1	rs589046	0.91	6:154393138	T/C	T	0.256	99.9	0.831
rs1636255	GNA12	rs1636249	0.9	7:2884283	T/G	G	0.302	99.8	0.59
rs864745	JAZF1	rs849138	0.93	7:28177338	G/A	NA	NA	NA	NA
rs3812265	CNOT4	NA	NA	7:135048804	T/C	NA	NA	NA	NA
rs1800783	NOS3	rs10247107	0.94	7:150683083	G/A	A	0.37	99.9	0.834
rs6999671	RPS20	rs16920326	1	8:56995782	A/G	A	0.035	99.9	0.144
rs2145923	NPR2	rs7873145	1	9:35786616	T/C	NA	NA	NA	NA
rs3814115	PCSK5	NA	NA	9:78504729	C/T	C	0.336	99.2	0.29
rs7853859	CENPP	rs1053441	0.99	9:95147840	T/A	NA	NA	NA	NA
rs3739707	LPAR1	NA	NA	9:113792706	C/A	A	0.264	100	0.452
rs7020782	PAPPA	NA	NA	9:119106881	A/C	C	0.298	100	0.087
rs803932	ASTN2	NA	NA	9:119458020	C/T	C	0.325	99.9	0.63

Continued on next page

Table A4 – *Continued from previous page*

SNP	Gene	Proxy SNP	R^2	Chr:Position	E/O	MA	MAF	CR	HWE
rs11102986	RXRA	NA	NA	9:137285503	G/A	A	0.185	99.8	0.42
rs291979	GRK5	rs291970	0.99	10:121123633	T/C	T	0.23	100	0.582
rs2735469	MRPL23	rs2735971	0.91	11:2021649	T/C	NA	NA	NA	NA
rs4320932	IGF2	NA	NA	11:2171601	T/C	C	0.206	100	0.457
rs900147	ARNTL	rs900145	1	11:13293905	C/T	C	0.298	99.6	0.726
rs948847	APLNR	NA	NA	11:57004344	G/T	G	0.432	99.8	0.766
rs174547	FADS1	NA	NA	11:61570783	T/C	C	0.337	99.8	0.561
rs3736228	LRP5	NA	NA	11:68201295	C/T	T	0.152	100	0.9
rs7396866	NEU3	rs10793108	0.91	11:74695212	G/A	NA	NA	NA	NA
rs674424	ABCG4	NA	NA	11:119030752	T/C	T	0.256	100	0.798
rs2282537	POU2F3	NA	NA	11:120187971	G/A	A	0.142	100	0.039
rs6487088	PDE3A	NA	NA	12:20588382	T/C	C	0.207	99.9	0.488
rs7137534	PDE3A	NA	NA	12:20831777	T/C	T	0.308	100	0.703
rs2066807	PAN2	NA	NA	12:56740682	G/C	G	0.063	100	0.053
rs2291617	METTL1	NA	NA	12:58166403	T/G	G	0.325	99.7	0.683
rs1042725	HMGA2	NA	NA	12:66358347	C/T	T	0.494	100	0.745
rs3782415	SOCS2	NA	NA	12:93967755	C/T	C	0.189	100	0.792
rs6219	IGF1	NA	NA	12:102790192	T/C	T	0.105	99.9	0.729
rs10861148	HSP90B1	rs4135054	0.99	12:104363610	T/C	T	0.111	99.9	0.934
rs907482	KNTC1	rs7970027	0.99	12:122780375	C/T	C	0.339	99.9	0.192
rs1051431	MPHOSPH9	NA	NA	12:123645803	G/A	G	0.226	100	0.151
rs1950500	NFATC4	NA	NA	14:24830850	T/C	T	0.275	99.9	0.308
rs696	NFKBIA	rs8904	0.99	14:35871217	G/A	A	0.369	100	0.117
rs709939	SAMD4A	rs2281652	0.98	14:55253864	A/G	G	0.449	100	0.189

Continued on next page

Table A4 – *Continued from previous page*

SNP	Gene	Proxy SNP	R^2	Chr:Position	E/O	MA	MAF	CR	HWE
rs3783937	FBLN5	NA	NA	14:92407693	C/T	T	0.261	100	1
rs1036477	FBN1	NA	NA	15:48914926	A/G	G	0.109	99.2	0.933
rs12050767	CYP19A1	rs749292	1	15:51558731	A/G	A	0.434	100	0.792
rs7163907	PTPN9	rs11636031	1	15:75815758	C/T	C	0.264	100	0.05
rs17599989	SEC11A	rs11637142	0.99	15:85295927	G/A	G	0.289	99.8	0.692
rs1516796	ACAN	rs8041863	1	15:89359689	A/T	NA	NA	NA	NA
rs8033670	IGF1R	rs9672965	0.96	15:99498433	T/C	T	0.478	100	0.091
rs5015437	LMF1	rs5015441	1	16:987256	T/C	T	0.375	99.9	0.488
rs258281	RAB26	NA	NA	16:2191734	G/A	NA	NA	NA	NA
rs2023693	ERI2	rs11074476	1	16:20886385	C/A	A	0.43	99.7	0.26
rs8055190	LRRC36	rs16957358	0.95	16:67394541	A/G	G	0.044	99.9	0.704
rs7359336	NFAT5	NA	NA	16:69733460	G/A	G	0.417	100	0.316
rs8071847	POLR2A	NA	NA	17:7407327	G/A	G	0.206	99.9	0.692
rs2909430	TP53	rs1625895	0.99	17:7578115	C/T	T	0.136	99.9	0.032
rs11080149	NF1	NA	NA	17:29623288	T/C	T	0.119	100	1
rs2715553	RARA	NA	NA	17:38496320	A/G	G	0.453	100	0.694
rs752313	EZH1	rs11868496	1	17:40881162	T/C	NA	NA	NA	NA
rs12603813	PLCD3	NA	NA	17:43196584	T/C	C	0.248	100	0.965
rs12603582	ITGB3	NA	NA	17:45377577	G/T	T	0.228	99.9	1
rs46522	UBE2Z	rs318095	1	17:46974734	T/C	T	0.472	100	0.948
rs9892365	TBX2	rs2079795	0.98	17:59496649	T/C	T	0.335	99.9	0.466
rs12940055	MAP3K3	NA	NA	17:61722142	C/T	T	0.114	100	0.468
rs2854207	CSH2	rs2854201	0.96	17:61947754	A/G	NA	NA	NA	NA
rs2053156	GRB2	rs959260	1	17:73369422	T/C	C	0.178	99.6	0.617

Continued on next page

Table A4 – *Continued from previous page*

SNP	Gene	Proxy SNP	R^2	Chr:Position	E/O	MA	MAF	CR	HWE
rs25656	NFATC1	NA	NA	18:77227476	A/G	NA	NA	NA	NA
rs891088	INSR	NA	NA	19:7184762	G/A	G	0.262	99.9	0.9
rs4808199	GATAD2A	NA	NA	19:19545099	G/A	A	0.18	99.9	0.956
rs4803520	GRIK5	NA	NA	19:42500373	G/A	NA	NA	NA	NA
rs2682552	XRCC1	rs2682587	0.97	19:44082429	A/C	A	0.184	100	0.33
rs158676	CDK5RAP1	NA	NA	20:31974395	A/G	G	0.322	100	0.158
rs2425019	MMP24	NA	NA	20:33819415	G/A	NA	NA	NA	NA
rs1780616	LBP	rs7273717	0.9	20:36971709	T/C	C	0.37	100	0.223
rs11086538	MC3R	NA	NA	20:54817822	G/T	T	0.471	99.8	0.648
rs2057291	GNAS	NA	NA	20:57472043	A/G	A	0.338	99.9	0.31

Table A5: Conditional quantile regression (CQR) models of BMI/obesity-associated SNPs and GS-BMI across the sample distribution. CQR models were fitted every 5th percentile of BMI and adjusted for age, age-squared, sex and study. β from ordinary least squares (OLS) and CQR models at each percentile are the effect sizes (kg/m^2 per Effect Allele). 95%CI are the 95% confidence intervals. In addition, the proportion of BMI variance that is explained by the GS-BMI was estimated, Variance Explained (%).

Part 1 - CQR Models 5-20%

SNP	Gene	N		5%	10%	15%	20%
rs1421085	FTO	75229	β	0.224	0.281	0.303	0.335
			$[95\%CI]$	$[0.157, 0.295]$	$[0.233, 0.344]$	$[0.249, 0.361]$	$[0.279, 0.389]$
			p-value	1.46E-10	2.41E-23	1.83E-26	8.67E-33
rs10767664	BDNF	74703	β	0.16	0.147	0.149	0.169
			$[95\%CI]$	$[0.077, 0.250]$	$[0.080, 0.211]$	$[0.080, 0.211]$	$[0.105, 0.236]$
			p-value	0.00028	0.0000117	0.0000114	0.0000004
rs11672660	GIPR	72569	β	0.124	0.126	0.098	0.087
			$[95\%CI]$	$[0.037, 0.202]$	$[0.053, 0.200]$	$[0.041, 0.171]$	$[0.020, 0.162]$
			p-value	0.00349	0.000826	0.00329	0.015
rs4788099	SH2B1	63924	β	0.081	0.077	0.082	0.097
			$[95\%CI]$	$[0.001, 0.147]$	$[0.020, 0.134]$	$[0.029, 0.144]$	$[0.043, 0.163]$
			p-value	0.027	0.00724	0.0051	0.0015
rs7903146	TCF7L2	75230	β	0.011	0.028	0.039	0.066
			$[95\%CI]$	$[-0.060, 0.079]$	$[-0.032, 0.085]$	$[-0.015, 0.100]$	$[0.010, 0.126]$
			p-value	0.76	0.35	0.184	0.025
rs2075650	TOMM40	74766	β	0.194	0.222	0.219	0.222
			$[95\%CI]$	$[0.092, 0.285]$	$[0.132, 0.295]$	$[0.141, 0.293]$	$[0.136, 0.303]$
			p-value	0.0000873	0.000000116	0.000000023	0.000000168
rs11873305	MC4R	75229	β	0.097	0.149	0.177	0.17
			$[95\%CI]$	$[-0.053, 0.276]$	$[0.013, 0.317]$	$[0.059, 0.305]$	$[0.043, 0.316]$
			p-value	0.245	0.056	0.00429	0.013
rs997295	MAP2K5	75214	β	0.03	0.033	0.05	0.064
			$[95\%CI]$	$[-0.035, 0.092]$	$[-0.022, 0.090]$	$[-0.003, 0.105]$	$[0.006, 0.116]$

Continued on next page

Table A5 – *Continued from previous page*

SNP	Gene	N		5%	10%	15%	20%
			p-value	0.364	0.246	0.067	0.022
rs3824755	NT5C2	75227	β	0.07	0.054	0.078	0.108
			$[95\%CI]$	$[-0.024, 0.184]$	$[-0.025, 0.166]$	$[-0.006, 0.169]$	$[0.015, 0.202]$
			p-value	0.183	0.265	0.083	0.024
rs12617233	FANCL	75230	β	0.017	0.024	0.019	0.067
			$[95\%CI]$	$[-0.051, 0.078]$	$[-0.032, 0.088]$	$[-0.033, 0.074]$	$[0.015, 0.123]$
			p-value	0.608	0.433	0.486	0.015
rs6499653	FTO	74894	β	0.012	0.027	0.037	0.057
			$[95\%CI]$	$[-0.063, 0.094]$	$[-0.032, 0.096]$	$[-0.025, 0.102]$	$[-0.010, 0.117]$
			p-value	0.771	0.407	0.247	0.078
rs1788826	NPC1	75220	β	0.066	0.079	0.088	0.088
			$[95\%CI]$	$[-0.009, 0.134]$	$[0.025, 0.142]$	$[0.033, 0.144]$	$[0.031, 0.145]$
			p-value	0.071	0.00828	0.00198	0.00236
rs17066846	MC4R	75120	β	-0.009	0.057	0.055	0.071
			$[95\%CI]$	$[-0.088, 0.073]$	$[-0.010, 0.130]$	$[-0.005, 0.131]$	$[0.005, 0.138]$
			p-value	0.832	0.112	0.113	0.037
rs6453133	HMGCR	75128	β	0.028	0.026	0.081	0.106
			$[95\%CI]$	$[-0.049, 0.096]$	$[-0.034, 0.093]$	$[0.024, 0.138]$	$[0.046, 0.163]$
			p-value	0.45	0.43	0.00504	0.000326
rs739564	IQCK	73065	β	0.07	0.1	0.09	0.067
			$[95\%CI]$	$[-0.010, 0.157]$	$[0.025, 0.170]$	$[0.017, 0.155]$	$[0.004, 0.142]$
			p-value	0.101	0.00713	0.011	0.058
rs2272903	TFAP2B	75228	β	0.144	0.071	0.07	0.092
			$[95\%CI]$	$[0.047, 0.244]$	$[-0.012, 0.177]$	$[-0.008, 0.158]$	$[0.011, 0.177]$
			p-value	0.00479	0.141	0.1	0.031
rs7553158	TNNI3K	75230	β	0.012	0.031	0.024	0.029
			$[95\%CI]$	$[-0.057, 0.073]$	$[-0.021, 0.092]$	$[-0.026, 0.080]$	$[-0.028, 0.082]$
			p-value	0.727	0.283	0.369	0.303
rs11570094	SPI1	75200	β	0.075	0.063	0.042	0.068

Continued on next page

Table A5 – *Continued from previous page*

SNP	Gene	N		5%	10%	15%	20%
			$[95\%CI]$	$[0.000, 0.142]$	$[0.010, 0.126]$	$[-0.012, 0.106]$	$[0.007, 0.128]$
			p-value	0.038	0.034	0.168	0.027
rs4946932	FOXO3	71439	β	0.051	0.042	0.068	0.08
			$[95\%CI]$	$[-0.022, 0.122]$	$[-0.011, 0.110]$	$[0.009, 0.129]$	$[0.022, 0.146]$
			p-value	0.16	0.179	0.026	0.011
rs2819347	LMOD1	75223	β	0.009	0.055	0.065	0.086
			$[95\%CI]$	$[-0.059, 0.090]$	$[0.001, 0.118]$	$[0.009, 0.125]$	$[0.029, 0.146]$
			p-value	0.81	0.065	0.028	0.00374
rs2836754	ETS2	66054	β	0.063	0.042	0.059	0.057
			$[95\%CI]$	$[-0.010, 0.126]$	$[-0.012, 0.104]$	$[0.001, 0.116]$	$[0.005, 0.122]$
			p-value	0.067	0.154	0.045	0.053
rs2984618	TAL1	75173	β	-0.001	0.042	0.039	0.055
			$[95\%CI]$	$[-0.076, 0.067]$	$[-0.009, 0.106]$	$[-0.010, 0.096]$	$[0.002, 0.115]$
			p-value	0.982	0.163	0.15	0.06
rs11208662	LEPR	75177	β	0.03	0.103	0.08	0.122
			$[95\%CI]$	$[-0.087, 0.155]$	$[0.005, 0.205]$	$[0.001, 0.190]$	$[0.025, 0.215]$
			p-value	0.619	0.039	0.099	0.012
rs6235	PCSK1	75183	β	-0.056	-0.045	-0.042	-0.024
			$[95\%CI]$	$[-0.128, 0.017]$	$[-0.118, 0.009]$	$[-0.101, 0.017]$	$[-0.085, 0.033]$
			p-value	0.132	0.167	0.157	0.42
rs9356744	CDKAL1	70863	β	-0.002	0.003	-0.001	0.005
			$[95\%CI]$	$[-0.078, 0.066]$	$[-0.059, 0.064]$	$[-0.059, 0.055]$	$[-0.055, 0.062]$
			p-value	0.959	0.921	0.985	0.854
rs7988412	MTIF3	61821	β	0.048	0.026	0.064	0.053
			$[95\%CI]$	$[-0.060, 0.135]$	$[-0.051, 0.097]$	$[-0.019, 0.142]$	$[-0.017, 0.135]$
			p-value	0.34	0.491	0.114	0.17
rs1780050	NEXN	75224	β	0.054	0.047	0.051	0.057
			$[95\%CI]$	$[-0.021, 0.108]$	$[-0.007, 0.109]$	$[-0.004, 0.102]$	$[0.002, 0.111]$
			p-value	0.102	0.107	0.054	0.034

Continued on next page

Table A5 – *Continued from previous page*

SNP	Gene	N		5%	10%	15%	20%
rs526134	USP37	63859	β	0.013	0.037	0.013	0.042
			$[95\%CI]$	$[-0.056, 0.092]$	$[-0.025, 0.096]$	$[-0.042, 0.076]$	$[-0.008, 0.103]$
			p-value	0.734	0.231	0.676	0.143
rs980828	NOS1AP	75222	β	-0.014	-0.018	-0.004	-0.01
			$[95\%CI]$	$[-0.076, 0.051]$	$[-0.076, 0.034]$	$[-0.058, 0.047]$	$[-0.059, 0.047]$
			p-value	0.678	0.518	0.896	0.72
rs17001561	SCARB2	66079	β	0.103	0.044	0.095	0.05
			$[95\%CI]$	$[0.010, 0.186]$	$[-0.034, 0.125]$	$[0.014, 0.166]$	$[-0.026, 0.127]$
			p-value	0.021	0.286	0.015	0.191
rs6232	PCSK1	75225	β	-0.059	0.001	0.042	0.056
			$[95\%CI]$	$[-0.193, 0.093]$	$[-0.112, 0.144]$	$[-0.071, 0.171]$	$[-0.058, 0.187]$
			p-value	0.425	0.988	0.494	0.372
rs749767	KAT8	70976	β	-0.027	0.027	0.041	0.043
			$[95\%CI]$	$[-0.105, 0.039]$	$[-0.030, 0.089]$	$[-0.017, 0.096]$	$[-0.010, 0.102]$
			p-value	0.453	0.369	0.152	0.133
rs1211166	NTRK2	75213	β	0.065	0.041	0.025	0.012
			$[95\%CI]$	$[-0.016, 0.145]$	$[-0.019, 0.115]$	$[-0.041, 0.090]$	$[-0.055, 0.083]$
			p-value	0.112	0.229	0.446	0.725
rs2535633	ITIH4	75189	β	-0.009	-0.001	-0.004	-0.008
			$[95\%CI]$	$[-0.072, 0.060]$	$[-0.064, 0.053]$	$[-0.056, 0.048]$	$[-0.063, 0.046]$
			p-value	0.796	0.97	0.879	0.78
rs10144353	PRKCH	65613	β	0.048	0.037	0.011	0
			$[95\%CI]$	$[-0.072, 0.185]$	$[-0.056, 0.124]$	$[-0.089, 0.105]$	$[-0.112, 0.091]$
			p-value	0.475	0.399	0.829	0.995
rs1561288	ADCY3	75226	β	0.042	0.016	0.048	0.03
			$[95\%CI]$	$[-0.041, 0.112]$	$[-0.049, 0.085]$	$[-0.009, 0.117]$	$[-0.037, 0.099]$
			p-value	0.284	0.633	0.143	0.383
rs2283228	KCNQ1	72933	β	-0.105	-0.061	-0.023	-0.001
			$[95\%CI]$	$[-0.221, 0.027]$	$[-0.176, 0.055]$	$[-0.138, 0.098]$	$[-0.108, 0.126]$

Continued on next page

Table A5 – *Continued from previous page*

SNP	Gene	N		5%	10%	15%	20%
			p-value	0.097	0.304	0.694	0.989
GS-BMI		75224	β	0.042	0.054	0.065	0.067
			$[95\%CI]$	$[0.029, 0.055]$	$[0.042, 0.065]$	$[0.054, 0.075]$	$[0.057, 0.077]$
			p-value	3.64E-10	5.58E-21	1.97E-31	1.01E-37
			VarianceExplained	0.00087	0.0013	0.0016	0.00171

Part 2 - CQR Models 25-40%

SNP	Gene	N		25%	30%	35%	40%
rs1421085	FTO	75229	β	0.354	0.372	0.406	0.41
			$[95\%CI]$	$[0.296, 0.410]$	$[0.318, 0.435]$	$[0.350, 0.465]$	$[0.348, 0.471]$
			p-value	1.82E-33	1.13E-35	1.18E-43	2.15E-38
rs10767664	BDNF	74703	β	0.205	0.211	0.216	0.223
			$[95\%CI]$	$[0.138, 0.265]$	$[0.147, 0.290]$	$[0.145, 0.288]$	$[0.148, 0.300]$
			p-value	2.73E-10	4.81E-09	2.72E-09	7.49E-09
rs11672660	GIPR	72569	β	0.149	0.153	0.15	0.164
			$[95\%CI]$	$[0.079, 0.213]$	$[0.086, 0.221]$	$[0.085, 0.229]$	$[0.092, 0.238]$
			p-value	1.26E-05	7.30E-06	4.63E-05	1.03E-05
rs4788099	SH2B1	63924	β	0.118	0.098	0.107	0.127
			$[95\%CI]$	$[0.051, 0.176]$	$[0.032, 0.161]$	$[0.041, 0.174]$	$[0.057, 0.197]$
			p-value	1.95E-04	2.83E-03	1.49E-03	3.82E-04
rs7903146	TCF7L2	75230	β	0.077	0.082	0.096	0.13
			$[95\%CI]$	$[0.012, 0.135]$	$[0.020, 0.144]$	$[0.042, 0.161]$	$[0.065, 0.197]$
			p-value	1.40E-02	8.03E-03	1.53E-03	1.15E-04
rs2075650	TOMM40	74766	β	0.223	0.206	0.227	0.266
			$[95\%CI]$	$[0.135, 0.296]$	$[0.127, 0.294]$	$[0.153, 0.312]$	$[0.183, 0.356]$
			p-value	3.52E-08	1.15E-06	1.82E-08	1.75E-09
rs11873305	MC4R	75229	β	0.245	0.22	0.253	0.286
			$[95\%CI]$	$[0.070, 0.379]$	$[0.091, 0.365]$	$[0.111, 0.409]$	$[0.121, 0.442]$
			p-value	1.68E-03	1.54E-03	7.61E-04	5.64E-04
rs997295	MAP2K5	75214	β	0.074	0.064	0.086	0.105
			$[95\%CI]$	$[0.014, 0.131]$	$[0.007, 0.125]$	$[0.028, 0.145]$	$[0.047, 0.170]$
			p-value	0.012	0.031	0.00375	8.93E-04
rs3824755	NT5C2	75227	β	0.13	0.16	0.182	0.18
			$[95\%CI]$	$[0.032, 0.231]$	$[0.073, 0.252]$	$[0.079, 0.284]$	$[0.084, 0.287]$
			p-value	1.00E-02	0.000458	4.56E-04	4.50E-04
rs12617233	FANCL	75230	β	0.08	0.08	0.109	0.123

Continued on next page

Table A5 – *Continued from previous page*

SNP	Gene	N		25%	30%	35%	40%
			$[95\%CI]$	$[0.021, 0.141]$	$[0.026, 0.139]$	$[0.052, 0.170]$	$[0.056, 0.184]$
			p-value	0.00845	0.00521	0.000269	0.000169
rs6499653	FTO	74894	β	0.059	0.068	0.051	0.078
			$[95\%CI]$	$[-0.007, 0.123]$	$[0.004, 0.128]$	$[-0.016, 0.118]$	$[-0.001, 0.148]$
			p-value	7.50E-02	0.032	0.129	0.039
rs1788826	NPC1	75220	β	0.08	0.079	0.075	0.074
			$[95\%CI]$	$[0.016, 0.136]$	$[0.020, 0.133]$	$[0.014, 0.129]$	$[0.011, 0.141]$
			p-value	0.00941	0.00609	0.011	2.30E-02
rs17066846	MC4R	75120	β	0.085	0.086	0.092	0.155
			$[95\%CI]$	$[0.001, 0.158]$	$[0.003, 0.149]$	$[0.020, 0.167]$	$[0.075, 0.235]$
			p-value	3.20E-02	0.021	0.014	0.000132
rs6453133	HMGCR	75128	β	0.104	0.088	0.114	0.127
			$[95\%CI]$	$[0.037, 0.156]$	$[0.028, 0.147]$	$[0.054, 0.175]$	$[0.063, 0.196]$
			p-value	0.000717	0.00419	0.000249	0.000218
rs739564	IQCK	73065	β	0.088	0.082	0.113	0.109
			$[95\%CI]$	$[0.011, 0.157]$	$[0.011, 0.156]$	$[0.034, 0.187]$	$[0.029, 0.199]$
			p-value	0.017	0.026	0.00375	0.012
rs2272903	TFAP2B	75228	β	0.095	0.098	0.102	0.109
			$[95\%CI]$	$[-0.003, 0.187]$	$[0.006, 0.190]$	$[0.012, 0.194]$	$[0.011, 0.207]$
			p-value	4.80E-02	3.60E-02	3.00E-02	3.00E-02
rs7553158	TNNI3K	75230	β	0.038	0.043	0.061	0.068
			$[95\%CI]$	$[-0.025, 0.090]$	$[-0.017, 0.095]$	$[0.007, 0.123]$	$[0.012, 0.136]$
			p-value	0.202	0.133	0.041	0.031
rs11570094	SPI1	75200	β	0.091	0.101	0.082	0.08
			$[95\%CI]$	$[0.024, 0.157]$	$[0.038, 0.155]$	$[0.018, 0.143]$	$[0.014, 0.152]$
			p-value	0.00816	7.14E-04	9.94E-03	0.022
rs4946932	FOXO3	71439	β	0.079	0.089	0.088	0.093
			$[95\%CI]$	$[0.018, 0.148]$	$[0.030, 0.148]$	$[0.027, 0.153]$	$[0.023, 0.157]$
			p-value	0.017	0.0031	0.00588	6.10E-03

Continued on next page

Table A5 – *Continued from previous page*

SNP	Gene	N		25%	30%	35%	40%
rs2819347	LMOD1	75223	β	0.077	0.069	0.09	0.124
			$[95\%CI]$	$[0.011, 0.134]$	$[0.006, 0.129]$	$[0.028, 0.154]$	$[0.054, 0.185]$
			p-value	0.014	0.026	0.00452	0.000216
rs2836754	ETS2	66054	β	0.076	0.095	0.088	0.088
			$[95\%CI]$	$[0.017, 0.136]$	$[0.029, 0.155]$	$[0.026, 0.155]$	$[0.025, 0.149]$
			p-value	0.012	0.0028	0.00717	0.00522
rs2984618	TAL1	75173	β	0.066	0.043	0.027	0.052
			$[95\%CI]$	$[0.006, 0.124]$	$[-0.014, 0.095]$	$[-0.029, 0.092]$	$[-0.008, 0.112]$
			p-value	0.028	0.121	0.385	0.091
rs11208662	LEPR	75177	β	0.149	0.149	0.167	0.15
			$[95\%CI]$	$[0.052, 0.248]$	$[0.055, 0.248]$	$[0.071, 0.266]$	$[0.054, 0.255]$
			p-value	0.00312	0.00234	0.000761	0.00334
rs6235	PCSK1	75183	β	0.003	0.015	0.022	0.038
			$[95\%CI]$	$[-0.061, 0.072]$	$[-0.046, 0.084]$	$[-0.042, 0.090]$	$[-0.032, 0.106]$
			p-value	0.924	0.661	0.522	0.282
rs9356744	CDKAL1	70863	β	0.016	0.013	0.021	0.039
			$[95\%CI]$	$[-0.050, 0.076]$	$[-0.041, 0.077]$	$[-0.040, 0.082]$	$[-0.027, 0.110]$
			p-value	0.607	0.666	0.505	0.268
rs7988412	MTIF3	61821	β	0.071	0.073	0.091	0.093
			$[95\%CI]$	$[-0.013, 0.142]$	$[-0.001, 0.151]$	$[0.009, 0.170]$	$[0.010, 0.193]$
			p-value	0.074	0.058	0.026	0.047
rs1780050	NEXN	75224	β	0.037	0.018	0.014	0.03
			$[95\%CI]$	$[-0.022, 0.098]$	$[-0.037, 0.071]$	$[-0.038, 0.076]$	$[-0.034, 0.090]$
			p-value	0.224	0.524	0.637	0.349
rs526134	USP37	63859	β	0.059	0.049	0.058	0.054
			$[95\%CI]$	$[-0.005, 0.119]$	$[-0.011, 0.111]$	$[0.002, 0.123]$	$[-0.007, 0.121]$
			p-value	0.067	0.122	0.062	0.105
rs980828	NOS1AP	75222	β	-0.011	0.004	0.011	0.022
			$[95\%CI]$	$[-0.071, 0.044]$	$[-0.049, 0.057]$	$[-0.041, 0.072]$	$[-0.039, 0.083]$

Continued on next page

Table A5 – *Continued from previous page*

SNP	Gene	N		25%	30%	35%	40%
			p-value	0.704	0.881	0.693	0.486
rs17001561	SCARB2	66079	β	0.075	0.076	0.108	0.081
			$[95\%CI]$	$[-0.010, 0.150]$	$[-0.010, 0.150]$	$[0.027, 0.197]$	$[0.000, 0.169]$
			p-value	0.065	0.062	0.014	0.058
rs6232	PCSK1	75225	β	0.101	0.14	0.121	0.127
			$[95\%CI]$	$[-0.033, 0.233]$	$[0.001, 0.252]$	$[0.003, 0.263]$	$[-0.011, 0.278]$
			p-value	0.137	0.032	0.064	0.085
rs749767	KAT8	70976	β	0.064	0.068	0.083	0.065
			$[95\%CI]$	$[0.003, 0.122]$	$[0.005, 0.124]$	$[0.024, 0.146]$	$[0.004, 0.134]$
			p-value	0.034	0.028	0.00795	0.05
rs1211166	NTRK2	75213	β	0.015	0.025	0.031	0.036
			$[95\%CI]$	$[-0.057, 0.092]$	$[-0.042, 0.091]$	$[-0.041, 0.104]$	$[-0.046, 0.111]$
			p-value	0.698	0.451	0.402	0.374
rs2535633	ITIH4	75189	β	0.006	0.004	0.019	0.031
			$[95\%CI]$	$[-0.052, 0.062]$	$[-0.053, 0.059]$	$[-0.038, 0.079]$	$[-0.032, 0.090]$
			p-value	0.83	0.878	0.535	0.324
rs10144353	PRKCH	65613	β	0.004	0.005	0.008	0.004
			$[95\%CI]$	$[-0.102, 0.113]$	$[-0.090, 0.115]$	$[-0.098, 0.098]$	$[-0.106, 0.116]$
			p-value	0.945	0.927	0.867	0.942
rs1561288	ADCY3	75226	β	0.014	0.012	0.017	0.048
			$[95\%CI]$	$[-0.054, 0.085]$	$[-0.052, 0.075]$	$[-0.044, 0.091]$	$[-0.024, 0.120]$
			p-value	0.7	0.703	0.616	0.19
rs2283228	KCNQ1	72933	β	0.012	-0.011	-0.008	-0.058
			$[95\%CI]$	$[-0.110, 0.120]$	$[-0.118, 0.110]$	$[-0.132, 0.086]$	$[-0.176, 0.059]$
			p-value	0.844	0.847	0.886	0.337
GS-BMI		75224	β	0.072	0.078	0.088	0.099
			$[95\%CI]$	$[0.063, 0.082]$	$[0.068, 0.089]$	$[0.077, 0.099]$	$[0.088, 0.109]$
			p-value	1.47E-47	3.56E-47	2.58E-52	4.41E-74
			VarianceExplained	0.19%	0.20%	0.23%	0.27%

Part 3 - CQR Models 45-60%

SNP	Gene	N		45%	50%	55%	60%
rs1421085	FTO	75229	β	0.423	0.429	0.454	0.481
			$[95\%CI]$	$[0.354, 0.484]$	$[0.364, 0.499]$	$[0.382, 0.521]$	$[0.402, 0.568]$
			p-value	7.05E-37	6.25E-36	4.40E-38	3.91E-30
rs10767664	BDNF	74703	β	0.252	0.231	0.266	0.278
			$[95\%CI]$	$[0.176, 0.330]$	$[0.156, 0.315]$	$[0.183, 0.338]$	$[0.190, 0.359]$
			p-value	9.86E-11	1.27E-08	2.05E-11	1.66E-10
rs11672660	GIPR	72569	β	0.196	0.203	0.256	0.253
			$[95\%CI]$	$[0.122, 0.269]$	$[0.126, 0.277]$	$[0.165, 0.326]$	$[0.171, 0.344]$
			p-value	1.85E-07	1.43E-07	3.41E-10	1.24E-08
rs4788099	SH2B1	63924	β	0.157	0.131	0.141	0.157
			$[95\%CI]$	$[0.082, 0.220]$	$[0.065, 0.198]$	$[0.069, 0.226]$	$[0.082, 0.243]$
			p-value	7.28E-06	1.08E-04	4.06E-04	1.06E-04
rs7903146	TCF7L2	75230	β	0.153	0.126	0.135	0.148
			$[95\%CI]$	$[0.080, 0.216]$	$[0.058, 0.198]$	$[0.065, 0.213]$	$[0.070, 0.227]$
			p-value	1.09E-05	3.86E-04	3.16E-04	2.61E-04
rs2075650	TOMM40	74766	β	0.297	0.244	0.261	0.272
			$[95\%CI]$	$[0.204, 0.382]$	$[0.151, 0.343]$	$[0.157, 0.351]$	$[0.164, 0.376]$
			p-value	5.71E-11	0.000000576	1.30E-07	4.86E-07
rs11873305	MC4R	75229	β	0.292	0.267	0.368	0.411
			$[95\%CI]$	$[0.114, 0.457]$	$[0.126, 0.435]$	$[0.193, 0.514]$	$[0.229, 0.589]$
			p-value	8.33E-04	6.79E-04	7.56E-06	8.80E-06
rs997295	MAP2K5	75214	β	0.136	0.133	0.15	0.165
			$[95\%CI]$	$[0.072, 0.198]$	$[0.071, 0.200]$	$[0.081, 0.219]$	$[0.087, 0.231]$
			p-value	1.73E-05	4.88E-05	2.14E-05	7.59E-06
rs3824755	NT5C2	75227	β	0.182	0.196	0.218	0.219
			$[95\%CI]$	$[0.081, 0.297]$	$[0.096, 0.298]$	$[0.105, 0.335]$	$[0.104, 0.357]$
			p-value	9.39E-04	1.43E-04	1.81E-04	6.49E-04
rs12617233	FANCL	75230	β	0.128	0.122	0.108	0.106

Continued on next page

Table A5 – *Continued from previous page*

SNP	Gene	N		45%	50%	55%	60%
			$[95\%CI]$	$[0.063, 0.189]$	$[0.050, 0.182]$	$[0.036, 0.175]$	$[0.040, 0.186]$
			p-value	0.0000632	0.000258	0.00242	0.00448
rs6499653	FTO	74894	β	0.099	0.085	0.116	0.142
			$[95\%CI]$	$[0.026, 0.166]$	$[0.018, 0.160]$	$[0.042, 0.196]$	$[0.062, 0.229]$
			p-value	5.02E-03	2.00E-02	3.18E-03	1.17E-03
rs1788826	NPC1	75220	β	0.089	0.081	0.089	0.113
			$[95\%CI]$	$[0.024, 0.152]$	$[0.011, 0.144]$	$[0.017, 0.156]$	$[0.036, 0.185]$
			p-value	6.37E-03	1.70E-02	1.30E-02	2.83E-03
rs17066846	MC4R	75120	β	0.189	0.169	0.184	0.178
			$[95\%CI]$	$[0.111, 0.267]$	$[0.089, 0.253]$	$[0.098, 0.271]$	$[0.094, 0.264]$
			p-value	2.12E-06	4.90E-05	0.0000302	0.0000363
rs6453133	HMGCR	75128	β	0.134	0.14	0.139	0.14
			$[95\%CI]$	$[0.064, 0.201]$	$[0.067, 0.202]$	$[0.065, 0.208]$	$[0.057, 0.217]$
			p-value	0.000108	0.0000435	0.00014	0.000611
rs739564	IQCK	73065	β	0.113	0.125	0.126	0.145
			$[95\%CI]$	$[0.028, 0.197]$	$[0.050, 0.209]$	$[0.036, 0.213]$	$[0.059, 0.244]$
			p-value	0.0085	0.00188	0.00523	2.34E-03
rs2272903	TFAP2B	75228	β	0.146	0.12	0.11	0.123
			$[95\%CI]$	$[0.041, 0.235]$	$[0.014, 0.213]$	$[0.006, 0.214]$	$[0.011, 0.240]$
			p-value	0.00308	0.018	0.039	0.035
rs7553158	TNNI3K	75230	β	0.091	0.096	0.109	0.13
			$[95\%CI]$	$[0.027, 0.149]$	$[0.033, 0.164]$	$[0.045, 0.181]$	$[0.065, 0.199]$
			p-value	0.00356	0.00352	0.00151	1.28E-04
rs11570094	SPI1	75200	β	0.081	0.075	0.108	0.145
			$[95\%CI]$	$[0.012, 0.151]$	$[0.009, 0.146]$	$[0.031, 0.189]$	$[0.067, 0.222]$
			p-value	2.20E-02	3.40E-02	6.37E-03	0.000221
rs4946932	FOXO3	71439	β	0.098	0.088	0.083	0.107
			$[95\%CI]$	$[0.029, 0.172]$	$[0.018, 0.159]$	$[0.012, 0.158]$	$[0.025, 0.180]$
			p-value	6.77E-03	1.40E-02	2.60E-02	6.23E-03

Continued on next page

Table A5 – *Continued from previous page*

SNP	Gene	N		45%	50%	55%	60%
rs2819347	LMOD1	75223	β	0.124	0.138	0.137	0.139
			$[95\%CI]$	$[0.055, 0.195]$	$[0.070, 0.200]$	$[0.065, 0.211]$	$[0.065, 0.216]$
			p-value	0.000474	0.0000227	2.01E-04	3.51E-04
rs2836754	ETS2	66054	β	0.077	0.081	0.09	0.085
			$[95\%CI]$	$[0.006, 0.144]$	$[0.016, 0.149]$	$[0.026, 0.168]$	$[0.013, 0.165]$
			p-value	0.031	0.017	0.013	0.029
rs2984618	TAL1	75173	β	0.07	0.072	0.061	0.078
			$[95\%CI]$	$[0.008, 0.134]$	$[0.009, 0.133]$	$[-0.002, 0.132]$	$[0.008, 0.153]$
			p-value	0.029	2.30E-02	7.30E-02	3.40E-02
rs11208662	LEPR	75177	β	0.11	0.129	0.091	0.122
			$[95\%CI]$	$[0.008, 0.218]$	$[0.014, 0.221]$	$[-0.011, 0.227]$	$[0.003, 0.242]$
			p-value	0.038	0.015	0.14	0.045
rs6235	PCSK1	75183	β	0.075	0.076	0.084	0.126
			$[95\%CI]$	$[0.005, 0.149]$	$[0.006, 0.142]$	$[0.011, 0.161]$	$[0.047, 0.206]$
			p-value	0.042	0.031	0.028	0.00203
rs9356744	CDKAL1	70863	β	0.051	0.077	0.078	0.069
			$[95\%CI]$	$[-0.019, 0.120]$	$[0.010, 0.147]$	$[0.001, 0.152]$	$[-0.010, 0.147]$
			p-value	0.15	0.027	0.045	8.60E-02
rs7988412	MTIF3	61821	β	0.151	0.16	0.179	0.175
			$[95\%CI]$	$[0.059, 0.245]$	$[0.075, 0.249]$	$[0.079, 0.271]$	$[0.082, 0.286]$
			p-value	0.00134	0.000295	0.000275	0.000938
rs1780050	NEXN	75224	β	0.026	0.024	0.034	0.051
			$[95\%CI]$	$[-0.042, 0.088]$	$[-0.040, 0.089]$	$[-0.036, 0.104]$	$[-0.025, 0.117]$
			p-value	0.439	0.472	0.346	0.153
rs526134	USP37	63859	β	0.065	0.061	0.053	0.066
			$[95\%CI]$	$[-0.002, 0.131]$	$[-0.009, 0.120]$	$[-0.014, 0.133]$	$[-0.005, 0.142]$
			p-value	0.057	0.065	0.158	0.077
rs980828	NOS1AP	75222	β	0.002	-0.005	0.007	0.018
			$[95\%CI]$	$[-0.056, 0.068]$	$[-0.065, 0.059]$	$[-0.055, 0.074]$	$[-0.051, 0.092]$

Continued on next page

Table A5 – *Continued from previous page*

SNP	Gene	N		45%	50%	55%	60%
			p-value	0.939	0.876	0.828	0.623
rs17001561	SCARB2	66079	β	0.058	0.077	0.058	0.088
			$[95\%CI]$	$[-0.028, 0.145]$	$[-0.018, 0.161]$	$[-0.035, 0.161]$	$[-0.026, 0.185]$
			p-value	0.177	0.089	0.253	0.097
rs6232	PCSK1	75225	β	0.132	0.084	0.109	0.083
			$[95\%CI]$	$[-0.002, 0.253]$	$[-0.051, 0.229]$	$[-0.047, 0.230]$	$[-0.095, 0.248]$
			p-value	0.04	0.236	0.135	0.34
rs749767	KAT8	70976	β	0.07	0.058	0.063	0.074
			$[95\%CI]$	$[0.004, 0.135]$	$[-0.008, 0.120]$	$[-0.008, 0.133]$	$[0.006, 0.156]$
			p-value	0.038	0.078	0.081	0.051
rs1211166	NTRK2	75213	β	0.031	0.051	0.053	0.053
			$[95\%CI]$	$[-0.048, 0.110]$	$[-0.030, 0.125]$	$[-0.031, 0.141]$	$[-0.037, 0.145]$
			p-value	0.455	0.201	0.224	0.252
rs2535633	ITIH4	75189	β	0.02	0.001	0.034	0.04
			$[95\%CI]$	$[-0.044, 0.082]$	$[-0.053, 0.072]$	$[-0.034, 0.095]$	$[-0.034, 0.112]$
			p-value	0.543	0.976	0.302	0.29
rs10144353	PRKCH	65613	β	-0.003	0.005	0.037	0.115
			$[95\%CI]$	$[-0.121, 0.117]$	$[-0.099, 0.129]$	$[-0.086, 0.166]$	$[-0.017, 0.252]$
			p-value	0.966	0.928	0.564	0.101
rs1561288	ADCY3	75226	β	0.08	0.06	0.051	0.036
			$[95\%CI]$	$[0.008, 0.148]$	$[-0.019, 0.133]$	$[-0.034, 0.126]$	$[-0.044, 0.122]$
			p-value	0.026	0.122	0.214	0.4
rs2283228	KCNQ1	72933	β	-0.032	-0.003	-0.007	0.001
			$[95\%CI]$	$[-0.159, 0.108]$	$[-0.133, 0.127]$	$[-0.124, 0.132]$	$[-0.141, 0.135]$
			p-value	0.634	0.958	0.92	0.983
GS-BMI		75224	β	0.102	0.109	0.119	0.13
			$[95\%CI]$	$[0.092, 0.114]$	$[0.097, 0.121]$	$[0.107, 0.133]$	$[0.117, 0.143]$
			p-value	7.88E-74	4.08E-72	7.63E-70	2.51E-82
			VarianceExplained	0.29%	0.29%	0.31%	0.35%

Part 4 - CQR Models 65-80%

SNP	Gene	N		65%	70%	75%	80%
rs1421085	FTO	75229	β	0.525	0.563	0.631	0.663
			$[95\%CI]$	$[0.444, 0.595]$	$[0.485, 0.646]$	$[0.544, 0.720]$	$[0.575, 0.762]$
			p-value	3.23E-42	1.36E-42	6.45E-44	4.19E-44
rs10767664	BDNF	74703	β	0.275	0.277	0.302	0.332
			$[95\%CI]$	$[0.183, 0.370]$	$[0.182, 0.365]$	$[0.189, 0.407]$	$[0.211, 0.447]$
			p-value	7.40E-09	4.38E-09	6.28E-08	3.70E-08
rs11672660	GIPR	72569	β	0.263	0.253	0.29	0.307
			$[95\%CI]$	$[0.173, 0.359]$	$[0.158, 0.354]$	$[0.179, 0.396]$	$[0.198, 0.424]$
			p-value	3.77E-08	2.88E-07	1.14E-07	1.61E-07
rs4788099	SH2B1	63924	β	0.17	0.185	0.217	0.23
			$[95\%CI]$	$[0.085, 0.252]$	$[0.092, 0.273]$	$[0.110, 0.314]$	$[0.124, 0.352]$
			p-value	5.73E-05	5.11E-05	3.24E-05	7.50E-05
rs7903146	TCF7L2	75230	β	0.164	0.215	0.248	0.247
			$[95\%CI]$	$[0.075, 0.241]$	$[0.119, 0.295]$	$[0.140, 0.337]$	$[0.140, 0.345]$
			p-value	1.14E-04	1.26E-06	5.99E-07	2.74E-06
rs2075650	TOMM40	74766	β	0.313	0.257	0.219	0.279
			$[95\%CI]$	$[0.210, 0.416]$	$[0.145, 0.380]$	$[0.100, 0.366]$	$[0.124, 0.404]$
			p-value	2.09E-09	1.73E-05	1.10E-03	9.63E-05
rs11873305	MC4R	75229	β	0.485	0.562	0.587	0.613
			$[95\%CI]$	$[0.322, 0.652]$	$[0.352, 0.750]$	$[0.392, 0.796]$	$[0.384, 0.835]$
			p-value	9.57E-09	3.46E-08	8.59E-09	8.01E-08
rs997295	MAP2K5	75214	β	0.162	0.189	0.183	0.177
			$[95\%CI]$	$[0.090, 0.250]$	$[0.107, 0.267]$	$[0.098, 0.279]$	$[0.089, 0.273]$
			p-value	6.39E-05	2.97E-06	7.93E-05	1.70E-04
rs3824755	NT5C2	75227	β	0.235	0.266	0.372	0.357
			$[95\%CI]$	$[0.095, 0.359]$	$[0.142, 0.425]$	$[0.207, 0.506]$	$[0.195, 0.542]$
			p-value	5.01E-04	3.25E-04	8.70E-07	4.98E-05
rs12617233	FANCL	75230	β	0.141	0.129	0.114	0.184

Continued on next page

Table A5 – *Continued from previous page*

SNP	Gene	N		65%	70%	75%	80%
			$[95\%CI]$	$[0.068, 0.222]$	$[0.049, 0.208]$	$[0.024, 0.215]$	$[0.083, 0.287]$
			p-value	0.000371	0.00157	2.10E-02	3.55E-04
rs6499653	FTO	74894	β	0.173	0.184	0.213	0.243
			$[95\%CI]$	$[0.083, 0.259]$	$[0.099, 0.284]$	$[0.100, 0.314]$	$[0.122, 0.355]$
			p-value	1.20E-04	8.35E-05	1.00E-04	0.0000359
rs1788826	NPC1	75220	β	0.13	0.143	0.184	0.185
			$[95\%CI]$	$[0.048, 0.210]$	$[0.068, 0.237]$	$[0.088, 0.279]$	$[0.085, 0.281]$
			p-value	1.58E-03	0.000778	1.62E-04	0.000174
rs17066846	MC4R	75120	β	0.166	0.195		0.177
			$[95\%CI]$	$[0.069, 0.264]$	$[0.097, 0.293]$		$[0.052, 0.306]$
			p-value	0.000698	7.97E-05		6.61E-03
rs6453133	HMGCR	75128	β	0.144	0.173	0.184	0.202
			$[95\%CI]$	$[0.057, 0.225]$	$[0.087, 0.259]$	$[0.088, 0.278]$	$[0.087, 0.298]$
			p-value	0.000827	0.0000953	1.40E-04	0.000203
rs739564	IQCK	73065	β	0.161	0.141	0.129	0.094
			$[95\%CI]$	$[0.062, 0.261]$	$[0.033, 0.249]$	$[0.016, 0.250]$	$[-0.035, 0.233]$
			p-value	1.59E-03	0.01	0.033	0.169
rs2272903	TFAP2B	75228	β	0.13	0.095	0.157	0.193
			$[95\%CI]$	$[0.010, 0.244]$	$[-0.024, 0.224]$	$[0.010, 0.280]$	$[0.039, 0.341]$
			p-value	0.032	0.133	0.021	0.013
rs7553158	TNNI3K	75230	β	0.142	0.135	0.157	0.151
			$[95\%CI]$	$[0.065, 0.222]$	$[0.055, 0.223]$	$[0.069, 0.245]$	$[0.042, 0.247]$
			p-value	4.13E-04	1.62E-03	5.07E-04	3.88E-03
rs11570094	SPI1	75200	β	0.138	0.124	0.122	0.129
			$[95\%CI]$	$[0.056, 0.222]$	$[0.033, 0.203]$	$[0.015, 0.217]$	$[0.027, 0.236]$
			p-value	9.93E-04	0.00431	1.70E-02	1.50E-02
rs4946932	FOXO3	71439	β	0.106	0.13	0.125	0.14
			$[95\%CI]$	$[0.020, 0.190]$	$[0.048, 0.212]$	$[0.028, 0.225]$	$[0.032, 0.250]$
			p-value	1.50E-02	2.09E-03	1.30E-02	1.20E-02

Continued on next page

Table A5 – *Continued from previous page*

SNP	Gene	N		65%	70%	75%	80%
rs2819347	LMOD1	75223	β	0.128	0.127	0.154	0.127
			$[95\%CI]$	$[0.049, 0.207]$	$[0.044, 0.216]$	$[0.065, 0.245]$	$[0.024, 0.227]$
			p-value	0.00148	0.00392	7.67E-04	0.013
rs2836754	ETS2	66054	β	0.093	0.102	0.121	0.091
			$[95\%CI]$	$[0.009, 0.175]$	$[0.017, 0.193]$	$[0.027, 0.223]$	$[-0.014, 0.197]$
			p-value	0.028	0.023	0.014	0.093
rs2984618	TAL1	75173	β	0.078	0.097	0.092	0.11
			$[95\%CI]$	$[0.001, 0.160]$	$[0.017, 0.187]$	$[-0.003, 0.182]$	$[0.013, 0.208]$
			p-value	5.50E-02	0.023	0.048	0.025
rs11208662	LEPR	75177	β	0.129	0.119	0.157	0.171
			$[95\%CI]$	$[-0.008, 0.268]$	$[0.003, 0.287]$	$[-0.004, 0.300]$	$[0.004, 0.335]$
			p-value	0.067	0.099	0.042	0.042
rs6235	PCSK1	75183	β	0.162	0.142	0.189	0.141
			$[95\%CI]$	$[0.079, 0.237]$	$[0.052, 0.230]$	$[0.082, 0.276]$	$[0.043, 0.255]$
			p-value	0.000073	0.00189	0.000125	0.00975
rs9356744	CDKAL1	70863	β	0.082	0.1	0.153	0.149
			$[95\%CI]$	$[-0.002, 0.167]$	$[0.024, 0.190]$	$[0.053, 0.253]$	$[0.034, 0.247]$
			p-value	0.058	0.019	0.00281	0.00607
rs7988412	MTIF3	61821	β	0.199	0.143	0.123	0.012
			$[95\%CI]$	$[0.085, 0.292]$	$[0.040, 0.256]$	$[-0.012, 0.239]$	$[-0.122, 0.142]$
			p-value	0.000163	0.01	0.056	0.854
rs1780050	NEXN	75224	β	0.072	0.084	0.069	0.037
			$[95\%CI]$	$[-0.014, 0.145]$	$[0.001, 0.164]$	$[-0.022, 0.154]$	$[-0.063, 0.138]$
			p-value	0.071	0.041	0.126	0.469
rs526134	USP37	63859	β	0.092	0.082	0.075	0.059
			$[95\%CI]$	$[0.013, 0.166]$	$[-0.008, 0.176]$	$[-0.020, 0.185]$	$[-0.044, 0.177]$
			p-value	0.018	0.082	0.148	0.289
rs980828	NOS1AP	75222	β	0.048	0.058	0.055	0.09
			$[95\%CI]$	$[-0.032, 0.122]$	$[-0.023, 0.135]$	$[-0.034, 0.145]$	$[-0.016, 0.179]$

Continued on next page

Table A5 – *Continued from previous page*

SNP	Gene	N		65%	70%	75%	80%
			p-value	0.217	0.15	0.224	7.10E-02
rs17001561	SCARB2	66079	β	0.109	0.08	0.092	0.056
			$[95\%CI]$	$[-0.002, 0.210]$	$[-0.029, 0.211]$	$[-0.050, 0.220]$	$[-0.078, 0.186]$
			p-value	0.04	0.197	0.182	0.402
rs6232	PCSK1	75225	β	0.094	0.129	0.244	0.054
			$[95\%CI]$	$[-0.081, 0.241]$	$[-0.067, 0.343]$	$[0.016, 0.447]$	$[-0.138, 0.277]$
			p-value	0.24	0.216	0.028	0.604
rs749767	KAT8	70976	β	0.069	0.033	0.066	0.064
			$[95\%CI]$	$[-0.012, 0.149]$	$[-0.046, 0.122]$	$[-0.030, 0.163]$	$[-0.042, 0.161]$
			p-value	0.093	0.437	0.173	0.218
rs1211166	NTRK2	75213	β	0.052	0.014	0.053	0.02
			$[95\%CI]$	$[-0.053, 0.135]$	$[-0.074, 0.117]$	$[-0.064, 0.169]$	$[-0.098, 0.151]$
			p-value	0.277	0.767	0.371	0.749
rs2535633	ITIH4	75189	β	0.041	0.023	0.023	0.06
			$[95\%CI]$	$[-0.044, 0.114]$	$[-0.058, 0.103]$	$[-0.064, 0.118]$	$[-0.039, 0.158]$
			p-value	0.308	0.57	0.618	0.234
rs10144353	PRKCH	65613	β	0.119	0.101	0.134	0.206
			$[95\%CI]$	$[-0.020, 0.269]$	$[-0.047, 0.265]$	$[-0.034, 0.309]$	$[0.020, 0.384]$
			p-value	0.104	0.203	0.125	0.026
rs1561288	ADCY3	75226	β	0.03	0.012	-0.002	-0.019
			$[95\%CI]$	$[-0.063, 0.116]$	$[-0.072, 0.109]$	$[-0.106, 0.106]$	$[-0.143, 0.091]$
			p-value	0.514	0.799	0.978	0.751
rs2283228	KCNQ1	72933	β	0.018	0.067	0.043	0.069
			$[95\%CI]$	$[-0.150, 0.157]$	$[-0.101, 0.208]$	$[-0.137, 0.260]$	$[-0.129, 0.242]$
			p-value	0.815	0.39	0.669	0.464
GS-BMI		75224	β	0.141	0.144	0.151	0.163
			$[95\%CI]$	$[0.127, 0.154]$	$[0.128, 0.161]$	$[0.136, 0.167]$	$[0.143, 0.181]$
			p-value	8.94E-91	1.85E-66	4.68E-84	2.40E-63
			VarianceExplained	0.38%	0.38%	0.42%	0.40%

Part 5 - CQR Models 85-95 and OLS estimates%

SNP	Gene	N		85%	90%	95%	OLS
rs1421085	FTO	75229	β	0.655	0.694	0.677	0.512
			$[95\%CI]$	$[0.552, 0.772]$	$[0.567, 0.836]$	$[0.500, 0.833]$	$[0.451, 0.572]$
			p-value	1.53E-31	4.27E-24	1.72E-15	5.88E-62
rs10767664	BDNF	74703	β	0.326	0.322	0.268	0.246
			$[95\%CI]$	$[0.198, 0.468]$	$[0.158, 0.474]$	$[0.069, 0.490]$	$[0.172, 0.319]$
			p-value	1.80E-06	8.31E-05	1.30E-02	5.89E-11
rs11672660	GIPR	72569	β	0.386	0.428	0.36	0.234
			$[95\%CI]$	$[0.262, 0.517]$	$[0.279, 0.591]$	$[0.159, 0.580]$	$[0.159, 0.309]$
			p-value	4.89E-09	8.51E-08	0.000676	8.16E-10
rs4788099	SH2B1	63924	β	0.317	0.277	0.322	0.18
			$[95\%CI]$	$[0.172, 0.422]$	$[0.139, 0.419]$	$[0.149, 0.504]$	$[0.113, 0.246]$
			p-value	8.00E-07	1.18E-04	0.000357	1.13E-07
rs7903146	TCF7L2	75230	β	0.281	0.272	0.363	0.167
			$[95\%CI]$	$[0.159, 0.400]$	$[0.120, 0.408]$	$[0.171, 0.531]$	$[0.102, 0.232]$
			p-value	5.44E-06	1.87E-04	8.11E-05	5.36E-07
rs2075650	TOMM40	74766	β	0.34	0.266	0.12	0.218
			$[95\%CI]$	$[0.178, 0.473]$	$[0.054, 0.468]$	$[-0.138, 0.388]$	$[0.131, 0.305]$
			p-value	7.20E-06	0.012	3.70E-01	9.75E-07
rs11873305	MC4R	75229	β	0.63	0.69	0.626	0.384
			$[95\%CI]$	$[0.391, 0.911]$	$[0.304, 1.017]$	$[0.158, 0.965]$	$[0.229, 0.539]$
			p-value	1.72E-06	0.000114	0.00217	1.23E-06
rs997295	MAP2K5	75214	β	0.174	0.216	0.182	0.131
			$[95\%CI]$	$[0.073, 0.280]$	$[0.091, 0.345]$	$[0.011, 0.369]$	$[0.070, 0.191]$
			p-value	9.86E-04	8.80E-04	0.047	2.40E-05
rs3824755	NT5C2	75227	β	0.385	0.386	0.262	0.218
			$[95\%CI]$	$[0.202, 0.597]$	$[0.161, 0.596]$	$[-0.006, 0.566]$	$[0.115, 0.321]$
			p-value	1.38E-04	0.000445	0.071	3.32E-05
rs12617233	FANCL	75230	β	0.243	0.266	0.364	0.128

Continued on next page

Table A5 – *Continued from previous page*

SNP	Gene	N		85%	90%	95%	OLS
			$[95\%CI]$	$[0.131, 0.351]$	$[0.125, 0.407]$	$[0.184, 0.524]$	$[0.067, 0.190]$
			p-value	1.60E-05	2.51E-04	0.0000263	4.34E-05
rs6499653	FTO	74894	β	0.252	0.243	0.249	0.142
			$[95\%CI]$	$[0.128, 0.378]$	$[0.109, 0.396]$	$[0.034, 0.443]$	$[0.073, 0.211]$
			p-value	8.49E-05	0.000914	0.018	5.19E-05
rs1788826	NPC1	75220	β	0.196	0.215	0.075	0.124
			$[95\%CI]$	$[0.080, 0.308]$	$[0.077, 0.333]$	$[-0.093, 0.247]$	$[0.061, 0.186]$
			p-value	0.000692	0.00098	0.387	1.08E-04
rs17066846	MC4R	75120	β	0.137	0.17	-0.01	0.144
			$[95\%CI]$	$[0.029, 0.296]$	$[0.013, 0.344]$	$[-0.201, 0.199]$	$[0.068, 0.220]$
			p-value	0.045	0.045	0.922	2.09E-04
rs6453133	HMGCR	75128	β	0.155	0.142	0.126	0.124
			$[95\%CI]$	$[0.042, 0.280]$	$[-0.003, 0.282]$	$[-0.070, 0.306]$	$[0.058, 0.189]$
			p-value	0.01	5.40E-02	0.192	2.18E-04
rs739564	IQCK	73065	β	0.126	0.111	0.172	0.147
			$[95\%CI]$	$[-0.011, 0.281]$	$[-0.058, 0.278]$	$[-0.080, 0.362]$	$[0.067, 0.227]$
			p-value	0.093	0.199	0.128	2.97E-04
rs2272903	TFAP2B	75228	β	0.176	0.366	0.354	0.173
			$[95\%CI]$	$[0.025, 0.371]$	$[0.165, 0.559]$	$[0.039, 0.627]$	$[0.076, 0.270]$
			p-value	4.80E-02	0.000249	0.017	4.77E-04
rs7553158	TNNI3K	75230	β	0.129	0.206	0.106	0.102
			$[95\%CI]$	$[0.023, 0.244]$	$[0.075, 0.335]$	$[-0.079, 0.286]$	$[0.042, 0.162]$
			p-value	2.20E-02	0.00218	0.253	8.40E-04
rs11570094	SPI1	75200	β	0.12	0.116	0.176	0.107
			$[95\%CI]$	$[0.014, 0.249]$	$[-0.028, 0.260]$	$[-0.027, 0.343]$	$[0.041, 0.172]$
			p-value	0.045	0.121	0.064	1.37E-03
rs4946932	FOXO3	71439	β	0.118	0.177	0.118	0.107
			$[95\%CI]$	$[-0.006, 0.242]$	$[0.035, 0.316]$	$[-0.081, 0.324]$	$[0.041, 0.174]$
			p-value	0.063	0.013	0.249	1.57E-03

Continued on next page

Table A5 – *Continued from previous page*

SNP	Gene	N		85%	90%	95%	OLS
rs2819347	LMOD1	75223	β	0.106	0.091	0.007	0.101
			$[95\%CI]$	$[0.001, 0.227]$	$[-0.059, 0.224]$	$[-0.169, 0.213]$	$[0.037, 0.165]$
			p-value	6.10E-02	2.04E-01	0.944	0.00189
rs2836754	ETS2	66054	β	0.082	0.132	0.129	0.099
			$[95\%CI]$	$[-0.029, 0.207]$	$[-0.005, 0.278]$	$[-0.065, 0.318]$	$[0.033, 0.164]$
			p-value	0.169	0.065	0.186	0.0032
rs2984618	TAL1	75173	β	0.131	0.119	0.103	0.087
			$[95\%CI]$	$[0.023, 0.242]$	$[-0.010, 0.249]$	$[-0.076, 0.275]$	$[0.026, 0.148]$
			p-value	0.019	0.071	0.244	0.00517
rs11208662	LEPR	75177	β	0.224	0.158	0.122	0.139
			$[95\%CI]$	$[0.016, 0.404]$	$[-0.067, 0.402]$	$[-0.117, 0.450]$	$[0.037, 0.242]$
			p-value	0.023	0.188	0.404	0.00766
rs6235	PCSK1	75183	β	0.132	0.135	0.204	0.09
			$[95\%CI]$	$[0.004, 0.256]$	$[-0.010, 0.291]$	$[0.024, 0.421]$	$[0.023, 0.158]$
			p-value	0.041	0.075	0.043	0.00882
rs9356744	CDKAL1	70863	β	0.19	0.186	0.175	0.071
			$[95\%CI]$	$[0.069, 0.315]$	$[0.034, 0.317]$	$[-0.003, 0.367]$	$[0.005, 0.137]$
			p-value	0.00222	0.00959	0.064	0.035
rs7988412	MTIF3	61821	β	0.04	-0.002	-0.002	0.09
			$[95\%CI]$	$[-0.120, 0.198]$	$[-0.176, 0.176]$	$[-0.240, 0.204]$	$[0.005, 0.175]$
			p-value	0.622	0.986	0.986	0.037
rs1780050	NEXN	75224	β	0.052	0.1	0.107	0.063
			$[95\%CI]$	$[-0.049, 0.169]$	$[-0.046, 0.223]$	$[-0.040, 0.312]$	$[0.002, 0.124]$
			p-value	0.354	0.143	0.233	0.042
rs526134	USP37	63859	β	0.021	0.049	0.15	0.066
			$[95\%CI]$	$[-0.081, 0.146]$	$[-0.085, 0.199]$	$[-0.010, 0.358]$	$[0.000, 0.132]$
			p-value	0.713	0.502	0.106	0.049
rs980828	NOS1AP	75222	β	0.062	0.073	0.145	0.05
			$[95\%CI]$	$[-0.044, 0.170]$	$[-0.058, 0.197]$	$[-0.010, 0.335]$	$[-0.010, 0.110]$

Continued on next page

Table A5 – *Continued from previous page*

SNP	Gene	N		85%	90%	95%	OLS
			p-value	2.54E-01	0.266	0.098	0.1
rs17001561	SCARB2	66079	β	0.057	0.039	0.014	0.07
			$[95\%CI]$	$[-0.096, 0.235]$	$[-0.153, 0.220]$	$[-0.223, 0.281]$	$[-0.017, 0.157]$
			p-value	0.495	0.678	0.909	0.113
rs6232	PCSK1	75225	β	0.071	0.028	0.115	0.095
			$[95\%CI]$	$[-0.165, 0.308]$	$[-0.296, 0.361]$	$[-0.217, 0.601]$	$[-0.041, 0.232]$
			p-value	0.563	0.869	0.577	0.172
rs749767	KAT8	70976	β	0.072	0.025	0.036	0.042
			$[95\%CI]$	$[-0.056, 0.177]$	$[-0.112, 0.165]$	$[-0.131, 0.234]$	$[-0.022, 0.105]$
			p-value	0.225	0.728	0.693	0.199
rs1211166	NTRK2	75213	β	0.044	0.055	0.044	0.041
			$[95\%CI]$	$[-0.083, 0.195]$	$[-0.101, 0.227]$	$[-0.141, 0.279]$	$[-0.034, 0.116]$
			p-value	0.536	0.516	0.686	0.289
rs2535633	ITIH4	75189	β	0.009	0.068	0.086	0.024
			$[95\%CI]$	$[-0.104, 0.123]$	$[-0.060, 0.206]$	$[-0.094, 0.251]$	$[-0.037, 0.085]$
			p-value	0.878	0.322	0.327	0.437
rs10144353	PRKCH	65613	β	0.128	-0.014	-0.224	0.044
			$[95\%CI]$	$[-0.013, 0.339]$	$[-0.227, 0.209]$	$[-0.484, 0.155]$	$[-0.067, 0.155]$
			p-value	0.148	0.897	0.175	0.441
rs1561288	ADCY3	75226	β	-0.027	-0.048	-0.104	0.024
			$[95\%CI]$	$[-0.170, 0.101]$	$[-0.212, 0.102]$	$[-0.305, 0.114]$	$[-0.047, 0.095]$
			p-value	0.693	0.55	0.329	0.507
rs2283228	KCNQ1	72933	β	-0.031	-0.047	-0.155	-0.037
			$[95\%CI]$	$[-0.246, 0.196]$	$[-0.295, 0.208]$	$[-0.487, 0.192]$	$[-0.159, 0.085]$
			p-value	0.78	0.712	0.381	0.55
GS-BMI		75224	β	0.163	0.167	0.161	0.119
			$[95\%CI]$	$[0.144, 0.184]$	$[0.140, 0.191]$	$[0.127, 0.194]$	$[0.108, 0.130]$
			p-value	1.07E-56	1.94E-37	6.24E-21	3.48E-93
			VarianceExplained	0.39%	0.36%	0.29%	0.52%

Table A6: Height-associated SNP information and results from ordinary least squares (OLS) models. 125 height associated SNPs were identified for analysis. The Effect / Other (E/O) alleles were based on original discovery studies (PMID) and SNPs were coded by height increasing alleles. Indicated positions were based on GRCh37 and all alleles were on the positive strand. The association of these SNPs with height was assessed using OLS models that were adjusted for age, sex and study. βOLS is the effect size (cm per Effect Allele) and 95%CI are the 95% confidence intervals.

SNP	Gene	E/O	PMID	β_{OLS}	$95\%CI$	p-value
rs1042725	HMGA2	C/T	17767157	0.565	0.500, 0.631	6.56E-64
rs2853977	HCP5	A/T	25282103	0.636	0.554, 0.717	1.01E-52
rs3782415	SOCS2	C/T	PMC3014369	0.464	0.383, 0.545	4.01E-29
rs780094	GCKR	C/T	25282103	0.372	0.306, 0.439	8.74E-28
rs9892365	TBX2	A/G	PMC3014369	0.356	0.286, 0.426	2.28E-23
rs7137534	PDE3A	T/C	25282103	0.352	0.282, 0.421	6.03E-23
rs1776897	HMGA1	G/T	20397748	0.610	0.488, 0.732	1.15E-22
rs572169	GHSR	T/C	20881960	0.351	0.280, 0.422	3.14E-22
rs2679178	NPPC	C/T	PMC3014369	0.527	0.410, 0.644	1.17E-18
rs2053156	GRB2	T/G	25282103	0.371	0.286, 0.456	1.15E-17
rs9930741	ERI2	T/C	25282103	0.285	0.219, 0.352	3.92E-17
rs2854207	CSH2	G/C	25282103	0.314	0.237, 0.392	2.11E-15
rs4320932	IGF2	T/C	25282103	0.336	0.252, 0.419	4.11E-15
rs752313	EZH1	C/T	25282103	0.258	0.191, 0.326	4.78E-14
rs709939	SAMD4A	T/C	25282103	0.248	0.181, 0.314	2.37E-13
rs1036477	FBN1	A/G	25282103	0.377	0.272, 0.483	2.34E-12
rs158676	CDK5RAP1	A/G	25282103	0.254	0.183, 0.325	2.39E-12
rs1822469	PPP3R1	C/T	25282103	0.243	0.175, 0.312	4.23E-12
rs258281	RAB26	G/A	25282103	0.308	0.220, 0.397	1.06E-11
rs9366637	HFE	C/T	25282103	0.470	0.334, 0.606	1.22E-11
rs551219	COL24A1	T/C	25282103	0.249	0.177, 0.321	1.30E-11
rs13072536	ITIH4	A/T	25282103	0.265	0.188, 0.342	1.71E-11
rs7522692	PIGC	G/A	25282103	0.290	0.205, 0.375	2.24E-11
rs13076290	CTNNB1	T/C	25282103	0.225	0.158, 0.291	3.23E-11
rs1636255	GNA12	C/A	PMC3014369	0.261	0.184, 0.338	3.58E-11
rs3796529	REST	T/C	25282103	0.277	0.194, 0.361	7.85E-11
rs1866146	POMC	G/A	25282103	0.229	0.160, 0.299	8.47E-11
rs9844666	PCCB	G/A	20881960	0.254	0.177, 0.331	9.23E-11
rs8071847	POLR2A	G/A	25282103	0.264	0.184, 0.344	1.10E-10
rs3783937	FBLN5	C/T	25282103	0.249	0.173, 0.326	1.72E-10
rs11080149	NF1	T/C	25282103	0.319	0.221, 0.418	2.48E-10
rs17472113	ZAR1	A/T	25282103	0.270	0.186, 0.354	3.23E-10
rs490634	CISH	C/T	25282103	0.346	0.238, 0.455	4.04E-10
rs17622208	SLC22A5	A/G	25282103	0.209	0.143, 0.275	5.08E-10
rs2982712	ESR1	C/T	23563607	0.209	0.143, 0.275	6.50E-10

Continued on next page

Table A6 – *Continued from previous page*

SNP	Gene	E/O	PMID	β_{OLS}	95%CI	p-value
rs1950500	NFATC4	T/C	20881960	0.240	0.163, 0.317	1.09E-09
rs1476387	PPIL6	G/T	25282103	0.206	0.139, 0.272	1.32E-09
rs4946932	FOXO3	C/A	25282103	0.220	0.148, 0.291	1.61E-09
rs1800783	NOS3	T/A	25282103	0.214	0.144, 0.284	1.71E-09
rs6718902	STAT1	T/C	25282103	0.233	0.157, 0.309	1.98E-09
rs2425019	MMP24	G/A	PMC3014369	0.205	0.138, 0.272	2.55E-09
rs6731022	EIF2AK3	C/G	25282103	0.220	0.147, 0.292	2.65E-09
rs12940055	MAP3K3	C/T	25282103	0.311	0.207, 0.416	5.15E-09
rs864745	JAZF1	T/C	25282103	0.213	0.141, 0.284	5.21E-09
rs6487088	PDE3A	T/C	25282103	0.246	0.163, 0.328	5.63E-09
rs4973410	NCL	C/T	25282103	0.205	0.136, 0.275	7.37E-09
rs451061	PRKCZ	C/G	25282103	0.201	0.132, 0.270	9.78E-09
rs832575	MAP3K1	A/G	25282103	0.287	0.187, 0.386	1.58E-08
rs4955526	EPHB1	C/T	25282103	0.207	0.135, 0.280	1.94E-08
rs8038415	IGF1R	C/T	25282103	0.189	0.123, 0.255	2.12E-08
rs7578199	HDLBP	T/C	25282103	0.215	0.139, 0.291	2.85E-08
rs7020782	PAPPA	A/C	25282103	0.202	0.130, 0.274	3.48E-08
rs2229712	RPS6KA1	A/C	25282103	0.263	0.168, 0.357	5.06E-08
rs7572476	BOK	C/T	25282103	0.186	0.119, 0.253	5.45E-08
rs2066807	PAN2	G/C	20881960	0.232	0.147, 0.317	8.76E-08
rs1516796	ACAN	A/C	25282103	0.180	0.112, 0.247	2.04E-07
rs6180	GHR	A/C	25429064	0.175	0.109, 0.240	2.11E-07
rs8055190	LRRC36	C/T	25282103	0.449	0.279, 0.620	2.30E-07
rs17106235	FAF1	G/C	25282103	0.316	0.195, 0.436	3.11E-07
rs3739707	LPAR1	C/A	25282103	0.197	0.121, 0.272	3.22E-07
rs674424	ABCG4	T/C	25282103	0.189	0.112, 0.265	1.32E-06
rs12225387	NEU3	G/A	25282103	0.199	0.118, 0.280	1.39E-06
rs3812265	CNOT4	T/C	25282103	0.191	0.113, 0.270	1.72E-06
rs10208728	IHH	A/G	25282103	0.285	0.167, 0.402	1.92E-06
rs291979	GRK5	A/G	25282103	0.189	0.110, 0.267	2.37E-06
rs2715553	RARA	A/G	25282103	0.156	0.088, 0.225	7.30E-06
rs2057291	GNAS	A/G	25282103	0.159	0.088, 0.230	1.18E-05
rs4803520	GRIK5	G/A	25282103	0.235	0.129, 0.341	1.45E-05
rs10736682	APLNR	G/A	25282103	0.143	0.077, 0.210	2.31E-05
rs2909430	TP53	T/C	25282103	0.209	0.110, 0.308	3.57E-05
rs12050767	CYP19A1	C/T	25282103	0.139	0.073, 0.205	3.62E-05
rs602633	PSRC1	T/G	25282103	0.179	0.094, 0.264	3.64E-05
rs1738475	HTR1D	C/G	20881960	0.142	0.072, 0.213	7.70E-05
rs17754	RFC1	C/G	25282103	0.130	0.063, 0.197	1.29E-04
rs17541471	NPR3	C/T	25282103	0.165	0.080, 0.250	1.34E-04
rs1342586	TGFB2	T/C	25282103	0.157	0.075, 0.238	1.69E-04
rs3814115	PCSK5	C/T	25282103	0.133	0.063, 0.203	2.02E-04

Continued on next page

Table A6 – *Continued from previous page*

SNP	Gene	E/O	PMID	β_{OLS}	95%CI	p-value
rs1780616	LBP	C/T	25282103	0.130	[0.061, 0.200]	2.28E-04
rs3736228	LRP5	C/T	25282103	0.175	[0.081, 0.269]	2.74E-04
rs212517	ECE1	A/T	25282103	0.126	[0.058, 0.195]	2.94E-04
rs7359336	NFAT5	G/A	25282103	0.120	[0.053, 0.187]	4.64E-04
rs2682552	XRCC1	A/T	25282103	0.148	[0.064, 0.232]	5.35E-04
rs17085675	PCSK1	T/A	25282103	0.131	[0.057, 0.206]	5.54E-04
rs11102986	RXRA	G/A	25282103	0.151	[0.064, 0.238]	6.49E-04
rs12603813	PLCD3	T/C	25282103	0.135	[0.057, 0.212]	6.61E-04
rs6219	IGF1	T/C	25282103	0.190	[0.080, 0.299]	7.01E-04
rs2234693	ESR1	C/T	25282103	0.113	[0.045, 0.180]	0.001
rs46522	UBE2Z	C/T	25282103	0.115	[0.046, 0.185]	0.001
rs891088	INSR	G/A	20881960	0.123	[0.049, 0.197]	0.001
rs9857730	VILL	C/T	25282103	0.132	[0.052, 0.213]	0.001
rs10185680	MFSD2B	G/A	25282103	0.108	[0.042, 0.173]	0.001
rs7163907	PTPN9	C/T	25282103	-0.121	[−0.196, −0.046]	0.002
rs2176167	NOP58	C/T	25282103	0.109	[0.040, 0.179]	0.002
rs7557989	THADA	T/C	25282103	0.114	[0.040, 0.188]	0.003
rs510769	OPRM1	T/C	25282103	0.123	[0.043, 0.204]	0.003
rs7756224	NMBR	C/T	25282103	0.111	[0.038, 0.184]	0.003
rs2282537	POU2F3	G/A	25282103	0.152	[0.052, 0.252]	0.003
rs7853859	CENPP	T/C	25282103	0.105	[0.036, 0.175]	0.003
rs2229642	ITPR3	G/C	25282103	0.112	[0.033, 0.190]	0.005
rs5015437	LMF1	A/G	25282103	0.102	[0.030, 0.174]	0.006
rs7963565	KNTC1	T/C	25282103	0.097	[0.028, 0.167]	0.006
rs696	NFKBIA	C/T	25282103	0.092	[0.024, 0.160]	0.008
rs25656	NFATC1	A/G	25282103	0.114	[0.029, 0.199]	0.009
rs3100776	IHH	C/T	25282103	0.243	[0.061, 0.426]	0.009
rs12145922	PKN2	A/C	25282103	0.088	[0.020, 0.156]	0.011
rs526134	USP37	A/G	25282103	0.094	[0.022, 0.167]	0.011
rs7004280	RPS20	C/G	25282103	0.207	[0.018, 0.396]	0.032
rs4808199	GATAD2A	G/A	25282103	0.098	[0.007, 0.189]	0.035
rs3821009	PDE11A	T/C	25282103	0.129	[0.006, 0.252]	0.040
rs2291617	METTL1	T/G	25282103	0.073	[0.003, 0.144]	0.042
rs1051168	SEC11A	T/G	25282103	0.080	[0.002, 0.157]	0.043
rs803932	ASTN2	C/T	25282103	0.073	[−0.001, 0.147]	0.054
rs2247870	GPR98	A/G	25282103	0.066	[−0.001, 0.134]	0.055
rs12503378	NUDT6	C/G	25282103	0.077	[−0.011, 0.166]	0.085
rs2145923	NPR2	C/T	25282103	0.076	[−0.011, 0.163]	0.088
rs12603582	ITGB3	G/T	25282103	0.060	[−0.019, 0.139]	0.137
rs4864548	CLOCK	A/G	25282103	0.045	[−0.022, 0.113]	0.190
rs2735469	MRPL23	A/G	25282103	0.061	[−0.033, 0.154]	0.204
rs1051431	MPHOSPH9	G/A	25282103	0.038	[−0.041, 0.117]	0.341

Continued on next page

Table A6 – *Continued from previous page*

SNP	Gene	E/O	PMID	β_{OLS}	$95\%CI$	p-value
rs41132	AP3B1	A/C	25282103	0.037	$-0.040, 0.115$	0.344
rs6127698	MC3R	G/T	25282103	0.030	$-0.035, 0.096$	0.368
rs1481892	ARNTL	G/C	25282103	0.033	$-0.039, 0.105$	0.370
rs2633442	MKRN2	G/A	25282103	0.023	$-0.049, 0.095$	0.527
rs1535	FADS1	A/G	25429064	-0.011	$-0.081, 0.059$	0.754
rs10861148	HSP90B1	A/C	25282103	0.005	$-0.099, 0.110$	0.921

Table A7: Conditional quantile regression (CQR) models of Height-associated SNPs and GS-Height across the sample distribution. CQR models were fitted every 5th percentile of height and adjusted for age, sex and study. β from ordinary least squares (OLS) and CQR models at each percentile are the effect sizes (cm per Effect Allele). 95%CI are the 95% confidence intervals. In addition, the proportion of height variance that is explained by the GS-Height was estimated, Variance Explained (%).

Part 1 - 5% to 20%

SNP	Gene	N		5%	10%	15%	20%
rs1042725	HMGA2	73105	β	0.594	0.603	0.586	0.607
			$[95\%CI]$	$[0.480, 0.739]$	$[0.487, 0.698]$	$[0.492, 0.688]$	$[0.513, 0.708]$
			p-value	1.80E-19	3.70E-29	2.79E-31	9.99E-34
rs2853977	HCP5	44699	β	0.757	0.706	0.652	0.642
			$[95\%CI]$	$[0.561, 0.904]$	$[0.565, 0.823]$	$[0.527, 0.775]$	$[0.525, 0.752]$
			p-value	3.71E-18	5.19E-27	1.85E-24	1.85E-28
rs3782415	SOCS2	73568	β	0.403	0.482	0.536	0.5
			$[95\%CI]$	$[0.243, 0.574]$	$[0.330, 0.619]$	$[0.401, 0.655]$	$[0.389, 0.608]$
			p-value	1.13E-06	5.57E-11	1.21E-16	4.74E-19
rs780094	GCKR	73548	β	0.293	0.359	0.395	0.392
			$[95\%CI]$	$[0.142, 0.426]$	$[0.248, 0.474]$	$[0.295, 0.492]$	$[0.302, 0.491]$
			p-value	4.25E-05	3.44E-10	7.03E-15	2.61E-16
rs9892365	TBX2	73566	β	0.244	0.25	0.305	0.316
			$[95\%CI]$	$[0.101, 0.374]$	$[0.135, 0.360]$	$[0.196, 0.406]$	$[0.219, 0.407]$
			p-value	4.16E-04	1.18E-05	1.43E-08	3.46E-11
rs7137534	PDE3A	73567	β	0.258	0.274	0.29	0.331
			$[95\%CI]$	$[0.126, 0.387]$	$[0.161, 0.398]$	$[0.175, 0.392]$	$[0.236, 0.431]$
			p-value	1.14E-04	6.09E-06	1.68E-07	1.55E-11
rs1776897	HMGA1	64379	β	0.38	0.404	0.498	0.538
			$[95\%CI]$	$[0.133, 0.647]$	$[0.208, 0.581]$	$[0.296, 0.721]$	$[0.392, 0.744]$
			p-value	3.80E-03	2.27E-05	3.56E-06	1.93E-09
rs572169	GHSR	73554	β	0.344	0.398	0.4	0.372
			$[95\%CI]$	$[0.184, 0.483]$	$[0.275, 0.510]$	$[0.286, 0.499]$	$[0.265, 0.468]$
			p-value	0.00000576	3.70E-11	1.63E-13	5.55E-13
rs2679178	NPPC	73567	β	0.572	0.448	0.465	0.516
			$[95\%CI]$	$[0.327, 0.809]$	$[0.260, 0.646]$	$[0.281, 0.635]$	$[0.354, 0.682]$
			p-value	2.05E-06	5.37E-06	3.18E-07	5.26E-10
rs2053156	GRB2	73536	β	0.393	0.373	0.328	0.375
			$[95\%CI]$	$[0.224, 0.536]$	$[0.238, 0.511]$	$[0.202, 0.448]$	$[0.232, 0.465]$
			p-value	7.34E-07	8.50E-08	1.58E-07	3.10E-10

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		5%	10%	15%	20%
rs9930741	ERI2	73143	β	0.301	0.382	0.356	0.334
			$[95\%CI]$	$[0.161, 0.438]$	$[0.264, 0.476]$	$[0.249, 0.459]$	$[0.250, 0.431]$
			p-value	1.89E-05	1.25E-12	1.74E-11	3.78E-13
rs2854207	CSH2	67567	β	0.315	0.354	0.259	0.227
			$[95\%CI]$	$[0.145, 0.487]$	$[0.235, 0.475]$	$[0.151, 0.384]$	$[0.121, 0.327]$
			p-value	2.71E-04	6.00E-09	1.65E-05	1.57E-05
rs4320932	IGF2	71204	β	0.313	0.35	0.359	0.359
			$[95\%CI]$	$[0.157, 0.497]$	$[0.205, 0.495]$	$[0.227, 0.481]$	$[0.239, 0.485]$
			p-value	3.63E-04	2.84E-06	2.62E-08	1.18E-08
rs752313	EZH1	69774	β	0.391	0.32	0.274	0.247
			$[95\%CI]$	$[0.244, 0.516]$	$[0.201, 0.424]$	$[0.175, 0.385]$	$[0.152, 0.337]$
			p-value	1.63E-08	2.13E-08	2.69E-07	1.20E-07
rs709939	SAMD4A	73569	β	0.207	0.202	0.205	0.204
			$[95\%CI]$	$[0.080, 0.339]$	$[0.084, 0.303]$	$[0.096, 0.301]$	$[0.125, 0.306]$
			p-value	0.00163	3.14E-04	1.08E-04	1.25E-05
rs1036477	FBN1	73539	β	0.381	0.344	0.464	0.453
			$[95\%CI]$	$[0.203, 0.603]$	$[0.151, 0.533]$	$[0.306, 0.599]$	$[0.292, 0.599]$
			p-value	1.61E-04	5.01E-04	6.57E-10	7.95E-09
rs158676	CDK5RAP1	73562	β	0.225	0.215	0.236	0.25
			$[95\%CI]$	$[0.107, 0.380]$	$[0.088, 0.331]$	$[0.111, 0.346]$	$[0.159, 0.348]$
			p-value	1.37E-03	6.14E-04	7.72E-05	2.00E-07
rs1822469	PPP3R1	69710	β	0.218	0.237	0.297	0.291
			$[95\%CI]$	$[0.083, 0.348]$	$[0.127, 0.348]$	$[0.196, 0.399]$	$[0.203, 0.389]$
			p-value	1.18E-03	2.64E-05	1.75E-08	1.16E-09
rs258281	RAB26	67411	β	0.36	0.477	0.423	0.318
			$[95\%CI]$	$[0.208, 0.542]$	$[0.340, 0.625]$	$[0.288, 0.558]$	$[0.222, 0.476]$
			p-value	1.66E-05	4.86E-11	5.73E-10	8.96E-07
rs9366637	HFE	73557	β	0.4	0.5	0.432	0.4
			$[95\%CI]$	$[0.164, 0.706]$	$[0.279, 0.741]$	$[0.246, 0.641]$	$[0.253, 0.622]$
			p-value	3.71E-03	2.00E-05	1.90E-05	2.20E-05
rs551219	COL24A1	73076	β	0.352	0.354	0.34	0.281
			$[95\%CI]$	$[0.210, 0.483]$	$[0.224, 0.464]$	$[0.225, 0.440]$	$[0.181, 0.391]$
			p-value	3.54E-07	4.51E-09	4.59E-10	1.04E-07
rs13072536	ITIH4	73519	β	0.215	0.296	0.304	0.269
			$[95\%CI]$	$[0.077, 0.386]$	$[0.172, 0.409]$	$[0.181, 0.412]$	$[0.186, 0.401]$
			p-value	6.00E-03	8.32E-07	2.72E-07	8.97E-07

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		5%	10%	15%	20%
rs7522692	PIGC	64398	β	0.4	0.308	0.314	0.254
			$[95\%CI]$	$[0.208, 0.558]$	$[0.158, 0.464]$	$[0.185, 0.454]$	$[0.165, 0.399]$
			p-value	7.61E-06	8.97E-05	6.02E-06	2.06E-05
rs13076290	CTNNB1	73154	β	0.269	0.307	0.278	0.243
			$[95\%CI]$	$[0.149, 0.398]$	$[0.194, 0.420]$	$[0.182, 0.371]$	$[0.145, 0.325]$
			p-value	1.87E-05	8.42E-08	1.22E-08	1.05E-07
rs1636255	GNA12	64371	β	0.217	0.224	0.286	0.262
			$[95\%CI]$	$[0.074, 0.418]$	$[0.107, 0.339]$	$[0.171, 0.404]$	$[0.170, 0.382]$
			p-value	1.40E-02	1.68E-04	1.08E-06	1.31E-06
rs3796529	REST	73119	β	0.393	0.345	0.36	0.304
			$[95\%CI]$	$[0.238, 0.541]$	$[0.201, 0.471]$	$[0.241, 0.479]$	$[0.202, 0.424]$
			p-value	3.47E-07	7.25E-07	3.85E-09	5.76E-08
rs1866146	POMC	73516	β	0.1	0.078	0.128	0.138
			$[95\%CI]$	$[-0.041, 0.242]$	$[-0.031, 0.195]$	$[0.019, 0.223]$	$[0.076, 0.256]$
			p-value	0.165	1.72E-01	1.30E-02	2.54E-03
rs9844666	PCCB	73511	β	0.175	0.232	0.333	0.317
			$[95\%CI]$	$[0.028, 0.335]$	$[0.108, 0.351]$	$[0.213, 0.438]$	$[0.215, 0.432]$
			p-value	0.025	1.99E-04	4.01E-09	1.13E-08
rs8071847	POLR2A	73565	β	0.177	0.227	0.32	0.338
			$[95\%CI]$	$[0.009, 0.324]$	$[0.100, 0.353]$	$[0.197, 0.454]$	$[0.222, 0.450]$
			p-value	0.027	0.000464	1.12E-06	8.97E-09
rs3783937	FBLN5	73104	β	0.235	0.22	0.226	0.241
			$[95\%CI]$	$[0.100, 0.393]$	$[0.099, 0.352]$	$[0.109, 0.338]$	$[0.115, 0.334]$
			p-value	0.00161	6.01E-04	9.62E-05	1.31E-05
rs11080149	NF1	73567	β	0.355	0.407	0.36	0.3
			$[95\%CI]$	$[0.114, 0.579]$	$[0.255, 0.556]$	$[0.199, 0.488]$	$[0.166, 0.429]$
			p-value	2.47E-03	7.12E-08	1.01E-06	6.10E-06
rs17472113	ZAR1	58807	β	0.094	0.242	0.307	0.282
			$[95\%CI]$	$[-0.071, 0.275]$	$[0.108, 0.379]$	$[0.179, 0.425]$	$[0.174, 0.415]$
			p-value	0.279	4.95E-04	9.24E-07	4.30E-06
rs490634	CISH	61124	β	0.423	0.424	0.471	0.38
			$[95\%CI]$	$[0.174, 0.593]$	$[0.219, 0.592]$	$[0.299, 0.633]$	$[0.262, 0.548]$
			p-value	0.0000592	1.25E-05	1.59E-08	1.84E-07
rs17622208	SLC22A5	73531	β	0.126	0.173	0.196	0.16
			$[95\%CI]$	$[0.003, 0.275]$	$[0.055, 0.267]$	$[0.091, 0.286]$	$[0.088, 0.275]$
			p-value	0.068	1.28E-03	8.48E-05	7.66E-04

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		5%	10%	15%	20%
rs2982712	ESR1	73566	β	0.25	0.235	0.217	0.176
			$[95\%CI]$	$[0.102, 0.389]$	$[0.117, 0.337]$	$[0.122, 0.316]$	$[0.089, 0.275]$
			p-value	6.56E-04	3.19E-05	9.54E-06	2.18E-04
rs1950500	NFATC4	64408	β	0.259	0.265	0.234	0.208
			$[95\%CI]$	$[0.103, 0.410]$	$[0.133, 0.368]$	$[0.111, 0.349]$	$[0.115, 0.311]$
			p-value	9.47E-04	8.79E-06	8.56E-05	2.88E-05
rs1476387	PPIL6	73563	β	0.195	0.274	0.293	0.27
			$[95\%CI]$	$[0.076, 0.339]$	$[0.168, 0.372]$	$[0.190, 0.387]$	$[0.180, 0.366]$
			p-value	0.00403	0.000000128	6.31E-09	8.15E-09
rs4946932	FOXO3	73565	β	0.217	0.199	0.243	0.244
			$[95\%CI]$	$[0.079, 0.358]$	$[0.087, 0.323]$	$[0.136, 0.343]$	$[0.138, 0.330]$
			p-value	2.20E-03	8.21E-04	3.85E-06	5.73E-07
rs1800783	NOS3	71277	β	0.249	0.212	0.238	0.215
			$[95\%CI]$	$[0.091, 0.365]$	$[0.094, 0.315]$	$[0.126, 0.331]$	$[0.109, 0.302]$
			p-value	0.000326	1.69E-04	5.25E-06	1.26E-05
rs6718902	STAT1	73557	β	0.393	0.265	0.306	0.291
			$[95\%CI]$	$[0.248, 0.537]$	$[0.140, 0.397]$	$[0.199, 0.412]$	$[0.191, 0.397]$
			p-value	7.61E-08	3.79E-05	1.25E-08	2.70E-08
rs2425019	MMP24	69770	β	0.24	0.212	0.169	0.185
			$[95\%CI]$	$[0.096, 0.377]$	$[0.096, 0.308]$	$[0.079, 0.281]$	$[0.093, 0.281]$
			p-value	7.85E-04	1.04E-04	1.19E-03	1.20E-04
rs6731022	EIF2AK3	69767	β	0.159	0.183	0.204	0.222
			$[95\%CI]$	$[0.005, 0.275]$	$[0.053, 0.299]$	$[0.106, 0.313]$	$[0.116, 0.312]$
			p-value	0.02	0.00371	1.18E-04	9.56E-06
rs12940055	MAP3K3	73566	β	0.376	0.304	0.269	0.269
			$[95\%CI]$	$[0.147, 0.556]$	$[0.125, 0.493]$	$[0.110, 0.438]$	$[0.140, 0.418]$
			p-value	0.000264	0.00141	0.00136	1.51E-04
rs864745	JAZF1	60616	β	0.203	0.22	0.187	0.132
			$[95\%CI]$	$[0.069, 0.356]$	$[0.113, 0.339]$	$[0.074, 0.281]$	$[0.045, 0.242]$
			p-value	5.77E-03	1.22E-04	3.81E-04	8.36E-03
rs6487088	PDE3A	73563	β	0.222	0.296	0.301	0.244
			$[95\%CI]$	$[0.063, 0.406]$	$[0.140, 0.419]$	$[0.167, 0.420]$	$[0.124, 0.359]$
			p-value	1.20E-02	3.66E-05	3.35E-06	5.27E-05
rs4973410	NCL	64408	β	0.252	0.2	0.193	0.195
			$[95\%CI]$	$[0.103, 0.391]$	$[0.089, 0.325]$	$[0.103, 0.303]$	$[0.106, 0.306]$
			p-value	0.000668	9.83E-04	1.62E-04	1.47E-04

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		5%	10%	15%	20%
rs451061	PRKCZ	71365	β	0.355	0.225	0.23	0.166
			$[95\%CI]$	$[0.209, 0.470]$	$[0.107, 0.341]$	$[0.122, 0.328]$	$[0.091, 0.267]$
			p-value	9.34E-08	0.000189	1.24E-05	1.77E-04
rs832575	MAP3K1	73539	β	0.164	0.348	0.339	0.324
			$[95\%CI]$	$[-0.002, 0.405]$	$[0.165, 0.503]$	$[0.172, 0.482]$	$[0.182, 0.453]$
			p-value	0.112	5.27E-05	1.89E-05	2.73E-06
rs4955526	EPHB1	64417	β	0.178	0.194	0.18	0.227
			$[95\%CI]$	$[0.009, 0.329]$	$[0.067, 0.321]$	$[0.083, 0.292]$	$[0.117, 0.319]$
			p-value	0.03	2.50E-03	8.20E-04	9.78E-06
rs8038415	IGF1R	73516	β	0.217	0.214	0.22	0.201
			$[95\%CI]$	$[0.075, 0.346]$	$[0.095, 0.331]$	$[0.110, 0.312]$	$[0.113, 0.292]$
			p-value	1.83E-03	3.53E-04	1.58E-05	1.13E-05
rs7578199	HDLBP	73569	β	0.193	0.239	0.296	0.249
			$[95\%CI]$	$[0.058, 0.364]$	$[0.114, 0.372]$	$[0.177, 0.411]$	$[0.160, 0.370]$
			p-value	1.40E-02	0.000301	6.48E-07	2.22E-06
rs7020782	PAPPA	73566	β	0.283	0.227	0.223	0.155
			$[95\%CI]$	$[0.130, 0.412]$	$[0.106, 0.338]$	$[0.107, 0.317]$	$[0.066, 0.272]$
			p-value	8.37E-05	1.30E-04	2.90E-05	3.28E-03
rs2229712	RPS6KA1	48240	β	0.216	0.286	0.316	0.331
			$[95\%CI]$	$[0.039, 0.416]$	$[0.128, 0.423]$	$[0.166, 0.443]$	$[0.182, 0.447]$
			p-value	0.025	1.55E-04	6.82E-06	1.03E-06
rs7572476	BOK	71367	β	0.235	0.236	0.217	0.188
			$[95\%CI]$	$[0.097, 0.359]$	$[0.111, 0.335]$	$[0.108, 0.310]$	$[0.097, 0.274]$
			p-value	3.94E-04	5.23E-05	2.54E-05	3.43E-05
rs2066807	PAN2	71822	β	0.262	0.231	0.194	0.147
			$[95\%CI]$	$[0.081, 0.428]$	$[0.078, 0.352]$	$[0.053, 0.303]$	$[0.023, 0.282]$
			p-value	0.00295	8.03E-04	2.76E-03	2.70E-02
rs1516796	ACAN	69308	β	0.219	0.181	0.204	0.15
			$[95\%CI]$	$[0.079, 0.357]$	$[0.068, 0.302]$	$[0.096, 0.302]$	$[0.071, 0.260]$
			p-value	0.00183	0.00219	1.08E-04	1.84E-03
rs6180	GHR	73552	β	0.149	0.219	0.251	0.245
			$[95\%CI]$	$[0.032, 0.271]$	$[0.099, 0.310]$	$[0.155, 0.354]$	$[0.150, 0.329]$
			p-value	1.40E-02	4.31E-05	6.27E-07	7.53E-08
rs8055190	LRRC36	64416	β	0.658	0.56	0.4	0.391
			$[95\%CI]$	$[0.306, 0.970]$	$[0.244, 0.854]$	$[0.169, 0.664]$	$[0.181, 0.627]$
			p-value	0.000104	0.000339	0.00165	0.000608

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		5%	10%	15%	20%
rs17106235	FAF1	60489	β	0.29	0.136	0.24	0.216
			$[95\%CI]$	$[0.023, 0.543]$	$[-0.074, 0.345]$	$[0.040, 0.392]$	$[0.060, 0.398]$
			p-value	0.029	2.05E-01	6.81E-03	1.30E-02
rs3739707	LPAR1	73568	β	0.215	0.264	0.261	0.223
			$[95\%CI]$	$[0.074, 0.366]$	$[0.138, 0.388]$	$[0.146, 0.380]$	$[0.110, 0.330]$
			p-value	0.00362	3.39E-05	1.23E-05	6.13E-05
rs674424	ABCG4	73568	β	0.182	0.142	0.166	0.188
			$[95\%CI]$	$[0.026, 0.327]$	$[0.013, 0.263]$	$[0.049, 0.279]$	$[0.109, 0.317]$
			p-value	0.018	0.025	0.00445	0.000386
rs12225387	NEU3	60621	β	0.211	0.132	0.159	0.236
			$[95\%CI]$	$[0.048, 0.377]$	$[0.009, 0.275]$	$[0.046, 0.313]$	$[0.096, 0.309]$
			p-value	0.011	5.00E-02	2.00E-02	1.37E-05
rs3812265	CNOT4	69776	β	0.136	0.115	0.144	0.165
			$[95\%CI]$	$[-0.017, 0.278]$	$[-0.012, 0.258]$	$[0.013, 0.259]$	$[0.064, 0.264]$
			p-value	0.073	0.094	2.10E-02	1.23E-03
rs10208728	IHH	64409	β	0.278	0.296	0.258	0.23
			$[95\%CI]$	$[0.074, 0.554]$	$[0.084, 0.473]$	$[0.088, 0.436]$	$[0.081, 0.400]$
			p-value	0.023	3.27E-03	4.01E-03	4.43E-03
rs291979	GRK5	73568	β	0.169	0.217	0.164	0.073
			$[95\%CI]$	$[-0.008, 0.326]$	$[0.082, 0.346]$	$[0.051, 0.271]$	$[-0.030, 0.182]$
			p-value	0.049	0.00127	0.00357	0.17
rs2715553	RARA	69206	β	0.241	0.187	0.195	0.172
			$[95\%CI]$	$[0.097, 0.364]$	$[0.070, 0.294]$	$[0.087, 0.298]$	$[0.084, 0.269]$
			p-value	3.40E-04	1.02E-03	2.92E-04	2.47E-04
rs2057291	GNAS	70100	β	0.111	0.193	0.192	0.153
			$[95\%CI]$	$[-0.040, 0.240]$	$[0.069, 0.306]$	$[0.086, 0.300]$	$[0.068, 0.262]$
			p-value	1.20E-01	1.67E-03	4.86E-04	2.05E-03
rs4803520	GRIK5	69747	β	0.227	0.335	0.404	0.267
			$[95\%CI]$	$[0.018, 0.435]$	$[0.152, 0.492]$	$[0.224, 0.557]$	$[0.139, 0.456]$
			p-value	0.033	1.27E-04	2.18E-06	9.33E-04
rs10736682	APLNR	73090	β	0.145	0.185	0.211	0.145
			$[95\%CI]$	$[-0.003, 0.266]$	$[0.064, 0.292]$	$[0.108, 0.300]$	$[0.078, 0.246]$
			p-value	0.033	1.10E-03	1.64E-05	0.000784
rs2909430	TP53	73412	β	0.217	0.317	0.286	0.229
			$[95\%CI]$	$[0.019, 0.407]$	$[0.132, 0.451]$	$[0.139, 0.430]$	$[0.092, 0.374]$
			p-value	2.80E-02	1.01E-04	1.09E-04	1.58E-03

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		5%	10%	15%	20%
rs12050767	CYP19A1	73563	β	0.141	0.17	0.155	0.125
			$[95\%CI]$	$[0.011, 0.277]$	$[0.050, 0.264]$	$[0.055, 0.259]$	$[0.045, 0.229]$
			p-value	0.037	0.00209	0.00285	0.00705
rs602633	PSRC1	64375	β	0.218	0.205	0.208	0.179
			$[95\%CI]$	$[0.072, 0.413]$	$[0.065, 0.344]$	$[0.073, 0.333]$	$[0.088, 0.296]$
			p-value	0.014	0.00383	1.49E-03	8.23E-04
rs1738475	HTR1D	64411	β	0.258	0.25	0.191	0.178
			$[95\%CI]$	$[0.123, 0.399]$	$[0.128, 0.360]$	$[0.094, 0.303]$	$[0.080, 0.288]$
			p-value	0.000248	0.0000212	0.000383	8.00E-04
rs17754	RFC1	73558	β	0.058	0.1	0.028	0.099
			$[95\%CI]$	$[-0.055, 0.192]$	$[-0.015, 0.205]$	$[-0.064, 0.137]$	$[0.002, 0.182]$
			p-value	0.349	0.072	0.588	0.03
rs17541471	NPR3	69777	β	0.151	0.119	0.121	0.106
			$[95\%CI]$	$[-0.047, 0.300]$	$[-0.020, 0.251]$	$[-0.013, 0.233]$	$[-0.010, 0.224]$
			p-value	0.091	0.084	0.053	7.60E-02
rs1342586	TGFB2	69745	β	0.161	0.145	0.214	0.181
			$[95\%CI]$	$[-0.008, 0.318]$	$[-0.004, 0.283]$	$[0.080, 0.330]$	$[0.088, 0.303]$
			p-value	0.054	0.044	0.000817	0.000967
rs3814115	PCSK5	73537	β	0.054	0.073	0.073	0.066
			$[95\%CI]$	$[-0.101, 0.182]$	$[-0.036, 0.189]$	$[-0.023, 0.180]$	$[-0.030, 0.163]$
			p-value	0.451	0.206	0.157	0.184
rs1780616	LBP	73560	β	0.113	0.098	0.124	0.124
			$[95\%CI]$	$[-0.013, 0.273]$	$[-0.011, 0.221]$	$[0.023, 0.236]$	$[0.034, 0.226]$
			p-value	1.19E-01	9.50E-02	2.30E-02	0.012
rs3736228	LRP5	73508	β	0.187	0.181	0.208	0.188
			$[95\%CI]$	$[0.009, 0.422]$	$[0.024, 0.345]$	$[0.071, 0.346]$	$[0.089, 0.328]$
			p-value	0.073	0.028	0.00301	0.00216
rs212517	ECE1	71344	β	0.139	0.092	0.12	0.091
			$[95\%CI]$	$[0.001, 0.274]$	$[-0.009, 0.210]$	$[0.017, 0.220]$	$[0.003, 0.187]$
			p-value	0.043	0.098	0.019	5.20E-02
rs7359336	NFAT5	73094	β	0.223	0.225	0.133	0.112
			$[95\%CI]$	$[0.080, 0.353]$	$[0.108, 0.326]$	$[0.037, 0.234]$	$[0.013, 0.190]$
			p-value	0.00123	0.0000508	0.00865	0.013
rs2682552	XRCC1	73106	β	0.036	0.059	0.117	0.125
			$[95\%CI]$	$[-0.104, 0.218]$	$[-0.062, 0.223]$	$[-0.028, 0.236]$	$[0.022, 0.231]$
			p-value	0.654	4.22E-01	0.085	0.018

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		5%	10%	15%	20%
rs17085675	PCSK1	69775	β	0.243	0.233	0.216	0.139
			$[95\%CI]$	$[0.071, 0.383]$	$[0.116, 0.356]$	$[0.098, 0.311]$	$[0.047, 0.242]$
			p-value	0.00217	0.00013	6.03E-05	5.20E-03
rs11102986	RXRA	71224	β	0.243	0.211	0.219	0.205
			$[95\%CI]$	$[0.074, 0.423]$	$[0.064, 0.344]$	$[0.083, 0.357]$	$[0.099, 0.326]$
			p-value	0.00631	2.81E-03	1.59E-03	3.67E-04
rs12603813	PLCD3	71116	β	0.143	0.146	0.144	0.088
			$[95\%CI]$	$[-0.014, 0.274]$	$[0.026, 0.290]$	$[0.008, 0.247]$	$[0.002, 0.194]$
			p-value	0.051	3.10E-02	1.90E-02	7.50E-02
rs6219	IGF1	73562	β	0	0.051	0.05	0.05
			$[95\%CI]$	$[-0.194, 0.206]$	$[-0.134, 0.230]$	$[-0.118, 0.222]$	$[-0.088, 0.189]$
			p-value	1	0.572	5.62E-01	4.84E-01
rs2234693	ESR1	69772	β	0.193	0.164	0.167	0.139
			$[95\%CI]$	$[0.043, 0.334]$	$[0.048, 0.284]$	$[0.064, 0.272]$	$[0.060, 0.238]$
			p-value	0.00899	0.00634	0.00163	0.00202
rs46522	UBE2Z	64204	β	0.219	0.161	0.152	0.127
			$[95\%CI]$	$[0.088, 0.365]$	$[0.051, 0.269]$	$[0.048, 0.256]$	$[0.030, 0.226]$
			p-value	0.002	0.00335	0.0044	0.01
rs891088	INSR	73562	β	0.084	0.071	0.1	0.093
			$[95\%CI]$	$[-0.056, 0.251]$	$[-0.047, 0.196]$	$[-0.025, 0.207]$	$[-0.019, 0.193]$
			p-value	0.283	2.54E-01	8.90E-02	0.087
rs9857730	VILL	73358	β	0.257	0.227	0.198	0.153
			$[95\%CI]$	$[0.098, 0.384]$	$[0.084, 0.348]$	$[0.078, 0.318]$	$[0.049, 0.261]$
			p-value	0.000436	0.000902	0.00136	4.87E-03
rs10185680	MFSD2B	73232	β	0.000	0.068	0.182	0.138
			$[95\%CI]$	$[-0.121, 0.164]$	$[-0.030, 0.180]$	$[0.076, 0.278]$	$[0.063, 0.243]$
			p-value	1.000	0.207	3.73E-04	2.72E-03
rs7163907	PTPN9	73234	β	-0.192	-0.236	-0.15	-0.13
			$[95\%CI]$	$[-0.325, -0.038]$	$[-0.362, -0.117]$	$[-0.280, -0.040]$	$[-0.239, -0.039]$
			p-value	0.00825	0.000147	0.014	0.011
rs2176167	NOP58	73508	β	0.173	0.105	0.109	0.090
			$[95\%CI]$	$[0.028, 0.284]$	$[-0.002, 0.217]$	$[-0.003, 0.205]$	$[-0.004, 0.189]$
			p-value	0.008	0.064	0.041	0.068
rs7557989	THADA	64417	β	0.228	0.186	0.152	0.189
			$[95\%CI]$	$[0.094, 0.364]$	$[0.058, 0.303]$	$[0.041, 0.264]$	$[0.084, 0.285]$
			p-value	0.001	0.003	0.007	2.14E-04

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		5%	10%	15%	20%
rs510769	OPRM1	64413	β	0.073	0.139	0.093	0.107
			$[95\%CI]$	$[-0.119, 0.234]$	$[-0.005, 0.262]$	$[-0.032, 0.226]$	$[-0.003, 0.218]$
			p-value	0.413	0.041	0.152	0.053
rs7756224	NMBR	60624	β	0.071	0.151	0.187	0.128
			$[95\%CI]$	$[-0.078, 0.214]$	$[0.008, 0.255]$	$[0.070, 0.290]$	$[0.046, 0.245]$
			p-value	0.337	0.016	0.001	0.012
rs2282537	POU2F3	64419	β	0.155	0.148	0.062	0.074
			$[95\%CI]$	$[-0.055, 0.373]$	$[-0.020, 0.317]$	$[-0.079, 0.225]$	$[-0.045, 0.229]$
			p-value	0.154	0.091	0.416	0.288
rs7853859	CENPP	69662	β	0.145	0.187	0.095	0.095
			$[95\%CI]$	$[-0.011, 0.282]$	$[0.067, 0.296]$	$[-0.016, 0.209]$	$[0.005, 0.184]$
			p-value	0.050	0.001	0.095	0.037
rs2229642	ITPR3	52677	β	0.051	0.055	0.112	0.110
			$[95\%CI]$	$[-0.091, 0.208]$	$[-0.082, 0.157]$	$[-0.007, 0.240]$	$[0.000, 0.226]$
			p-value	0.498	0.367	0.076	0.055
rs5015437	LMF1	64185	β	0.076	0.100	0.066	0.083
			$[95\%CI]$	$[-0.093, 0.209]$	$[-0.035, 0.202]$	$[-0.043, 0.186]$	$[0.000, 0.183]$
			p-value	0.317	0.098	0.262	0.076
rs7963565	KNTC1	73568	β	-0.015	0.046	0.048	0.080
			$[95\%CI]$	$[-0.176, 0.121]$	$[-0.066, 0.162]$	$[-0.056, 0.163]$	$[-0.018, 0.167]$
			p-value	0.839	0.427	0.392	0.093
rs696	NFKBIA	73461	β	0.036	0.108	0.109	0.063
			$[95\%CI]$	$[-0.102, 0.179]$	$[-0.013, 0.210]$	$[-0.005, 0.208]$	$[-0.023, 0.157]$
			p-value	0.614	0.058	0.044	0.180
rs25656	NFATC1	48455	β	0.160	0.158	0.171	0.161
			$[95\%CI]$	$[0.003, 0.344]$	$[0.033, 0.314]$	$[0.051, 0.293]$	$[0.048, 0.296]$
			p-value	0.065	0.027	0.006	0.011
rs3100776	IHH	64334	β	-0.101	0.152	0.113	0.117
			$[95\%CI]$	$[-0.589, 0.266]$	$[-0.156, 0.411]$	$[-0.131, 0.432]$	$[-0.111, 0.367]$
			p-value	0.639	0.298	0.436	0.343
rs12145922	PKN2	69346	β	0.000	0.075	0.090	0.063
			$[95\%CI]$	$[-0.140, 0.123]$	$[-0.054, 0.165]$	$[-0.022, 0.193]$	$[-0.025, 0.154]$
			p-value	1.000	0.180	0.102	0.166
rs526134	USP37	60625	β	-0.085	-0.035	0.058	0.057
			$[95\%CI]$	$[-0.243, 0.053]$	$[-0.151, 0.085]$	$[-0.066, 0.157]$	$[-0.031, 0.149]$
			p-value	0.260	0.563	0.305	0.213

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		5%	10%	15%	20%
rs7004280	RPS20	64415	β	0.076	0.100	0.009	0.100
			$[95\%CI]$	$[-0.227, 0.495]$	$[-0.209, 0.351]$	$[-0.244, 0.338]$	$[-0.206, 0.310]$
			p-value	0.686	0.476	0.948	0.446
rs4808199	GATAD2A	64415	β	0.038	-0.094	0.006	0.102
			$[95\%CI]$	$[-0.140, 0.215]$	$[-0.231, 0.065]$	$[-0.134, 0.140]$	$[-0.028, 0.210]$
			p-value	0.677	0.206	0.937	0.09
rs3821009	PDE11A	69745	β	-0.073	-0.048	-0.015	0.014
			$[95\%CI]$	$[-0.281, 0.144]$	$[-0.299, 0.185]$	$[-0.206, 0.166]$	$[-0.139, 0.203]$
			p-value	0.492	0.701	0.872	0.874
rs2291617	METTL1	73099	β	-0.027	0.000	0.075	0.072
			$[95\%CI]$	$[-0.166, 0.116]$	$[-0.107, 0.134]$	$[-0.028, 0.179]$	$[-0.023, 0.165]$
			p-value	0.708	1.000	0.156	0.139
rs1051168	SEC11A	64399	β	0.054	-0.020	0.062	0.065
			$[95\%CI]$	$[-0.102, 0.186]$	$[-0.133, 0.130]$	$[-0.065, 0.178]$	$[-0.032, 0.172]$
			p-value	0.458	0.763	0.322	0.209
rs803932	ASTN2	64105	β	0.000	0.129	0.122	0.130
			$[95\%CI]$	$[-0.163, 0.152]$	$[0.007, 0.251]$	$[0.019, 0.249]$	$[0.026, 0.222]$
			p-value	1.000	0.038	0.036	0.00816
rs2247870	GPR98	69775	β	0.000	0.111	0.125	0.113
			$[95\%CI]$	$[-0.120, 0.163]$	$[-0.004, 0.226]$	$[0.022, 0.235]$	$[0.015, 0.200]$
			p-value	1.000	0.057	0.022	0.016
rs12503378	NUDT6	73569	β	0.047	0.03	0.094	0.088
			$[95\%CI]$	$[-0.116, 0.242]$	$[-0.119, 0.174]$	$[-0.034, 0.228]$	$[-0.031, 0.210]$
			p-value	0.610	0.687	0.162	0.154
rs2145923	NPR2	69777	β	0.061	0.052	0.052	0.070
			$[95\%CI]$	$[-0.133, 0.197]$	$[-0.116, 0.184]$	$[-0.082, 0.188]$	$[-0.054, 0.174]$
			p-value	0.469	0.498	0.455	2.33E-01
rs12603582	ITGB3	73541	β	0.200	0.176	0.121	0.071
			$[95\%CI]$	$[0.064, 0.394]$	$[0.061, 0.322]$	$[-0.008, 0.247]$	$[-0.032, 0.186]$
			p-value	0.017	0.00801	0.06	0.207
rs4864548	CLOCK	73568	β	-0.039	0.029	0.009	0.000
			$[95\%CI]$	$[-0.161, 0.106]$	$[-0.070, 0.157]$	$[-0.092, 0.117]$	$[-0.099, 0.077]$
			p-value	0.559	0.618	0.866	1.000
rs2735469	MRPL23	67563	β	0.080	-0.039	-0.020	0.043
			$[95\%CI]$	$[-0.133, 0.287]$	$[-0.173, 0.149]$	$[-0.154, 0.137]$	$[-0.095, 0.149]$
			p-value	0.462	0.635	0.785	0.483

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		5%	10%	15%	20%
rs1051431	MPHOSPH9	73561	β	-0.016	-0.029	-0.021	0.000
			$[95\%CI]$	$[-0.182, 0.117]$	$[-0.167, 0.103]$	$[-0.139, 0.114]$	$[-0.100, 0.104]$
			p-value	0.835	0.675	0.742	1.000
rs41132	AP3B1	73441	β	-0.073	-0.004	0.02	0.025
			$[95\%CI]$	$[-0.226, 0.085]$	$[-0.126, 0.129]$	$[-0.091, 0.137]$	$[-0.070, 0.135]$
			p-value	0.353	0.946	0.734	0.631
rs6127698	MC3R	73486	β	-0.035	-0.004	-0.002	0.025
			$[95\%CI]$	$[-0.176, 0.084]$	$[-0.105, 0.107]$	$[-0.102, 0.098]$	$[-0.057, 0.114]$
			p-value	0.593	9.39E-01	0.976	0.569
rs1481892	ARNTL	73543	β	0.000	0.000	0.041	0.046
			$[95\%CI]$	$[-0.155, 0.125]$	$[-0.130, 0.113]$	$[-0.075, 0.153]$	$[-0.044, 0.140]$
			p-value	1.000	1.000	0.484	0.328
rs2633442	MKRN2	60624	β	0.154	0.080	0.000	-0.051
			$[95\%CI]$	$[0.021, 0.318]$	$[-0.054, 0.200]$	$[-0.118, 0.096]$	$[-0.149, 0.042]$
			p-value	0.043	0.222	1.000	0.293
rs1535	FADS1	73553	β	-0.008	-0.016	0.030	0.000
			$[95\%CI]$	$[-0.156, 0.124]$	$[-0.133, 0.096]$	$[-0.079, 0.141]$	$[-0.092, 0.103]$
			p-value	0.915	0.79	0.592	1.000
rs10861148	HSP90B1	73557	β	0.087	-0.082	-0.017	-0.025
			$[95\%CI]$	$[-0.117, 0.274]$	$[-0.241, 0.083]$	$[-0.180, 0.137]$	$[-0.144, 0.123]$
			p-value	0.389	0.321	0.831	0.711
GS-Height		73570	β	0.173	0.172	0.176	0.18
			$[95\%CI]$	$[0.159, 0.185]$	$[0.161, 0.182]$	$[0.166, 0.185]$	$[0.170, 0.188]$
			p-value	6.02E-146	4.50E-227	1.40E-294	1.28e-322
			VarianceExplained	1.817%	1.825%	1.971%	1.996%

Part 2 - 25% to 40%

SNP	Gene	N		25%	30%	35%	40%
rs1042725	HMGA2	73105	β	0.613	0.579	0.555	0.53
			$[95\%CI]$	$[0.514, 0.693]$	$[0.496, 0.654]$	$[0.470, 0.636]$	$[0.447, 0.616]$
			p-value	9.28E-41	7.77E-46	5.51E-39	3.29E-35
rs2853977	HCP5	44699	β	0.688	0.65	0.7	0.674
			$[95\%CI]$	$[0.574, 0.784]$	$[0.563, 0.764]$	$[0.582, 0.792]$	$[0.577, 0.778]$
			p-value	3.15E-37	1.58E-35	8.74E-39	2.09E-39
rs3782415	SOCS2	73568	β	0.466	0.456	0.436	0.43
			$[95\%CI]$	$[0.355, 0.574]$	$[0.358, 0.558]$	$[0.354, 0.545]$	$[0.320, 0.536]$
			p-value	3.67E-17	2.05E-19	3.30E-19	1.21E-14
rs780094	GCKR	73548	β	0.403	0.409	0.409	0.39
			$[95\%CI]$	$[0.314, 0.488]$	$[0.324, 0.490]$	$[0.323, 0.500]$	$[0.312, 0.481]$
			p-value	2.59E-19	5.05E-22	9.24E-20	3.20E-20
rs9892365	TBX2	73566	β	0.328	0.356	0.351	0.365
			$[95\%CI]$	$[0.232, 0.421]$	$[0.251, 0.440]$	$[0.256, 0.435]$	$[0.269, 0.458]$
			p-value	9.81E-12	9.25E-14	9.98E-15	1.59E-14
rs7137534	PDE3A	73567	β	0.35	0.353	0.363	0.371
			$[95\%CI]$	$[0.258, 0.446]$	$[0.258, 0.439]$	$[0.264, 0.437]$	$[0.279, 0.459]$
			p-value	2.33E-13	2.76E-14	2.81E-16	3.33E-16
rs1776897	HMGA1	64379	β	0.616	0.591	0.606	0.617
			$[95\%CI]$	$[0.451, 0.751]$	$[0.444, 0.723]$	$[0.449, 0.768]$	$[0.457, 0.774]$
			p-value	8.22E-16	2.64E-17	1.14E-13	1.77E-14
rs572169	GHSR	73554	β	0.36	0.338	0.347	0.353
			$[95\%CI]$	$[0.256, 0.444]$	$[0.260, 0.432]$	$[0.256, 0.431]$	$[0.267, 0.438]$
			p-value	7.66E-14	2.11E-14	8.66E-15	4.15E-16
rs2679178	NPPC	73567	β	0.499	0.481	0.514	0.514
			$[95\%CI]$	$[0.332, 0.662]$	$[0.324, 0.614]$	$[0.398, 0.669]$	$[0.370, 0.657]$
			p-value	2.14E-09	8.45E-11	1.12E-13	1.8E-12
rs2053156	GRB2	73536	β	0.342	0.345	0.359	0.37
			$[95\%CI]$	$[0.213, 0.443]$	$[0.234, 0.438]$	$[0.249, 0.457]$	$[0.273, 0.474]$
			p-value	8.13E-09	2.02E-11	1.36E-11	1.72E-13
rs9930741	ERI2	73143	β	0.312	0.3	0.3	0.301
			$[95\%CI]$	$[0.225, 0.400]$	$[0.215, 0.383]$	$[0.212, 0.380]$	$[0.215, 0.381]$
			p-value	5.41E-12	3.51E-12	2.42E-12	6.42E-13
rs2854207	CSH2	67567	β	0.225	0.241	0.207	0.248
			$[95\%CI]$	$[0.124, 0.326]$	$[0.138, 0.319]$	$[0.132, 0.315]$	$[0.139, 0.340]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		25%	30%	35%	40%
			p-value	1.13E-05	1.75E-07	8.46E-06	1.31E-06
rs4320932	IGF2	71204	β	0.352	0.35	0.372	0.363
			$[95\%CI]$	$[0.247, 0.469]$	$[0.242, 0.457]$	$[0.264, 0.457]$	$[0.260, 0.469]$
			p-value	5.70E-10	1.23E-10	2.98E-14	7.03E-12
rs752313	EZH1	69774	β	0.251	0.232	0.242	0.24
			$[95\%CI]$	$[0.165, 0.340]$	$[0.135, 0.302]$	$[0.159, 0.329]$	$[0.154, 0.323]$
			p-value	1.72E-08	4.45E-08	2.27E-08	2.41E-08
rs709939	SAMD4A	73569	β	0.213	0.252	0.222	0.219
			$[95\%CI]$	$[0.116, 0.289]$	$[0.153, 0.321]$	$[0.129, 0.298]$	$[0.137, 0.292]$
			p-value	1.46E-06	3.76E-09	2.89E-07	3.62E-08
rs1036477	FBN1	73539	β	0.437	0.398	0.453	0.456
			$[95\%CI]$	$[0.292, 0.574]$	$[0.265, 0.550]$	$[0.313, 0.579]$	$[0.333, 0.582]$
			p-value	1.02E-09	4.74E-08	2.73E-11	6.97E-13
rs158676	CDK5RAP1	73562	β	0.247	0.243	0.253	0.249
			$[95\%CI]$	$[0.149, 0.343]$	$[0.158, 0.342]$	$[0.166, 0.344]$	$[0.164, 0.344]$
			p-value	4.15E-07	2.61E-07	2.31E-08	6.04E-08
rs1822469	PPP3R1	69710	β	0.31	0.306	0.311	0.29
			$[95\%CI]$	$[0.226, 0.406]$	$[0.207, 0.378]$	$[0.222, 0.394]$	$[0.208, 0.369]$
			p-value	1.60E-11	3.10E-12	1.99E-12	1.71E-12
rs258281	RAB26	67411	β	0.315	0.305	0.372	0.321
			$[95\%CI]$	$[0.197, 0.447]$	$[0.205, 0.434]$	$[0.242, 0.459]$	$[0.203, 0.438]$
			p-value	7.30E-07	2.51E-07	2.00E-11	7.64E-08
rs9366637	HFE	73557	β	0.497	0.447	0.5	0.5
			$[95\%CI]$	$[0.305, 0.691]$	$[0.275, 0.608]$	$[0.331, 0.662]$	$[0.336, 0.686]$
			p-value	6.00E-07	1.18E-07	3.34E-09	2.69E-08
rs551219	COL24A1	73076	β	0.3	0.3	0.268	0.233
			$[95\%CI]$	$[0.217, 0.401]$	$[0.224, 0.398]$	$[0.192, 0.373]$	$[0.156, 0.333]$
			p-value	1.07E-10	1.21E-11	6.78E-09	2.01E-07
rs13072536	ITIH4	73519	β	0.295	0.28	0.256	0.22
			$[95\%CI]$	$[0.205, 0.420]$	$[0.186, 0.379]$	$[0.165, 0.356]$	$[0.123, 0.317]$
			p-value	1.05E-07	1.71E-08	1.02E-07	8.88E-06
rs7522692	PIGC	64398	β	0.255	0.243	0.232	0.266
			$[95\%CI]$	$[0.140, 0.354]$	$[0.137, 0.347]$	$[0.128, 0.352]$	$[0.173, 0.372]$
			p-value	2.81E-06	6.06E-06	4.79E-05	2.23E-07
rs13076290	CTNNB1	73154	β	0.237	0.24	0.236	0.212
			$[95\%CI]$	$[0.149, 0.331]$	$[0.152, 0.325]$	$[0.140, 0.310]$	$[0.129, 0.290]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		25%	30%	35%	40%
			p-value	2.66E-07	3.73E-08	4.91E-08	2.86E-07
rs1636255	GNA12	64371	β	0.255	0.216	0.229	0.217
			$[95\%CI]$	$[0.146, 0.361]$	$[0.117, 0.313]$	$[0.135, 0.319]$	$[0.123, 0.315]$
			p-value	3.05E-06	1.40E-05	8.81E-07	7.72E-06
rs3796529	REST	73119	β	0.278	0.246	0.265	0.297
			$[95\%CI]$	$[0.171, 0.378]$	$[0.149, 0.357]$	$[0.169, 0.375]$	$[0.187, 0.393]$
			p-value	1.23E-07	3.54E-06	3.27E-07	2.03E-08
rs1866146	POMC	73516	β	0.161	0.181	0.183	0.231
			$[95\%CI]$	$[0.074, 0.259]$	$[0.087, 0.277]$	$[0.087, 0.277]$	$[0.147, 0.309]$
			p-value	5.58E-04	1.69E-04	1.46E-04	1.61E-08
rs9844666	PCCB	73511	β	0.325	0.269	0.298	0.29
			$[95\%CI]$	$[0.218, 0.431]$	$[0.192, 0.398]$	$[0.205, 0.396]$	$[0.187, 0.384]$
			p-value	2.46E-09	2.90E-07	6.69E-10	5.57E-09
rs8071847	POLR2A	73565	β	0.301	0.275	0.274	0.29
			$[95\%CI]$	$[0.192, 0.417]$	$[0.174, 0.372]$	$[0.181, 0.378]$	$[0.186, 0.386]$
			p-value	1.81E-07	5.49E-08	4.35E-08	9.53E-09
rs3783937	FBLN5	73104	β	0.25	0.247	0.26	0.28
			$[95\%CI]$	$[0.133, 0.343]$	$[0.146, 0.342]$	$[0.164, 0.357]$	$[0.178, 0.368]$
			p-value	3.39E-06	7.46E-07	9.34E-08	6.96E-09
rs11080149	NF1	73567	β	0.25	0.278	0.308	0.29
			$[95\%CI]$	$[0.145, 0.387]$	$[0.157, 0.406]$	$[0.167, 0.430]$	$[0.172, 0.401]$
			p-value	4.09E-05	1.24E-05	4.08E-06	6.01E-07
rs17472113	ZAR1	58807	β	0.301	0.339	0.323	0.293
			$[95\%CI]$	$[0.195, 0.419]$	$[0.233, 0.444]$	$[0.208, 0.411]$	$[0.192, 0.395]$
			p-value	1.49E-07	3.37E-10	5.53E-10	1.42E-08
rs490634	CISH	61124	β	0.36	0.404	0.385	0.335
			$[95\%CI]$	$[0.217, 0.489]$	$[0.241, 0.523]$	$[0.250, 0.522]$	$[0.205, 0.478]$
			p-value	0.000000279	1.96E-08	1.50E-08	1.42E-06
rs17622208	SLC22A5	73531	β	0.215	0.191	0.206	0.195
			$[95\%CI]$	$[0.127, 0.299]$	$[0.103, 0.276]$	$[0.122, 0.277]$	$[0.123, 0.281]$
			p-value	1.07E-06	1.53E-05	2.68E-07	1.39E-06
rs2982712	ESR1	73566	β	0.187	0.199	0.233	0.168
			$[95\%CI]$	$[0.096, 0.275]$	$[0.122, 0.293]$	$[0.132, 0.296]$	$[0.103, 0.259]$
			p-value	3.66E-05	4.92E-06	2.40E-08	3.37E-05
rs1950500	NFATC4	64408	β	0.205	0.197	0.206	0.204
			$[95\%CI]$	$[0.104, 0.316]$	$[0.100, 0.301]$	$[0.123, 0.312]$	$[0.111, 0.301]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		25%	30%	35%	40%
			p-value	1.12E-04	1.14E-04	1.56E-05	2.12E-05
rs1476387	PPIL6	73563	β	0.265	0.25	0.237	0.226
			$[95\%CI]$	$[0.185, 0.364]$	$[0.150, 0.327]$	$[0.145, 0.314]$	$[0.145, 0.297]$
			p-value	3.50E-09	3.20E-08	2.89E-08	5.56E-09
rs4946932	FOXO3	73565	β	0.203	0.21	0.198	0.18
			$[95\%CI]$	$[0.128, 0.312]$	$[0.117, 0.304]$	$[0.116, 0.284]$	$[0.093, 0.278]$
			p-value	1.47E-05	9.92E-06	4.91E-06	1.58E-04
rs1800783	NOS3	71277	β	0.189	0.184	0.191	0.221
			$[95\%CI]$	$[0.100, 0.286]$	$[0.107, 0.273]$	$[0.108, 0.279]$	$[0.135, 0.309]$
			p-value	6.42E-05	1.76E-05	1.39E-05	5.65E-07
rs6718902	STAT1	73557	β	0.293	0.245	0.227	0.221
			$[95\%CI]$	$[0.193, 0.386]$	$[0.139, 0.337]$	$[0.124, 0.317]$	$[0.135, 0.314]$
			p-value	1.77E-09	1.61E-06	4.09E-06	1.38E-06
rs2425019	MMP24	69770	β	0.216	0.211	0.193	0.213
			$[95\%CI]$	$[0.129, 0.297]$	$[0.128, 0.298]$	$[0.108, 0.277]$	$[0.129, 0.295]$
			p-value	4.75E-07	1.08E-06	1.01E-05	5.17E-07
rs6731022	EIF2AK3	69767	β	0.204	0.2	0.199	0.205
			$[95\%CI]$	$[0.107, 0.297]$	$[0.115, 0.300]$	$[0.113, 0.291]$	$[0.120, 0.295]$
			p-value	1.86E-05	2.52E-05	1.40E-05	4.15E-06
rs12940055	MAP3K3	73566	β	0.312	0.267	0.275	0.299
			$[95\%CI]$	$[0.165, 0.457]$	$[0.126, 0.420]$	$[0.151, 0.398]$	$[0.162, 0.434]$
			p-value	2.77E-05	3.64E-04	1.27E-05	1.63E-05
rs864745	JAZF1	60616	β	0.176	0.185	0.201	0.183
			$[95\%CI]$	$[0.077, 0.264]$	$[0.096, 0.279]$	$[0.111, 0.289]$	$[0.097, 0.285]$
			p-value	2.28E-04	6.35E-05	9.58E-06	1.16E-04
rs6487088	PDE3A	73563	β	0.251	0.2	0.238	0.233
			$[95\%CI]$	$[0.158, 0.371]$	$[0.095, 0.312]$	$[0.105, 0.311]$	$[0.132, 0.327]$
			p-value	3.77E-06	0.000308	6.26E-06	2.64E-06
rs4973410	NCL	64408	β	0.236	0.209	0.223	0.201
			$[95\%CI]$	$[0.139, 0.325]$	$[0.118, 0.300]$	$[0.132, 0.308]$	$[0.113, 0.293]$
			p-value	5.87E-07	5.43E-06	7.80E-07	9.51E-06
rs451061	PRKCZ	71365	β	0.184	0.204	0.19	0.186
			$[95\%CI]$	$[0.093, 0.277]$	$[0.120, 0.281]$	$[0.107, 0.270]$	$[0.107, 0.273]$
			p-value	9.14E-05	6.22E-07	4.12E-06	1.35E-05
rs832575	MAP3K1	73539	β	0.312	0.248	0.267	0.295
			$[95\%CI]$	$[0.176, 0.439]$	$[0.122, 0.370]$	$[0.147, 0.405]$	$[0.169, 0.414]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		25%	30%	35%	40%
			p-value	2.63E-06	8.66E-05	6.16E-05	2.86E-06
rs4955526	EPHB1	64417	β	0.23	0.232	0.226	0.234
			$[95\%CI]$	$[0.134, 0.327]$	$[0.141, 0.330]$	$[0.139, 0.315]$	$[0.145, 0.320]$
			p-value	0.00000265	1.51E-06	4.16E-07	1.07E-07
rs8038415	IGF1R	73516	β	0.213	0.246	0.188	0.183
			$[95\%CI]$	$[0.122, 0.305]$	$[0.151, 0.312]$	$[0.109, 0.274]$	$[0.108, 0.269]$
			p-value	5.35E-06	1.40E-09	8.47E-06	8.83E-06
rs7578199	HDLBP	73569	β	0.25	0.227	0.192	0.162
			$[95\%CI]$	$[0.140, 0.351]$	$[0.116, 0.315]$	$[0.091, 0.286]$	$[0.071, 0.258]$
			p-value	3.18E-06	9.80E-06	9.79E-05	0.000676
rs7020782	PAPPA	73566	β	0.175	0.191	0.19	0.176
			$[95\%CI]$	$[0.089, 0.277]$	$[0.107, 0.287]$	$[0.104, 0.282]$	$[0.095, 0.258]$
			p-value	2.67E-04	2.51E-05	2.51E-05	2.10E-05
rs2229712	RPS6KA1	48240	β	0.305	0.26	0.279	0.293
			$[95\%CI]$	$[0.181, 0.431]$	$[0.148, 0.390]$	$[0.160, 0.403]$	$[0.169, 0.404]$
			p-value	1.63E-06	2.64E-05	8.25E-06	1.05E-06
rs7572476	BOK	71367	β	0.158	0.127	0.127	0.129
			$[95\%CI]$	$[0.078, 0.255]$	$[0.039, 0.224]$	$[0.052, 0.214]$	$[0.036, 0.203]$
			p-value	5.67E-04	6.88E-03	2.32E-03	2.41E-03
rs2066807	PAN2	71822	β	0.142	0.214	0.194	0.196
			$[95\%CI]$	$[0.041, 0.277]$	$[0.094, 0.313]$	$[0.093, 0.297]$	$[0.091, 0.307]$
			p-value	1.90E-02	1.34E-04	1.76E-04	4.60E-04
rs1516796	ACAN	69308	β	0.125	0.128	0.128	0.14
			$[95\%CI]$	$[0.049, 0.225]$	$[0.051, 0.217]$	$[0.052, 0.224]$	$[0.055, 0.232]$
			p-value	5.12E-03	2.50E-03	3.27E-03	1.74E-03
rs6180	GHR	73552	β	0.175	0.211	0.229	0.23
			$[95\%CI]$	$[0.095, 0.275]$	$[0.119, 0.287]$	$[0.143, 0.304]$	$[0.145, 0.306]$
			p-value	1.21E-04	7.65E-07	2.73E-08	1.92E-08
rs8055190	LRRC36	64416	β	0.455	0.44	0.488	0.447
			$[95\%CI]$	$[0.204, 0.678]$	$[0.237, 0.671]$	$[0.291, 0.681]$	$[0.246, 0.681]$
			p-value	1.90E-04	9.10E-05	8.96E-07	6.81E-05
rs17106235	FAF1	60489	β	0.284	0.367	0.367	0.317
			$[95\%CI]$	$[0.117, 0.436]$	$[0.188, 0.498]$	$[0.215, 0.501]$	$[0.164, 0.472]$
			p-value	5.28E-04	5.37E-06	3.43E-07	6.29E-05
rs3739707	LPAR1	73568	β	0.21	0.22	0.195	0.218
			$[95\%CI]$	$[0.092, 0.307]$	$[0.128, 0.315]$	$[0.099, 0.279]$	$[0.126, 0.307]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		25%	30%	35%	40%
			p-value	1.55E-04	3.81E-06	1.74E-05	2.63E-06
rs674424	ABCG4	73568	β	0.25	0.249	0.254	0.27
			$[95\%CI]$	$[0.135, 0.335]$	$[0.149, 0.355]$	$[0.161, 0.354]$	$[0.180, 0.364]$
			p-value	1.04E-06	1.25E-06	2.78E-07	7.15E-09
rs12225387	NEU3	60621	β	0.229	0.2	0.187	0.144
			$[95\%CI]$	$[0.110, 0.329]$	$[0.089, 0.299]$	$[0.081, 0.289]$	$[0.052, 0.256]$
			p-value	4.25E-05	1.78E-04	4.40E-04	5.43E-03
rs3812265	CNOT4	69776	β	0.15	0.156	0.173	0.175
			$[95\%CI]$	$[0.050, 0.254]$	$[0.067, 0.263]$	$[0.072, 0.275]$	$[0.074, 0.266]$
			p-value	3.68E-03	1.88E-03	8.46E-04	3.84E-04
rs10208728	IHH	64409	β	0.254	0.263	0.271	0.262
			$[95\%CI]$	$[0.126, 0.439]$	$[0.111, 0.425]$	$[0.110, 0.415]$	$[0.119, 0.403]$
			p-value	1.42E-03	1.14E-03	5.23E-04	3.02E-04
rs291979	GRK5	73568	β		0.157	0.235	0.222
			$[95\%CI]$		$[0.076, 0.281]$	$[0.121, 0.325]$	$[0.132, 0.314]$
			p-value		3.22E-03	4.00E-06	1.36E-06
rs2715553	RARA	69206	β	0.192	0.165	0.177	0.182
			$[95\%CI]$	$[0.098, 0.277]$	$[0.083, 0.260]$	$[0.090, 0.262]$	$[0.105, 0.264]$
			p-value	2.54E-05	2.85E-04	7.19E-05	8.07E-06
rs2057291	GNAS	70100	β	0.174	0.193	0.208	0.193
			$[95\%CI]$	$[0.084, 0.280]$	$[0.100, 0.278]$	$[0.113, 0.286]$	$[0.105, 0.279]$
			p-value	0.000513	2.09E-05	1.95E-06	1.28E-05
rs4803520	GRIK5	69747	β	0.256	0.211	0.176	0.223
			$[95\%CI]$	$[0.101, 0.424]$	$[0.077, 0.349]$	$[0.055, 0.330]$	$[0.102, 0.362]$
			p-value	2.02E-03	2.32E-03	1.30E-02	8.10E-04
rs10736682	APLNR	73090	β	0.183	0.181	0.14	0.155
			$[95\%CI]$	$[0.091, 0.260]$	$[0.089, 0.262]$	$[0.082, 0.240]$	$[0.072, 0.234]$
			p-value	0.0000183	0.0000517	5.53E-04	1.56E-04
rs2909430	TP53	73412	β	0.249	0.198	0.242	0.233
			$[95\%CI]$	$[0.096, 0.367]$	$[0.070, 0.331]$	$[0.112, 0.346]$	$[0.119, 0.369]$
			p-value	3.34E-04	3.19E-03	4.71E-05	0.000258
rs12050767	CYP19A1	73563	β	0.141	0.113	0.111	0.123
			$[95\%CI]$	$[0.052, 0.227]$	$[0.024, 0.188]$	$[0.014, 0.181]$	$[0.036, 0.199]$
			p-value	0.0015	0.00648	8.53E-03	0.00277
rs602633	PSRC1	64375	β	0.171	0.181	0.166	0.134
			$[95\%CI]$	$[0.059, 0.283]$	$[0.074, 0.278]$	$[0.073, 0.279]$	$[0.034, 0.241]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		25%	30%	35%	40%
			p-value	2.63E-03	4.62E-04	1.28E-03	1.10E-02
rs1738475	HTR1D	64411	β	0.16	0.143	0.134	0.123
			$[95\%CI]$	$[0.068, 0.262]$	$[0.060, 0.234]$	$[0.044, 0.230]$	$[0.037, 0.210]$
			p-value	1.32E-03	1.37E-03	4.59E-03	5.26E-03
rs17754	RFC1	73558	β	0.134	0.128	0.142	0.15
			$[95\%CI]$	$[0.052, 0.233]$	$[0.049, 0.219]$	$[0.068, 0.234]$	$[0.071, 0.233]$
			p-value	0.00359	0.00297	0.000822	0.000288
rs17541471	NPR3	69777	β	0.127	0.14	0.175	0.187
			$[95\%CI]$	$[0.026, 0.254]$	$[0.034, 0.256]$	$[0.061, 0.282]$	$[0.090, 0.296]$
			p-value	2.90E-02	1.30E-02	0.00188	0.000371
rs1342586	TGFB2	69745	β	0.158	0.124	0.147	0.15
			$[95\%CI]$	$[0.059, 0.272]$	$[0.029, 0.228]$	$[0.040, 0.234]$	$[0.054, 0.248]$
			p-value	4.01E-03	0.016	0.00308	0.00238
rs3814115	PCSK5	73537	β	0.062	0.091	0.123	0.12
			$[95\%CI]$	$[-0.032, 0.152]$	$[0.002, 0.185]$	$[0.027, 0.204]$	$[0.033, 0.214]$
			p-value	0.181	0.052	0.00646	0.00883
rs1780616	LBP	73560	β	0.136	0.147	0.16	0.167
			$[95\%CI]$	$[0.053, 0.238]$	$[0.064, 0.245]$	$[0.076, 0.247]$	$[0.087, 0.251]$
			p-value	0.00356	0.00123	0.000207	0.0000647
rs3736228	LRP5	73508	β	0.215	0.147	0.132	0.129
			$[95\%CI]$	$[0.070, 0.333]$	$[0.034, 0.281]$	$[0.038, 0.261]$	$[0.016, 0.246]$
			p-value	0.00158	0.02	0.022	2.70E-02
rs212517	ECE1	71344	β	0.079	0.127	0.128	0.152
			$[95\%CI]$	$[-0.013, 0.167]$	$[0.032, 0.198]$	$[0.047, 0.205]$	$[0.066, 0.226]$
			p-value	8.80E-02	2.15E-03	1.18E-03	2.23E-04
rs7359336	NFAT5	73094	β	0.124	0.107	0.09	0.084
			$[95\%CI]$	$[0.038, 0.211]$	$[0.007, 0.186]$	$[-0.007, 0.161]$	$[0.006, 0.168]$
			p-value	0.00389	0.019	3.60E-02	0.044
rs2682552	XRCC1	73106	β	0.133	0.127	0.12	0.115
			$[95\%CI]$	$[0.039, 0.254]$	$[0.017, 0.224]$	$[0.010, 0.208]$	$[0.016, 0.216]$
			p-value	0.014	0.016	0.016	0.024
rs17085675	PCSK1	69775	β	0.187	0.131	0.138	0.139
			$[95\%CI]$	$[0.070, 0.274]$	$[0.046, 0.236]$	$[0.047, 0.235]$	$[0.051, 0.237]$
			p-value	0.000293	0.00656	0.00381	0.00327
rs11102986	RXRA	71224	β	0.252	0.203	0.154	
			$[95\%CI]$	$[0.145, 0.366]$	$[0.083, 0.321]$	$[0.059, 0.265]$	

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		25%	30%	35%	40%
			p-value	5.64E-06	9.24E-04	0.00367	
rs12603813	PLCD3	71116	β	0.053	0.086	0.125	0.116
			$[95\%CI]$	$[-0.049, 0.162]$	$[-0.007, 0.181]$	$[0.025, 0.208]$	$[0.013, 0.204]$
			p-value	3.32E-01	7.70E-02	7.45E-03	0.017
rs6219	IGF1	73562	β	0.1	0.027	0.026	0.082
			$[95\%CI]$	$[-0.076, 0.216]$	$[-0.098, 0.165]$	$[-0.096, 0.184]$	$[-0.044, 0.216]$
			p-value	1.86E-01	6.85E-01	7.15E-01	2.16E-01
rs2234693	ESR1	69772	β	0.119	0.087	0.084	0.092
			$[95\%CI]$	$[0.017, 0.193]$	$[-0.001, 0.159]$	$[0.002, 0.163]$	$[0.007, 0.170]$
			p-value	0.00738	0.034	0.041	0.028
rs46522	UBE2Z	64204	β	0.121	0.134	0.138	0.122
			$[95\%CI]$	$[0.013, 0.203]$	$[0.049, 0.230]$	$[0.052, 0.214]$	$[0.039, 0.209]$
			p-value	0.013	0.00371	9.60E-04	0.00435
rs891088	INSR	73562	β	0.127	0.129	0.134	0.109
			$[95\%CI]$	$[0.030, 0.220]$	$[0.042, 0.221]$	$[0.051, 0.220]$	$[0.022, 0.193]$
			p-value	9.36E-03	4.64E-03	1.67E-03	1.20E-02
rs9857730	VILL	73358	β	0.15	0.13	0.135	0.101
			$[95\%CI]$	$[0.059, 0.274]$	$[0.051, 0.247]$	$[0.041, 0.240]$	$[0.006, 0.197]$
			p-value	5.99E-03	0.00898	0.00799	0.038
rs10185680	MFSD2B	73232	β	0.145	0.133	0.094	0.121
			$[95\%CI]$	$[0.077, 0.244]$	$[0.057, 0.221]$	$[0.013, 0.173]$	$[0.036, 0.190]$
			p-value	0.001	0.001	0.020	0.00212
rs7163907	PTPN9	73234	β	-0.125	-0.107	-0.116	-0.100
			$[95\%CI]$	$[-0.235, -0.026]$	$[-0.195, -0.014]$	$[-0.203, -0.017]$	$[-0.187, -0.003]$
			p-value	0.019	0.019	0.014	0.033
rs2176167	NOP58	73508	β	0.118	0.127	0.117	0.108
			$[95\%CI]$	$[0.033, 0.212]$	$[0.045, 0.216]$	$[0.028, 0.194]$	$[0.030, 0.186]$
			p-value	0.009	0.004	0.006	0.007
rs7557989	THADA	64417	β	0.199	0.132	0.136	0.109
			$[95\%CI]$	$[0.106, 0.298]$	$[0.043, 0.224]$	$[0.045, 0.233]$	$[0.020, 0.197]$
			p-value	4.50E-05	0.004	0.004	0.017
rs510769	OPRM1	64413	β	0.134	0.086	0.105	0.134
			$[95\%CI]$	$[0.034, 0.241]$	$[-0.001, 0.199]$	$[0.001, 0.196]$	$[0.023, 0.233]$
			p-value	0.009	0.092	0.034	0.012
rs7756224	NMBR	60624	β	0.176	0.168	0.134	0.128
			$[95\%CI]$	$[0.081, 0.282]$	$[0.074, 0.254]$	$[0.037, 0.222]$	$[0.040, 0.210]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		25%	30%	35%	40%
			p-value	6.96E-04	0.000	0.005	0.003
rs2282537	POU2F3	64419	β	0.107	0.139	0.137	0.093
			$[95\%CI]$	$[-0.043, 0.232]$	$[0.017, 0.251]$	$[-0.001, 0.245]$	$[-0.020, 0.223]$
			p-value	0.126	0.021	0.028	0.131
rs7853859	CENPP	69662	β	0.124	0.103	0.052	0.061
			$[95\%CI]$	$[0.019, 0.199]$	$[0.014, 0.181]$	$[-0.032, 0.140]$	$[-0.019, 0.146]$
			p-value	0.007	0.015	0.233	0.143
rs2229642	ITPR3	52677	β	0.036	0.05	0.095	0.087
			$[95\%CI]$	$[-0.069, 0.149]$	$[-0.048, 0.158]$	$[-0.007, 0.188]$	$[0.000, 0.187]$
			p-value	0.513	0.344	0.057	0.066
rs5015437	LMF1	64185	β	0.126	0.147	0.152	0.15
			$[95\%CI]$	$[0.021, 0.212]$	$[0.056, 0.242]$	$[0.069, 0.245]$	$[0.067, 0.244]$
			p-value	0.009	0.00173	0.001	0.001
rs7963565	KNTC1	73568	β	0.072	0.054	0.083	0.060
			$[95\%CI]$	$[-0.002, 0.180]$	$[-0.018, 0.156]$	$[-0.006, 0.166]$	$[-0.019, 0.154]$
			p-value	0.120	0.225	0.058	0.167
rs696	NFKBIA	73461	β	0.051	0.100	0.120	0.082
			$[95\%CI]$	$[-0.030, 0.147]$	$[0.002, 0.177]$	$[0.026, 0.194]$	$[-0.003, 0.164]$
			p-value	0.262	0.024	0.00444	0.051
rs25656	NFATC1	48455	β	0.145	0.131	0.079	0.089
			$[95\%CI]$	$[0.024, 0.259]$	$[0.026, 0.233]$	$[-0.022, 0.210]$	$[-0.019, 0.205]$
			p-value	0.015	0.014	0.181	0.113
rs3100776	IHH	64334	β	0.168	0.18	0.151	0.199
			$[95\%CI]$	$[-0.078, 0.377]$	$[-0.045, 0.425]$	$[-0.074, 0.372]$	$[-0.050, 0.425]$
			p-value	0.155	0.141	0.181	0.097
rs12145922	PKN2	69346	β	0.065	0.093	0.122	0.095
			$[95\%CI]$	$[-0.025, 0.155]$	$[0.008, 0.180]$	$[0.039, 0.200]$	$[0.015, 0.174]$
			p-value	0.154	0.036	0.003	0.018
rs526134	USP37	60625	β	0.057	0.067	0.079	0.074
			$[95\%CI]$	$[-0.041, 0.153]$	$[-0.022, 0.154]$	$[-0.011, 0.165]$	$[-0.016, 0.164]$
			p-value	0.247	0.142	0.080	0.115
rs7004280	RPS20	64415	β	0.097	0.085	0.199	0.311
			$[95\%CI]$	$[-0.199, 0.322]$	$[-0.162, 0.342]$	$[-0.062, 0.476]$	$[0.033, 0.537]$
			p-value	0.469	0.511	0.145	0.015
rs4808199	GATAD2A	64415	β	0.100	0.091	0.103	0.081
			$[95\%CI]$	$[-0.042, 0.210]$	$[-0.017, 0.195]$	$[-0.024, 0.199]$	$[-0.029, 0.195]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		25%	30%	35%	40%
			p-value	0.117	0.099	0.070	0.161
rs3821009	PDE11A	69745	β	0.104	0.127	0.163	0.131
			$[95\%CI]$	$[-0.086, 0.257]$	$[-0.041, 0.275]$	$[0.009, 0.314]$	$[-0.009, 0.274]$
			p-value	0.228	0.117	0.036	0.073
rs2291617	METTL1	73099	β	0.09	0.043	0.056	0.078
			$[95\%CI]$	$[-0.017, 0.168]$	$[-0.030, 0.137]$	$[-0.026, 0.142]$	$[-0.004, 0.161]$
			p-value	0.056	0.309	0.194	0.066
rs1051168	SEC11A	64399	β	0.052	0.029	0.072	0.067
			$[95\%CI]$	$[-0.050, 0.149]$	$[-0.076, 0.120]$	$[-0.038, 0.157]$	$[-0.035, 0.153]$
			p-value	0.307	0.557	0.148	0.162
rs803932	ASTN2	64105	β	0.127	0.113	0.093	0.084
			$[95\%CI]$	$[0.024, 0.220]$	$[0.027, 0.210]$	$[0.009, 0.200]$	$[0.004, 0.189]$
			p-value	0.012	0.015	0.054	0.079
rs2247870	GPR98	69775	β	0.100	0.081	0.129	0.094
			$[95\%CI]$	$[-0.006, 0.178]$	$[-0.006, 0.166]$	$[0.041, 0.206]$	$[0.014, 0.179]$
			p-value	0.036	0.069	0.002	0.026
rs12503378	NUDT6	73569	β	0.025	0.036	0.111	0.076
			$[95\%CI]$	$[-0.079, 0.167]$	$[-0.051, 0.164]$	$[-0.015, 0.203]$	$[-0.022, 0.190]$
			p-value	0.686	0.503	0.044	0.166
rs2145923	NPR2	69777	β	0.092	0.089	0.054	0.051
			$[95\%CI]$	$[-0.030, 0.202]$	$[-0.027, 0.193]$	$[-0.057, 0.166]$	$[-0.057, 0.163]$
			p-value	0.118	0.109	0.341	0.363
rs12603582	ITGB3	73541	β	0.070	0.007	0.010	0.045
			$[95\%CI]$	$[-0.021, 0.185]$	$[-0.087, 0.123]$	$[-0.068, 0.128]$	$[-0.041, 0.139]$
			p-value	0.184	0.891	0.843	0.334
rs4864548	CLOCK	73568	β	0.000	-0.001	-0.003	0.034
			$[95\%CI]$	$[-0.088, 0.097]$	$[-0.093, 0.077]$	$[-0.097, 0.072]$	$[-0.060, 0.110]$
			p-value	1.000	0.99	0.951	0.428
rs2735469	MRPL23	67563	β	0.026	0.049	0.019	0.028
			$[95\%CI]$	$[-0.089, 0.179]$	$[-0.058, 0.174]$	$[-0.079, 0.150]$	$[-0.083, 0.143]$
			p-value	0.705	0.403	0.741	0.631
rs1051431	MPHOSPH9	73561	β	0.000	0.000	0.000	0.000
			$[95\%CI]$	$[-0.098, 0.117]$	$[-0.093, 0.098]$	$[-0.078, 0.111]$	$[-0.104, 0.084]$
			p-value	1.000	0.998	1.000	1.000
rs41132	AP3B1	73441	β	0.037	0.008	0.000	0.000
			$[95\%CI]$	$[-0.063, 0.144]$	$[-0.070, 0.128]$	$[-0.087, 0.095]$	$[-0.091, 0.083]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		25%	30%	35%	40%
			p-value	0.487	0.879	1.000	1.000
rs6127698	MC3R	73486	β	0.015	0.023	0.033	0.050
			$[95\%CI]$	$[-0.062, 0.110]$	$[-0.047, 0.117]$	$[-0.039, 0.122]$	$[-0.032, 0.127]$
			p-value	0.733	0.584	0.415	0.220
rs1481892	ARNTL	73543	β	0.025	-0.010	0.007	0.032
			$[95\%CI]$	$[-0.067, 0.118]$	$[-0.105, 0.072]$	$[-0.067, 0.108]$	$[-0.053, 0.117]$
			p-value	0.593	0.823	0.876	0.457
rs2633442	MKRN2	60624	β	-0.040	-0.032	-0.005	0.008
			$[95\%CI]$	$[-0.136, 0.058]$	$[-0.117, 0.061]$	$[-0.100, 0.084]$	$[-0.077, 0.094]$
			p-value	0.423	0.493	0.91	0.85
rs1535	FADS1	73553	β	-0.025	0.002	0.000	0.030
			$[95\%CI]$	$[-0.116, 0.063]$	$[-0.087, 0.088]$	$[-0.072, 0.104]$	$[-0.056, 0.115]$
			p-value	0.585	0.972	1.000	0.487
rs10861148	HSP90B1	73557	β	0.000	0.000	0.000	0.020
			$[95\%CI]$	$[-0.134, 0.148]$	$[-0.135, 0.133]$	$[-0.156, 0.112]$	$[-0.122, 0.134]$
			p-value	1.000	1.000	1.000	0.758
GS-Height		73570	β	0.179	0.178	0.174	0.173
			$[95\%CI]$	$[0.170, 0.187]$	$[0.168, 0.187]$	$[0.166, 0.183]$	$[0.165, 0.181]$
			p-value	<2.2E-308	<2.2E-308	<2.2E-308	<2.2E-308
			VarianceExplained	1.967%	1.897%	1.863%	1.849%

Part 3 - 45% to 60%

SNP	Gene	N		45%	50%	55%	60%
rs1042725	HMGA2	73105	β	0.55	0.543	0.561	0.552
			$[95\%CI]$	$[0.470, 0.629]$	$[0.465, 0.633]$	$[0.482, 0.642]$	$[0.469, 0.630]$
			p-value	5.17E-42	5.65E-37	1.29E-43	6.94E-42
rs2853977	HCP5	44699	β	0.678	0.656	0.632	0.589
			$[95\%CI]$	$[0.581, 0.784]$	$[0.559, 0.758]$	$[0.528, 0.735]$	$[0.497, 0.687]$
			p-value	7.64E-39	4.05E-38	1.08E-32	2.48E-33
rs3782415	SOCS2	73568	β	0.45	0.461	0.477	0.523
			$[95\%CI]$	$[0.349, 0.550]$	$[0.355, 0.557]$	$[0.385, 0.590]$	$[0.411, 0.621]$
			p-value	1.23E-18	3.83E-19	3.46E-20	1.38E-22
rs780094	GCKR	73548	β	0.393	0.36	0.36	0.372
			$[95\%CI]$	$[0.312, 0.474]$	$[0.265, 0.432]$	$[0.270, 0.445]$	$[0.303, 0.464]$
			p-value	1.17E-21	2.98E-17	4.21E-16	9.05E-20
rs9892365	TBX2	73566	β	0.413	0.391	0.387	0.393
			$[95\%CI]$	$[0.319, 0.497]$	$[0.302, 0.473]$	$[0.296, 0.472]$	$[0.302, 0.471]$
			p-value	5.42E-20	1.47E-19	8.78E-18	4.35E-20
rs7137534	PDE3A	73567	β	0.393	0.361	0.387	0.397
			$[95\%CI]$	$[0.311, 0.480]$	$[0.280, 0.451]$	$[0.295, 0.472]$	$[0.305, 0.481]$
			p-value	2.90E-20	2.34E-16	3.84E-18	6.95E-19
rs1776897	HMGA1	64379	β	0.635	0.668	0.686	0.662
			$[95\%CI]$	$[0.484, 0.797]$	$[0.511, 0.816]$	$[0.529, 0.843]$	$[0.514, 0.808]$
			p-value	2.41E-15	1.15E-17	7.24E-18	8.12E-19
rs572169	GHSR	73554	β	0.367	0.359	0.35	0.323
			$[95\%CI]$	$[0.274, 0.449]$	$[0.266, 0.435]$	$[0.256, 0.431]$	$[0.237, 0.413]$
			p-value	1.61E-16	9.03E-17	7.91E-15	9.09E-13
rs2679178	NPPC	73567	β	0.528	0.551	0.597	0.65
			$[95\%CI]$	$[0.372, 0.673]$	$[0.420, 0.686]$	$[0.446, 0.734]$	$[0.491, 0.784]$
			p-value	7.10E-12	8.30E-16	4.37E-16	2.77E-18
rs2053156	GRB2	73536	β	0.374	0.359	0.345	0.378
			$[95\%CI]$	$[0.275, 0.494]$	$[0.251, 0.466]$	$[0.243, 0.460]$	$[0.248, 0.477]$
			p-value	1.53E-11	5.25E-11	5.11E-10	1.74E-10
rs9930741	ERI2	73143	β	0.341	0.322	0.319	0.31
			$[95\%CI]$	$[0.260, 0.418]$	$[0.238, 0.400]$	$[0.239, 0.404]$	$[0.227, 0.401]$
			p-value	3.03E-17	3.18E-15	3.74E-14	3.57E-12
rs2854207	CSH2	67567	β	0.284	0.271	0.277	0.267
			$[95\%CI]$	$[0.174, 0.383]$	$[0.176, 0.366]$	$[0.187, 0.370]$	$[0.184, 0.374]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		45%	50%	55%	60%
			p-value	5.07E-08	2.13E-08	2.60E-09	1.98E-08
rs4320932	IGF2	71204	β	0.377	0.358	0.359	0.387
			$[95\%CI]$	$[0.274, 0.482]$	$[0.255, 0.469]$	$[0.257, 0.465]$	$[0.286, 0.494]$
			p-value	1.43E-12	8.03E-11	1.44E-11	3.08E-13
rs752313	EZH1	69774	β	0.246	0.258	0.239	0.223
			$[95\%CI]$	$[0.161, 0.328]$	$[0.170, 0.331]$	$[0.147, 0.319]$	$[0.134, 0.307]$
			p-value	6.71E-09	1.77E-10	2.15E-08	3.34E-07
rs709939	SAMD4A	73569	β	0.233	0.24	0.249	0.247
			$[95\%CI]$	$[0.151, 0.308]$	$[0.161, 0.314]$	$[0.159, 0.324]$	$[0.157, 0.328]$
			p-value	8.43E-09	7.34E-10	2.35E-09	1.09E-08
rs1036477	FBN1	73539	β	0.474	0.429	0.436	0.453
			$[95\%CI]$	$[0.338, 0.609]$	$[0.296, 0.571]$	$[0.289, 0.562]$	$[0.271, 0.571]$
			p-value	7.45E-12	1.25E-09	2.86E-10	3.17E-09
rs158676	CDK5RAP1	73562	β	0.255	0.245	0.215	0.243
			$[95\%CI]$	$[0.166, 0.343]$	$[0.152, 0.325]$	$[0.129, 0.304]$	$[0.146, 0.325]$
			p-value	2.26E-08	2.94E-08	1.59E-06	1.47E-07
rs1822469	PPP3R1	69710	β	0.269	0.253	0.231	0.24
			$[95\%CI]$	$[0.187, 0.351]$	$[0.159, 0.324]$	$[0.138, 0.318]$	$[0.151, 0.319]$
			p-value	2.08E-10	1.42E-09	4.01E-07	1.80E-08
rs258281	RAB26	67411	β	0.324	0.293	0.267	0.26
			$[95\%CI]$	$[0.198, 0.431]$	$[0.181, 0.400]$	$[0.148, 0.382]$	$[0.138, 0.368]$
			p-value	4.01E-08	1.19E-07	7.21E-06	1.05E-05
rs9366637	HFE	73557	β	0.486	0.45	0.477	0.546
			$[95\%CI]$	$[0.318, 0.679]$	$[0.295, 0.631]$	$[0.316, 0.647]$	$[0.371, 0.719]$
			p-value	1.02E-07	1.22E-07	7.23E-09	6.28E-10
rs551219	COL24A1	73076	β	0.26	0.261	0.264	0.254
			$[95\%CI]$	$[0.175, 0.359]$	$[0.176, 0.358]$	$[0.181, 0.353]$	$[0.164, 0.344]$
			p-value	3.13E-08	1.00E-08	1.73E-09	2.95E-08
rs13072536	ITIH4	73519	β	0.221	0.246	0.233	0.242
			$[95\%CI]$	$[0.126, 0.322]$	$[0.156, 0.336]$	$[0.129, 0.328]$	$[0.142, 0.329]$
			p-value	1.07E-05	6.59E-08	4.03E-06	3.38E-07
rs7522692	PIGC	64398	β	0.29	0.313	0.316	0.307
			$[95\%CI]$	$[0.194, 0.411]$	$[0.208, 0.412]$	$[0.230, 0.440]$	$[0.204, 0.411]$
			p-value	2.14E-07	1.73E-09	5.24E-09	9.29E-09
rs13076290	CTNNB1	73154	β	0.203	0.166	0.176	0.213
			$[95\%CI]$	$[0.120, 0.278]$	$[0.082, 0.240]$	$[0.088, 0.259]$	$[0.130, 0.298]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		45%	50%	55%	60%
			p-value	6.02E-07	3.47E-05	4.91E-05	7.00E-07
rs1636255	GNA12	64371	β	0.232	0.25	0.263	0.265
			$[95\%CI]$	$[0.135, 0.324]$	$[0.157, 0.340]$	$[0.160, 0.356]$	$[0.175, 0.370]$
			p-value	1.58E-06	1.02E-07	1.73E-07	1.14E-07
rs3796529	REST	73119	β	0.297	0.287	0.262	0.222
			$[95\%CI]$	$[0.199, 0.410]$	$[0.191, 0.394]$	$[0.158, 0.369]$	$[0.117, 0.327]$
			p-value	5.19E-08	3.70E-08	7.82E-07	3.09E-05
rs1866146	POMC	73516	β	0.261	0.263	0.303	0.286
			$[95\%CI]$	$[0.176, 0.344]$	$[0.185, 0.346]$	$[0.208, 0.378]$	$[0.198, 0.376]$
			p-value	8.57E-10	2.25E-10	1.88E-12	2.29E-10
rs9844666	PCCB	73511	β	0.254	0.207	0.189	0.202
			$[95\%CI]$	$[0.156, 0.348]$	$[0.097, 0.297]$	$[0.099, 0.290]$	$[0.092, 0.298]$
			p-value	2.20E-07	5.75E-05	1.14E-04	1.32E-04
rs8071847	POLR2A	73565	β	0.303	0.285	0.304	0.264
			$[95\%CI]$	$[0.193, 0.393]$	$[0.193, 0.390]$	$[0.214, 0.404]$	$[0.173, 0.364]$
			p-value	2.31E-09	1.87E-08	4.69E-10	5.00E-08
rs3783937	FBLN5	73104	β	0.304	0.278	0.273	0.262
			$[95\%CI]$	$[0.207, 0.396]$	$[0.185, 0.376]$	$[0.165, 0.356]$	$[0.163, 0.357]$
			p-value	2.33E-10	1.02E-08	1.71E-08	8.38E-08
rs11080149	NF1	73567	β	0.296	0.284	0.306	0.296
			$[95\%CI]$	$[0.182, 0.430]$	$[0.171, 0.424]$	$[0.181, 0.434]$	$[0.174, 0.412]$
			p-value	2.75E-06	1.33E-05	1.75E-06	9.07E-07
rs17472113	ZAR1	58807	β	0.286	0.302	0.292	0.296
			$[95\%CI]$	$[0.190, 0.407]$	$[0.191, 0.400]$	$[0.195, 0.397]$	$[0.201, 0.404]$
			p-value	3.00E-07	1.66E-08	1.71E-08	1.16E-08
rs490634	CISH	61124	β	0.414		0.341	0.34
			$[95\%CI]$	$[0.271, 0.524]$		$[0.216, 0.494]$	$[0.228, 0.466]$
			p-value	2.11E-10		1.32E-06	3.11E-08
rs17622208	SLC22A5	73531	β	0.22	0.211	0.21	0.237
			$[95\%CI]$	$[0.143, 0.303]$	$[0.131, 0.288]$	$[0.126, 0.294]$	$[0.163, 0.330]$
			p-value	8.69E-08	1.35E-07	1.00E-06	2.95E-08
rs2982712	ESR1	73566	β	0.166	0.153	0.178	0.174
			$[95\%CI]$	$[0.078, 0.249]$	$[0.075, 0.235]$	$[0.096, 0.263]$	$[0.096, 0.260]$
			p-value	1.40E-04	2.08E-04	3.31E-05	3.16E-05
rs1950500	NFATC4	64408	β	0.237	0.26	0.278	0.283
			$[95\%CI]$	$[0.136, 0.340]$	$[0.174, 0.368]$	$[0.174, 0.368]$	$[0.200, 0.384]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		45%	50%	55%	60%
			p-value	4.63E-06	1.70E-07	3.02E-08	1.34E-09
rs1476387	PPIL6	73563	β	0.217	0.181	0.157	0.159
			$[95\%CI]$	$[0.127, 0.303]$	$[0.095, 0.259]$	$[0.082, 0.244]$	$[0.068, 0.236]$
			p-value	1.24E-06	1.31E-05	1.61E-04	2.27E-04
rs4946932	FOXO3	73565	β	0.205	0.212	0.216	0.262
			$[95\%CI]$	$[0.122, 0.297]$	$[0.124, 0.293]$	$[0.132, 0.307]$	$[0.166, 0.355]$
			p-value	4.76E-06	1.04E-06	1.16E-06	4.65E-08
rs1800783	NOS3	71277	β	0.226	0.195	0.202	0.212
			$[95\%CI]$	$[0.138, 0.318]$	$[0.112, 0.277]$	$[0.120, 0.288]$	$[0.124, 0.294]$
			p-value	1.07E-06	4.14E-06	2.48E-06	1.05E-06
rs6718902	STAT1	73557	β	0.22	0.217	0.23	0.229
			$[95\%CI]$	$[0.120, 0.314]$	$[0.127, 0.317]$	$[0.139, 0.331]$	$[0.131, 0.320]$
			p-value	8.31E-06	9.85E-06	2.75E-06	2.40E-06
rs2425019	MMP24	69770	β	0.209	0.222	0.217	0.204
			$[95\%CI]$	$[0.127, 0.300]$	$[0.126, 0.290]$	$[0.130, 0.293]$	$[0.120, 0.286]$
			p-value	2.22E-06	1.08E-07	1.76E-07	1.62E-06
rs6731022	EIF2AK3	69767	β	0.232	0.239	0.236	0.222
			$[95\%CI]$	$[0.138, 0.312]$	$[0.157, 0.333]$	$[0.148, 0.335]$	$[0.134, 0.316]$
			p-value	2.20E-07	9.98E-08	5.92E-07	1.76E-06
rs12940055	MAP3K3	73566	β	0.303	0.299	0.302	0.374
			$[95\%CI]$	$[0.177, 0.435]$	$[0.176, 0.421]$	$[0.188, 0.435]$	$[0.231, 0.495]$
			p-value	4.14E-06	1.50E-06	1.47E-06	3.88E-08
rs864745	JAZF1	60616	β	0.205	0.216	0.22	0.227
			$[95\%CI]$	$[0.118, 0.290]$	$[0.127, 0.300]$	$[0.135, 0.307]$	$[0.132, 0.316]$
			p-value	2.67E-06	1.21E-06	4.95E-07	1.34E-06
rs6487088	PDE3A	73563	β	0.26	0.259	0.248	0.232
			$[95\%CI]$	$[0.153, 0.357]$	$[0.135, 0.346]$	$[0.129, 0.340]$	$[0.118, 0.337]$
			p-value	6.07E-07	1.20E-06	3.72E-06	3.71E-05
rs4973410	NCL	64408	β	0.186	0.198	0.189	0.2
			$[95\%CI]$	$[0.103, 0.273]$	$[0.101, 0.275]$	$[0.107, 0.281]$	$[0.110, 0.289]$
			p-value	1.33E-05	9.89E-06	2.13E-05	9.82E-06
rs451061	PRKCZ	71365	β	0.18	0.198	0.21	0.179
			$[95\%CI]$	$[0.100, 0.277]$	$[0.118, 0.286]$	$[0.129, 0.291]$	$[0.099, 0.268]$
			p-value	6.38E-05	4.40E-06	4.48E-07	2.75E-05
rs832575	MAP3K1	73539	β	0.267	0.261	0.287	0.248
			$[95\%CI]$	$[0.148, 0.401]$	$[0.142, 0.378]$	$[0.154, 0.397]$	$[0.124, 0.376]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		45%	50%	55%	60%
			p-value	3.59E-05	1.52E-05	3.47E-06	1.26E-04
rs4955526	EPHB1	64417	β	0.221	0.195	0.225	0.243
			$[95\%CI]$	$[0.128, 0.309]$	$[0.106, 0.292]$	$[0.132, 0.312]$	$[0.148, 0.328]$
			p-value	1.35E-06	4.26E-05	9.46E-07	1.43E-07
rs8038415	IGF1R	73516	β	0.194		0.196	0.202
			$[95\%CI]$	$[0.107, 0.273]$		$[0.117, 0.285]$	$[0.119, 0.284]$
			p-value	5.46E-06		5.67E-06	2.03E-06
rs7578199	HDLBP	73569	β	0.175	0.154	0.164	0.196
			$[95\%CI]$	$[0.075, 0.269]$	$[0.051, 0.237]$	$[0.072, 0.269]$	$[0.087, 0.286]$
			p-value	3.92E-04	1.26E-03	1.12E-03	1.16E-04
rs7020782	PAPPA	73566	β	0.191	0.191	0.209	0.189
			$[95\%CI]$	$[0.097, 0.273]$	$[0.103, 0.274]$	$[0.116, 0.287]$	$[0.099, 0.281]$
			p-value	1.78E-05	1.18E-05	1.85E-06	4.23E-05
rs2229712	RPS6KA1	48240	β	0.286	0.254	0.242	0.224
			$[95\%CI]$	$[0.163, 0.408]$	$[0.137, 0.367]$	$[0.116, 0.360]$	$[0.108, 0.338]$
			p-value	4.33E-06	1.34E-05	1.03E-04	1.53E-04
rs7572476	BOK	71367	β	0.13	0.128	0.156	0.213
			$[95\%CI]$	$[0.051, 0.223]$	$[0.043, 0.209]$	$[0.079, 0.249]$	$[0.133, 0.299]$
			p-value	2.76E-03	2.25E-03	2.88E-04	3.63E-07
rs2066807	PAN2	71822	β	0.24	0.258	0.275	0.298
			$[95\%CI]$	$[0.115, 0.342]$	$[0.149, 0.366]$	$[0.167, 0.409]$	$[0.191, 0.408]$
			p-value	4.30E-05	3.77E-06	8.05E-06	7.53E-08
rs1516796	ACAN	69308	β	0.141	0.156	0.184	
			$[95\%CI]$	$[0.064, 0.239]$	$[0.083, 0.248]$	$[0.095, 0.262]$	
			p-value	1.76E-03	1.95E-04	1.43E-05	
rs6180	GHR	73552	β	0.229	0.218	0.193	0.16
			$[95\%CI]$	$[0.144, 0.314]$	$[0.137, 0.290]$	$[0.105, 0.275]$	$[0.079, 0.245]$
			p-value	1.12E-07	2.55E-08	5.78E-06	1.43E-04
rs8055190	LRRC36	64416	β	0.476	0.414	0.414	0.5
			$[95\%CI]$	$[0.249, 0.675]$	$[0.194, 0.620]$	$[0.230, 0.622]$	$[0.288, 0.730]$
			p-value	1.59E-05	1.14E-04	3.24E-05	8.15E-06
rs17106235	FAF1	60489	β	0.301	0.294	0.329	0.325
			$[95\%CI]$	$[0.152, 0.444]$	$[0.159, 0.446]$	$[0.176, 0.477]$	$[0.178, 0.494]$
			p-value	5.76E-05	4.90E-05	2.08E-05	5.43E-05
rs3739707	LPAR1	73568	β	0.211	0.197	0.192	0.203
			$[95\%CI]$	$[0.115, 0.310]$	$[0.095, 0.282]$	$[0.090, 0.277]$	$[0.098, 0.294]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		45%	50%	55%	60%
			p-value	2.67E-05	3.34E-05	5.41E-05	5.81E-05
rs674424	ABCG4	73568	β	0.277	0.253	0.259	0.246
			$[95\%CI]$	$[0.179, 0.373]$	$[0.163, 0.339]$	$[0.166, 0.356]$	$[0.137, 0.339]$
			p-value	1.84E-08	9.91E-09	8.08E-08	1.37E-06
rs12225387	NEU3	60621	β	0.176	0.184	0.21	0.238
			$[95\%CI]$	$[0.068, 0.278]$	$[0.092, 0.281]$	$[0.100, 0.305]$	$[0.124, 0.340]$
			p-value	9.75E-04	1.36E-04	5.65E-05	1.85E-05
rs3812265	CNOT4	69776	β	0.185	0.169	0.186	0.2
			$[95\%CI]$	$[0.084, 0.284]$	$[0.062, 0.259]$	$[0.091, 0.279]$	$[0.108, 0.309]$
			p-value	2.71E-04	7.68E-04	9.29E-05	1.08E-04
rs10208728	IHH	64409	β	0.308	0.277	0.283	0.263
			$[95\%CI]$	$[0.157, 0.449]$	$[0.118, 0.430]$	$[0.154, 0.423]$	$[0.103, 0.410]$
			p-value	2.16E-05	5.04E-04	3.73E-05	7.36E-04
rs291979	GRK5	73568	β	0.215	0.202	0.19	0.201
			$[95\%CI]$	$[0.123, 0.318]$	$[0.105, 0.294]$	$[0.094, 0.288]$	$[0.101, 0.300]$
			p-value	1.31E-05	2.97E-05	1.05E-04	6.67E-05
rs2715553	RARA	69206	β	0.164	0.145	0.129	0.128
			$[95\%CI]$	$[0.084, 0.261]$	$[0.075, 0.237]$	$[0.040, 0.212]$	$[0.034, 0.204]$
			p-value	2.51E-04	4.74E-04	3.16E-03	0.00291
rs2057291	GNAS	70100	β	0.188	0.182	0.182	0.212
			$[95\%CI]$	$[0.098, 0.281]$	$[0.091, 0.260]$	$[0.087, 0.269]$	$[0.125, 0.299]$
			p-value	5.88E-05	3.10E-05	8.77E-05	1.67E-06
rs4803520	GRIK5	69747	β	0.253	0.214	0.201	0.191
			$[95\%CI]$	$[0.128, 0.389]$	$[0.095, 0.368]$	$[0.069, 0.330]$	$[0.059, 0.329]$
			p-value	1.31E-04	2.02E-03	2.23E-03	5.16E-03
rs10736682	APLNR	73090	β	0.174	0.15	0.174	0.176
			$[95\%CI]$	$[0.092, 0.260]$	$[0.074, 0.233]$	$[0.089, 0.257]$	$[0.098, 0.266]$
			p-value	5.17E-05	1.69E-04	4.52E-05	4.75E-05
rs2909430	TP53	73412	β	0.253	0.231	0.21	0.178
			$[95\%CI]$	$[0.120, 0.384]$	$[0.117, 0.350]$	$[0.090, 0.350]$	$[0.058, 0.303]$
			p-value	1.90E-04	0.0000941	0.00139	0.00384
rs12050767	CYP19A1	73563	β	0.108		0.15	0.172
			$[95\%CI]$	$[0.024, 0.193]$		$[0.072, 0.245]$	$[0.088, 0.257]$
			p-value	0.012		5.69E-04	6.03E-05
rs602633	PSRC1	64375	β	0.172	0.185	0.187	0.194
			$[95\%CI]$	$[0.068, 0.286]$	$[0.084, 0.298]$	$[0.086, 0.309]$	$[0.091, 0.295]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		45%	50%	55%	60%
			p-value	1.80E-03	7.09E-04	1.22E-03	1.90E-04
rs1738475	HTR1D	64411	β	0.127	0.083	0.087	0.1
			$[95\%CI]$	$[0.045, 0.218]$	$[-0.002, 0.179]$	$[0.002, 0.174]$	$[0.018, 0.186]$
			p-value	4.02E-03	7.50E-02	5.20E-02	2.20E-02
rs17754	RFC1	73558	β	0.163	0.15	0.154	0.173
			$[95\%CI]$	$[0.080, 0.247]$	$[0.069, 0.231]$	$[0.068, 0.246]$	$[0.087, 0.259]$
			p-value	0.000129	0.00033	6.85E-04	7.45E-05
rs17541471	NPR3	69777	β	0.17	0.152	0.143	0.133
			$[95\%CI]$	$[0.073, 0.273]$	$[0.048, 0.252]$	$[0.035, 0.247]$	$[0.030, 0.250]$
			p-value	0.000836	3.31E-03	7.92E-03	1.80E-02
rs1342586	TGFB2	69745	β	0.172	0.163	0.154	0.133
			$[95\%CI]$	$[0.059, 0.280]$	$[0.068, 0.262]$	$[0.054, 0.253]$	$[0.006, 0.214]$
			p-value	0.00271	0.000912	0.00225	0.011
rs3814115	PCSK5	73537	β	0.126	0.111	0.117	0.143
			$[95\%CI]$	$[0.032, 0.214]$	$[0.022, 0.202]$	$[0.028, 0.215]$	$[0.052, 0.232]$
			p-value	0.00602	0.015	1.40E-02	1.75E-03
rs1780616	LBP	73560	β	0.16	0.146	0.142	0.133
			$[95\%CI]$	$[0.075, 0.252]$	$[0.059, 0.233]$	$[0.054, 0.228]$	$[0.050, 0.222]$
			p-value	0.000411	0.000995	0.00107	0.00213
rs3736228	LRP5	73508	β	0.126	0.109	0.115	0.087
			$[95\%CI]$	$[0.006, 0.252]$	$[-0.014, 0.216]$	$[0.002, 0.237]$	$[-0.034, 0.207]$
			p-value	4.60E-02	0.064	0.054	1.56E-01
rs212517	ECE1	71344	β	0.13	0.139	0.14	0.151
			$[95\%CI]$	$[0.045, 0.213]$	$[0.061, 0.221]$	$[0.068, 0.230]$	$[0.072, 0.247]$
			p-value	2.16E-03	6.51E-04	6.46E-04	7.01E-04
rs7359336	NFAT5	73094	β	0.087	0.082	0.11	0.1
			$[95\%CI]$	$[0.003, 0.168]$	$[-0.005, 0.160]$	$[0.027, 0.194]$	$[0.022, 0.188]$
			p-value	4.00E-02	5.30E-02	0.01	1.80E-02
rs2682552	XRCC1	73106	β	0.11	0.09	0.127	0.152
			$[95\%CI]$	$[0.008, 0.207]$	$[-0.017, 0.191]$	$[0.010, 0.221]$	$[0.045, 0.259]$
			p-value	0.033	8.80E-02	1.80E-02	5.52E-03
rs17085675	PCSK1	69775	β	0.14	0.131	0.132	0.142
			$[95\%CI]$	$[0.049, 0.238]$	$[0.037, 0.226]$	$[0.038, 0.234]$	$[0.048, 0.234]$
			p-value	0.00362	0.00649	0.00775	0.00257
rs11102986	RXRA	71224	β	0.163	0.109	0.128	0.086
			$[95\%CI]$	$[0.042, 0.271]$	$[0.010, 0.227]$	$[0.015, 0.236]$	$[-0.021, 0.202]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		45%	50%	55%	60%
			p-value	0.00556	4.70E-02	2.20E-02	1.32E-01
rs12603813	PLCD3	71116	β	0.09	0.115	0.129	0.158
			$[95\%CI]$	$[0.001, 0.203]$	$[0.017, 0.210]$	$[0.040, 0.225]$	$[0.064, 0.256]$
			p-value	0.08	0.019	0.00646	0.00129
rs6219	IGF1	73562	β	0.1	0.146	0.241	0.247
			$[95\%CI]$	$[-0.035, 0.238]$	$[0.006, 0.302]$	$[0.098, 0.377]$	$[0.108, 0.381]$
			p-value	1.51E-01	5.60E-02	7.06E-04	4.55E-04
rs2234693	ESR1	69772	β	0.079	0.066	0.1	0.098
			$[95\%CI]$	$[-0.015, 0.162]$	$[-0.020, 0.144]$	$[0.003, 0.176]$	$[0.006, 0.174]$
			p-value	0.084	0.122	0.024	0.022
rs46522	UBE2Z	64204	β	0.136	0.132	0.096	0.069
			$[95\%CI]$	$[0.058, 0.234]$	$[0.040, 0.207]$	$[0.007, 0.180]$	$[-0.021, 0.156]$
			p-value	0.00198	1.95E-03	0.03	0.125
rs891088	INSR	73562	β	0.116	0.1	0.11	0.08
			$[95\%CI]$	$[0.018, 0.206]$	$[0.017, 0.201]$	$[0.020, 0.204]$	$[-0.009, 0.177]$
			p-value	1.60E-02	0.032	1.70E-02	9.20E-02
rs9857730	VILL	73358	β	0.103	0.106	0.053	0.087
			$[95\%CI]$	$[-0.006, 0.206]$	$[0.010, 0.206]$	$[-0.029, 0.157]$	$[-0.028, 0.173]$
			p-value	0.056	0.037	0.267	0.092
rs10185680	MFSD2B	73232	β	0.13	0.138	0.131	0.131
			$[95\%CI]$	$[0.053, 0.221]$	$[0.065, 0.221]$	$[0.054, 0.215]$	$[0.043, 0.208]$
			p-value	0.00238	0.000518	0.00153	0.00179
rs7163907	PTPN9	73234	β	-0.100	-0.095	-0.096	-0.131
			$[95\%CI]$	$[-0.195, -0.004]$	$[-0.188, -0.002]$	$[-0.197, -0.014]$	$[-0.229, -0.031]$
			p-value	0.041	0.047	3.90E-02	9.33E-03
rs2176167	NOP58	73508	β	0.08	0.065	0.065	0.062
			$[95\%CI]$	$[-0.009, 0.170]$	$[-0.018, 0.157]$	$[-0.017, 0.154]$	$[-0.029, 0.146]$
			p-value	0.075	0.142	0.136	0.158
rs7557989	THADA	64417	β	0.096	0.092	0.088	0.094
			$[95\%CI]$	$[0.008, 0.189]$	$[0.005, 0.172]$	$[0.002, 0.183]$	$[0.000, 0.189]$
			p-value	0.036	0.031	0.054	0.049
rs510769	OPRM1	64413	β	0.136	0.174	0.168	0.187
			$[95\%CI]$	$[0.049, 0.241]$	$[0.062, 0.260]$	$[0.066, 0.265]$	$[0.093, 0.295]$
			p-value	4.74E-03	5.39E-04	9.27E-04	2.90E-04
rs7756224	NMBR	60624	β	0.139	0.138	0.132	0.125
			$[95\%CI]$	$[0.050, 0.233]$	$[0.044, 0.216]$	$[0.035, 0.216]$	$[0.027, 0.212]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		45%	50%	55%	60%
			p-value	0.003	0.001	0.005	0.008
rs2282537	POU2F3	64419	β	0.093	0.112	0.124	0.161
			$[95\%CI]$	$[-0.047, 0.212]$	$[-0.008, 0.236]$	$[0.006, 0.241]$	$[0.025, 0.283]$
			p-value	0.157	0.070	0.039	0.014
rs7853859	CENPP	69662	β	0.076	0.088	0.070	0.100
			$[95\%CI]$	$[-0.018, 0.161]$	$[-0.004, 0.173]$	$[-0.013, 0.165]$	$[0.012, 0.190]$
			p-value	0.096	0.051	0.118	0.027
rs2229642	ITPR3	52677	β	0.113	0.133	0.137	0.161
			$[95\%CI]$	$[0.010, 0.209]$	$[0.037, 0.223]$	$[0.036, 0.237]$	$[0.057, 0.263]$
			p-value	0.027	0.005	0.007	0.002
rs5015437	LMF1	64185	β	0.147	0.155	0.138	0.13
			$[95\%CI]$	$[0.061, 0.242]$	$[0.062, 0.241]$	$[0.047, 0.230]$	$[0.041, 0.218]$
			p-value	0.001	0.001	0.003	4.37E-03
rs7963565	KNTC1	73568	β	0.096	0.110	0.133	0.120
			$[95\%CI]$	$[0.008, 0.188]$	$[0.027, 0.196]$	$[0.040, 0.217]$	$[0.023, 0.202]$
			p-value	0.036	0.011	0.003	0.00814
rs696	NFKBIA	73461	β	0.072	0.094	0.125	0.154
			$[95\%CI]$	$[-0.009, 0.167]$	$[0.007, 0.172]$	$[0.039, 0.208]$	$[0.055, 0.234]$
			p-value	0.111	0.025	0.003	6.26E-04
rs25656	NFATC1	48455	β	0.100	0.100	0.123	0.100
			$[95\%CI]$	$[-0.005, 0.212]$	$[0.010, 0.202]$	$[0.015, 0.227]$	$[-0.016, 0.190]$
			p-value	0.070	0.039	0.023	0.060
rs3100776	IHH	64334	β	0.201	0.253	0.399	0.374
			$[95\%CI]$	$[-0.016, 0.426]$	$[0.016, 0.554]$	$[0.169, 0.641]$	$[0.134, 0.569]$
			p-value	0.079	0.062	6.89E-04	8.30E-04
rs12145922	PKN2	69346	β	0.107	0.080	0.041	0.063
			$[95\%CI]$	$[0.016, 0.191]$	$[-0.012, 0.161]$	$[-0.041, 0.134]$	$[-0.032, 0.133]$
			p-value	0.016	0.070	0.364	0.134
rs526134	USP37	60625	β	0.097	0.112	0.141	0.139
			$[95\%CI]$	$[0.006, 0.192]$	$[0.015, 0.194]$	$[0.034, 0.209]$	$[0.042, 0.225]$
			p-value	0.042	0.015	0.002	0.0025
rs7004280	RPS20	64415	β	0.318	0.323	0.346	0.360
			$[95\%CI]$	$[0.095, 0.551]$	$[0.088, 0.585]$	$[0.143, 0.605]$	$[0.094, 0.588]$
			p-value	0.006	0.012	0.00253	0.00435
rs4808199	GATAD2A	64415	β	0.064	0.098	0.113	0.146
			$[95\%CI]$	$[-0.039, 0.190]$	$[-0.015, 0.202]$	$[0.012, 0.230]$	$[0.038, 0.266]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		45%	50%	55%	60%
			p-value	0.274	0.080	0.041	0.011
rs3821009	PDE11A	69745	β	0.133	0.082	0.117	0.133
			$[95\%CI]$	$[-0.045, 0.257]$	$[-0.070, 0.212]$	$[-0.061, 0.264]$	$[-0.037, 0.283]$
			p-value	0.089	0.259	0.160	0.101
rs2291617	METTL1	73099	β	0.081	0.100	0.131	0.130
			$[95\%CI]$	$[-0.003, 0.182]$	$[0.019, 0.186]$	$[0.046, 0.215]$	$[0.038, 0.218]$
			p-value	0.086	0.019	0.0024	0.00455
rs1051168	SEC11A	64399	β	0.076	0.064	0.071	0.086
			$[95\%CI]$	$[-0.030, 0.166]$	$[-0.039, 0.148]$	$[-0.017, 0.172]$	$[-0.011, 0.183]$
			p-value	0.120	0.176	0.139	0.086
rs803932	ASTN2	64105	β	0.117	0.128	0.118	0.126
			$[95\%CI]$	$[0.016, 0.200]$	$[0.017, 0.195]$	$[0.027, 0.219]$	$[0.032, 0.213]$
			p-value	0.013	0.00449	0.015	0.00577
rs2247870	GPR98	69775	β	0.097	0.114	0.108	0.081
			$[95\%CI]$	$[0.013, 0.180]$	$[0.030, 0.194]$	$[0.026, 0.196]$	$[-0.006, 0.158]$
			p-value	0.026	0.00678	0.014	0.057
rs12503378	NUDT6	73569	β	0.102	0.098	0.065	0.07
			$[95\%CI]$	$[-0.016, 0.200]$	$[-0.012, 0.205]$	$[-0.035, 0.189]$	$[-0.043, 0.191]$
			p-value	0.068	0.077	0.259	0.242
rs2145923	NPR2	69777	β	0.075	0.065	0.067	0.083
			$[95\%CI]$	$[-0.030, 0.186]$	$[-0.039, 0.167]$	$[-0.041, 0.173]$	$[-0.046, 0.181]$
			p-value	0.177	0.221	0.22	0.15
rs12603582	ITGB3	73541	β	0.061	0.070	0.060	0.032
			$[95\%CI]$	$[-0.033, 0.173]$	$[-0.026, 0.168]$	$[-0.029, 0.161]$	$[-0.071, 0.137]$
			p-value	0.251	0.152	0.218	0.547
rs4864548	CLOCK	73568	β	0.071	0.074	0.102	0.085
			$[95\%CI]$	$[-0.012, 0.155]$	$[-0.003, 0.172]$	$[0.022, 0.192]$	$[-0.012, 0.158]$
			p-value	0.098	0.09	0.018	0.051
rs2735469	MRPL23	67563	β	0.047	0.046	0.071	0.075
			$[95\%CI]$	$[-0.078, 0.174]$	$[-0.069, 0.164]$	$[-0.058, 0.188]$	$[-0.049, 0.181]$
			p-value	0.453	0.44	0.258	0.208
rs1051431	MPHOSPH9	73561	β	0.000	0.028	0.009	0.023
			$[95\%CI]$	$[-0.089, 0.109]$	$[-0.075, 0.120]$	$[-0.082, 0.115]$	$[-0.071, 0.129]$
			p-value	1.000	0.574	0.864	0.652
rs41132	AP3B1	73441	β	-0.011	-0.012	0.004	0.07
			$[95\%CI]$	$[-0.110, 0.086]$	$[-0.101, 0.077]$	$[-0.076, 0.104]$	$[-0.036, 0.159]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		45%	50%	55%	60%
			p-value	0.829	0.796	0.926	0.166
rs6127698	MC3R	73486	β	0.040	0.068	0.043	0.057
			$[95\%CI]$	$[-0.036, 0.132]$	$[-0.006, 0.147]$	$[-0.040, 0.130]$	$[-0.031, 0.128]$
			p-value	0.348	0.082	0.314	0.167
rs1481892	ARNTL	73543	β	0.026	0.040	0.043	0.067
			$[95\%CI]$	$[-0.058, 0.118]$	$[-0.056, 0.122]$	$[-0.042, 0.139]$	$[-0.030, 0.152]$
			p-value	0.564	0.380	0.342	0.150
rs2633442	MKRN2	60624	β	0.011	0.034	0.040	0.010
			$[95\%CI]$	$[-0.082, 0.096]$	$[-0.056, 0.120]$	$[-0.039, 0.142]$	$[-0.075, 0.099]$
			p-value	0.813	0.446	0.388	0.816
rs1535	FADS1	73553	β	0.015	0.04	0.038	0.008
			$[95\%CI]$	$[-0.072, 0.103]$	$[-0.043, 0.119]$	$[-0.040, 0.131]$	$[-0.070, 0.104]$
			p-value	0.737	0.335	0.380	0.859
rs10861148	HSP90B1	73557	β	0.045	0.056	0.100	0.065
			$[95\%CI]$	$[-0.093, 0.173]$	$[-0.084, 0.192]$	$[-0.036, 0.222]$	$[-0.060, 0.176]$
			p-value	0.503	0.426	0.131	0.274
GS-Height		73570	β	0.174	0.176	0.177	0.176
			$[95\%CI]$	$[0.165, 0.182]$	$[0.168, 0.183]$	$[0.169, 0.185]$	$[0.169, 0.184]$
			p-value	<2.2E-308	<2.2E-308	<2.2E-308	<2.2E-308
			VarianceExplained	1.844%	1.86%	1.889%	1.919%

Part 4 - 65% to 80%

SNP	Gene	N		65%	70%	75%	80%
rs1042725	HMGA2	73105	β	0.526	0.559	0.572	0.556
			$[95\%CI]$	$[0.447, 0.617]$	$[0.467, 0.643]$	$[0.474, 0.646]$	$[0.453, 0.640]$
			p-value	2.41E-34	2.17E-35	6.27E-39	1.50E-31
rs2853977	HCP5	44699	β	0.596	0.621	0.56	0.515
			$[95\%CI]$	$[0.484, 0.700]$	$[0.513, 0.714]$	$[0.453, 0.671]$	$[0.402, 0.644]$
			p-value	1.40E-27	2.80E-34	2.50E-23	5.15E-17
rs3782415	SOCS2	73568	β	0.499	0.471	0.424	0.406
			$[95\%CI]$	$[0.393, 0.588]$	$[0.357, 0.569]$	$[0.313, 0.532]$	$[0.299, 0.514]$
			p-value	1.19E-24	4.27E-18	1.47E-14	1.75E-13
rs780094	GCKR	73548	β	0.386	0.397	0.36	0.326
			$[95\%CI]$	$[0.299, 0.466]$	$[0.293, 0.476]$	$[0.265, 0.442]$	$[0.243, 0.418]$
			p-value	1.37E-19	1.11E-17	1.97E-15	4.24E-13
rs9892365	TBX2	73566	β	0.379	0.4	0.374	0.387
			$[95\%CI]$	$[0.288, 0.466]$	$[0.312, 0.489]$	$[0.288, 0.465]$	$[0.290, 0.489]$
			p-value	8.06E-17	6.85E-19	7.25E-17	2.10E-14
rs7137534	PDE3A	73567	β	0.384	0.382	0.365	0.369
			$[95\%CI]$	$[0.295, 0.467]$	$[0.274, 0.464]$	$[0.276, 0.461]$	$[0.274, 0.471]$
			p-value	1.71E-18	1.76E-15	1.13E-14	1.33E-13
rs1776897	HMGA1	64379	β	0.632	0.609	0.628	0.588
			$[95\%CI]$	$[0.479, 0.777]$	$[0.475, 0.775]$	$[0.462, 0.786]$	$[0.437, 0.781]$
			p-value	1.62E-16	2.30E-15	5.12E-14	1.22E-11
rs572169	GHSR	73554	β	0.326	0.308	0.272	0.313
			$[95\%CI]$	$[0.224, 0.405]$	$[0.207, 0.392]$	$[0.173, 0.374]$	$[0.210, 0.406]$
			p-value	2.26E-12	7.26E-11	1.11E-07	6.40E-10
rs2679178	NPPC	73567	β	0.625	0.594	0.552	0.528
			$[95\%CI]$	$[0.460, 0.760]$	$[0.433, 0.748]$	$[0.407, 0.737]$	$[0.376, 0.709]$
			p-value	3.37E-16	1.28E-13	8.58E-11	4.83E-10
rs2053156	GRB2	73536	β	0.331	0.37	0.411	0.393
			$[95\%CI]$	$[0.235, 0.433]$	$[0.265, 0.471]$	$[0.302, 0.525]$	$[0.273, 0.505]$
			p-value	7.39E-11	1.43E-12	3.15E-13	2.34E-11
rs9930741	ERI2	73143	β	0.281	0.253	0.223	0.179
			$[95\%CI]$	$[0.182, 0.361]$	$[0.171, 0.341]$	$[0.137, 0.316]$	$[0.096, 0.273]$
			p-value	5.92E-10	5.50E-09	1.20E-06	7.87E-05
rs2854207	CSH2	67567	β	0.314	0.348	0.343	0.393
			$[95\%CI]$	$[0.220, 0.412]$	$[0.242, 0.449]$	$[0.238, 0.452]$	$[0.264, 0.494]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		65%	70%	75%	80%
			p-value	2.09E-10	4.50E-11	2.02E-10	2.29E-11
rs4320932	IGF2	71204	β	0.334	0.313	0.297	0.296
			$[95\%CI]$	$[0.224, 0.427]$	$[0.195, 0.425]$	$[0.183, 0.412]$	$[0.164, 0.400]$
			p-value	7.12E-11	9.26E-08	3.34E-07	6.44E-07
rs752313	EZH1	69774	β	0.229	0.254	0.232	0.24
			$[95\%CI]$	$[0.130, 0.303]$	$[0.161, 0.336]$	$[0.144, 0.322]$	$[0.154, 0.333]$
			p-value	2.07E-07	1.30E-08	4.32E-07	1.43E-07
rs709939	SAMD4A	73569	β	0.249	0.255	0.227	0.24
			$[95\%CI]$	$[0.171, 0.338]$	$[0.173, 0.342]$	$[0.141, 0.321]$	$[0.155, 0.338]$
			p-value	4.44E-09	3.12E-09	5.23E-07	2.21E-07
rs1036477	FBN1	73539	β	0.362	0.354	0.288	0.269
			$[95\%CI]$	$[0.234, 0.500]$	$[0.219, 0.489]$	$[0.159, 0.439]$	$[0.137, 0.430]$
			p-value	6.92E-08	2.37E-07	4.13E-05	3.18E-04
rs158676	CDK5RAP1	73562	β	0.22	0.229	0.258	0.248
			$[95\%CI]$	$[0.136, 0.316]$	$[0.150, 0.322]$	$[0.165, 0.359]$	$[0.161, 0.360]$
			p-value	1.58E-06	2.29E-07	1.70E-07	1.17E-06
rs1822469	PPP3R1	69710	β	0.225	0.223	0.209	0.231
			$[95\%CI]$	$[0.134, 0.311]$	$[0.134, 0.315]$	$[0.116, 0.296]$	$[0.135, 0.325]$
			p-value	6.13E-07	1.13E-06	5.52E-06	0.00000187
rs258281	RAB26	67411	β	0.233	0.236	0.258	0.27
			$[95\%CI]$	$[0.124, 0.333]$	$[0.130, 0.365]$	$[0.133, 0.381]$	$[0.141, 0.391]$
			p-value	1.06E-05	7.93E-05	3.67E-05	2.14E-05
rs9366637	HFE	73557	β	0.557	0.595	0.563	0.502
			$[95\%CI]$	$[0.396, 0.716]$	$[0.418, 0.784]$	$[0.380, 0.741]$	$[0.276, 0.705]$
			p-value	1.25E-11	2.31E-10	1.09E-09	3.50E-06
rs551219	COL24A1	73076	β	0.217	0.194	0.151	
			$[95\%CI]$	$[0.134, 0.315]$	$[0.108, 0.287]$	$[0.075, 0.258]$	
			p-value	2.38E-06	2.10E-05	1.05E-03	
rs13072536	ITIH4	73519	β	0.238	0.272	0.308	0.288
			$[95\%CI]$	$[0.148, 0.321]$	$[0.186, 0.378]$	$[0.196, 0.403]$	$[0.170, 0.385]$
			p-value	4.14E-08	2.02E-08	6.73E-09	1.43E-07
rs7522692	PIGC	64398	β	0.31	0.273	0.286	0.282
			$[95\%CI]$	$[0.192, 0.401]$	$[0.166, 0.387]$	$[0.161, 0.395]$	$[0.171, 0.403]$
			p-value	7.53E-09	9.12E-07	0.00000185	2.07E-06
rs13076290	CTNNB1	73154	β	0.195	0.2	0.237	0.238
			$[95\%CI]$	$[0.116, 0.281]$	$[0.117, 0.287]$	$[0.152, 0.317]$	$[0.145, 0.328]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		65%	70%	75%	80%
			p-value	3.37E-06	4.33E-06	1.60E-08	4.27E-07
rs1636255	GNA12	64371	β	0.279	0.282	0.284	0.249
			$[95\%CI]$	$[0.185, 0.380]$	$[0.194, 0.396]$	$[0.177, 0.374]$	$[0.146, 0.361]$
			p-value	3.10E-08	4.44E-08	1.35E-08	5.25E-06
rs3796529	REST	73119	β	0.21	0.248	0.211	0.222
			$[95\%CI]$	$[0.093, 0.310]$	$[0.134, 0.349]$	$[0.116, 0.343]$	$[0.120, 0.343]$
			p-value	1.54E-04	7.38E-06	2.92E-04	1.04E-04
rs1866146	POMC	73516	β	0.271	0.269	0.268	0.275
			$[95\%CI]$	$[0.190, 0.352]$	$[0.180, 0.353]$	$[0.181, 0.360]$	$[0.182, 0.374]$
			p-value	4.61E-11	6.11E-10	2.39E-09	1.50E-08
rs9844666	PCCB	73511	β	0.173	0.221	0.208	0.248
			$[95\%CI]$	$[0.083, 0.271]$	$[0.123, 0.321]$	$[0.108, 0.316]$	$[0.130, 0.349]$
			p-value	3.48E-04	1.15E-05	9.39E-05	8.41E-06
rs8071847	POLR2A	73565	β	0.222	0.213	0.195	0.234
			$[95\%CI]$	$[0.129, 0.326]$	$[0.113, 0.314]$	$[0.096, 0.299]$	$[0.122, 0.364]$
			p-value	7.83E-06	4.14E-05	1.89E-04	1.19E-04
rs3783937	FBLN5	73104	β	0.25	0.272	0.212	0.2
			$[95\%CI]$	$[0.150, 0.339]$	$[0.175, 0.371]$	$[0.123, 0.327]$	$[0.102, 0.306]$
			p-value	2.42E-07	3.85E-08	0.0000445	1.37E-04
rs11080149	NF1	73567	β	0.262	0.268	0.267	0.33
			$[95\%CI]$	$[0.131, 0.373]$	$[0.124, 0.400]$	$[0.139, 0.415]$	$[0.205, 0.473]$
			p-value	2.00E-05	1.31E-04	1.51E-04	8.35E-07
rs17472113	ZAR1	58807	β	0.284	0.261	0.298	0.325
			$[95\%CI]$	$[0.180, 0.385]$	$[0.148, 0.369]$	$[0.186, 0.411]$	$[0.209, 0.437]$
			p-value	4.31E-08	3.99E-06	2.04E-07	2.27E-08
rs490634	CISH	61124	β	0.314	0.32		0.314
			$[95\%CI]$	$[0.181, 0.447]$	$[0.178, 0.462]$		$[0.164, 0.465]$
			p-value	4.10E-06	5.87E-06		3.31E-05
rs17622208	SLC22A5	73531	β	0.207	0.222	0.25	0.204
			$[95\%CI]$	$[0.140, 0.295]$	$[0.133, 0.309]$	$[0.148, 0.334]$	$[0.107, 0.298]$
			p-value	1.73E-07	1.00E-06	1.33E-07	3.01E-05
rs2982712	ESR1	73566	β	0.16	0.191	0.212	0.232
			$[95\%CI]$	$[0.078, 0.239]$	$[0.104, 0.274]$	$[0.130, 0.302]$	$[0.141, 0.329]$
			p-value	9.98E-05	1.23E-05	1.18E-06	1.01E-06
rs1950500	NFATC4	64408	β	0.258	0.281	0.274	0.185
			$[95\%CI]$	$[0.166, 0.370]$	$[0.182, 0.377]$	$[0.158, 0.377]$	$[0.091, 0.303]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		65%	70%	75%	80%
			p-value	6.28E-07	9.11E-09	6.35E-07	5.78E-04
rs1476387	PPIL6	73563	β	0.132	0.145	0.152	0.158
			$[95\%CI]$	$[0.051, 0.214]$	$[0.065, 0.229]$	$[0.065, 0.239]$	$[0.064, 0.250]$
			p-value	1.25E-03	4.17E-04	6.51E-04	8.53E-04
rs4946932	FOXO3	73565	β	0.261	0.257	0.24	0.236
			$[95\%CI]$	$[0.173, 0.343]$	$[0.165, 0.347]$	$[0.148, 0.328]$	$[0.146, 0.333]$
			p-value	1.53E-09	3.19E-08	1.51E-07	8.37E-07
rs1800783	NOS3	71277	β	0.174	0.201		0.234
			$[95\%CI]$	$[0.094, 0.267]$	$[0.103, 0.290]$		$[0.145, 0.338]$
			p-value	7.99E-05	2.01E-05		2.07E-06
rs6718902	STAT1	73557	β	0.216	0.215	0.24	0.247
			$[95\%CI]$	$[0.119, 0.312]$	$[0.119, 0.321]$	$[0.138, 0.353]$	$[0.139, 0.355]$
			p-value	1.14E-05	2.22E-05	1.35E-05	6.60E-06
rs2425019	MMP24	69770	β	0.2	0.228	0.23	0.203
			$[95\%CI]$	$[0.115, 0.290]$	$[0.150, 0.323]$	$[0.137, 0.322]$	$[0.110, 0.312]$
			p-value	6.64E-06	2.68E-07	1.14E-06	7.83E-05
rs6731022	EIF2AK3	69767	β	0.234	0.254	0.253	0.274
			$[95\%CI]$	$[0.145, 0.323]$	$[0.157, 0.347]$	$[0.165, 0.367]$	$[0.177, 0.386]$
			p-value	2.45E-07	2.14E-07	7.20E-07	2.29E-07
rs12940055	MAP3K3	73566	β	0.349	0.372	0.336	0.315
			$[95\%CI]$	$[0.211, 0.462]$	$[0.242, 0.510]$	$[0.193, 0.491]$	$[0.172, 0.453]$
			p-value	5.68E-08	3.75E-08	9.59E-06	1.04E-05
rs864745	JAZF1	60616	β	0.191	0.178	0.213	0.212
			$[95\%CI]$	$[0.097, 0.269]$	$[0.103, 0.283]$	$[0.112, 0.311]$	$[0.120, 0.325]$
			p-value	1.18E-05	1.22E-04	3.38E-05	5.62E-05
rs6487088	PDE3A	73563	β	0.224	0.241	0.233	0.267
			$[95\%CI]$	$[0.128, 0.317]$	$[0.136, 0.348]$	$[0.123, 0.348]$	$[0.153, 0.377]$
			p-value	4.50E-06	7.56E-06	5.13E-05	3.25E-06
rs4973410	NCL	64408	β	0.191	0.241	0.242	0.23
			$[95\%CI]$	$[0.112, 0.279]$	$[0.132, 0.318]$	$[0.147, 0.337]$	$[0.141, 0.325]$
			p-value	1.14E-05	3.47E-07	6.45E-07	8.84E-07
rs451061	PRKCZ	71365	β	0.17	0.18	0.182	0.162
			$[95\%CI]$	$[0.090, 0.252]$	$[0.094, 0.270]$	$[0.087, 0.271]$	$[0.071, 0.258]$
			p-value	3.60E-05	5.89E-05	1.03E-04	6.80E-04
rs832575	MAP3K1	73539	β	0.234	0.269	0.267	0.258
			$[95\%CI]$	$[0.109, 0.361]$	$[0.138, 0.396]$	$[0.131, 0.405]$	$[0.116, 0.397]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		65%	70%	75%	80%
			p-value	2.60E-04	2.86E-05	1.28E-04	3.02E-04
rs4955526	EPHB1	64417	β	0.227	0.27	0.231	0.195
			$[95\%CI]$	$[0.143, 0.322]$	$[0.170, 0.351]$	$[0.143, 0.334]$	$[0.100, 0.290]$
			p-value	6.43E-07	4.83E-09	2.68E-06	7.38E-05
rs8038415	IGF1R	73516	β	0.181	0.175	0.196	0.178
			$[95\%CI]$	$[0.100, 0.264]$	$[0.101, 0.272]$	$[0.102, 0.287]$	$[0.087, 0.266]$
			p-value	1.25E-05	6.66E-05	3.41E-05	8.83E-05
rs7578199	HDLBP	73569	β	0.188	0.223	0.234	0.228
			$[95\%CI]$	$[0.093, 0.281]$	$[0.129, 0.322]$	$[0.141, 0.349]$	$[0.127, 0.340]$
			p-value	8.29E-05	4.92E-06	9.69E-06	0.0000276
rs7020782	PAPPA	73566	β	0.176	0.182	0.162	0.17
			$[95\%CI]$	$[0.089, 0.265]$	$[0.094, 0.271]$	$[0.076, 0.274]$	$[0.072, 0.268]$
			p-value	9.34E-05	6.21E-05	1.40E-03	6.52E-04
rs2229712	RPS6KA1	48240	β	0.229	0.268	0.233	0.229
			$[95\%CI]$	$[0.123, 0.354]$	$[0.132, 0.380]$	$[0.102, 0.357]$	$[0.089, 0.357]$
			p-value	8.75E-05	1.96E-05	3.19E-04	8.32E-04
rs7572476	BOK	71367	β	0.231	0.262	0.237	0.18
			$[95\%CI]$	$[0.147, 0.308]$	$[0.171, 0.335]$	$[0.149, 0.331]$	$[0.092, 0.271]$
			p-value	1.85E-08	3.48E-10	2.62E-07	9.16E-05
rs2066807	PAN2	71822	β	0.262	0.287	0.253	0.261
			$[95\%CI]$	$[0.165, 0.367]$	$[0.156, 0.394]$	$[0.117, 0.353]$	$[0.128, 0.382]$
			p-value	4.44E-07	2.21E-06	1.98E-05	0.000066
rs1516796	ACAN	69308	β	0.185	0.178	0.155	0.165
			$[95\%CI]$	$[0.095, 0.273]$	$[0.096, 0.266]$	$[0.073, 0.258]$	$[0.070, 0.261]$
			p-value	4.05E-05	4.75E-05	9.77E-04	6.91E-04
rs6180	GHR	73552	β	0.162	0.136	0.11	0.134
			$[95\%CI]$	$[0.078, 0.237]$	$[0.049, 0.217]$	$[0.030, 0.208]$	$[0.049, 0.219]$
			p-value	7.39E-05	1.51E-03	1.40E-02	1.99E-03
rs8055190	LRRC36	64416	β	0.475	0.557	0.457	0.54
			$[95\%CI]$	$[0.297, 0.722]$	$[0.280, 0.737]$	$[0.220, 0.700]$	$[0.254, 0.709]$
			p-value	1.57E-05	1.89E-06	1.83E-04	3.49E-06
rs17106235	FAF1	60489	β	0.357	0.321	0.365	0.375
			$[95\%CI]$	$[0.192, 0.481]$	$[0.183, 0.503]$	$[0.221, 0.563]$	$[0.212, 0.550]$
			p-value	9.38E-07	6.13E-05	3.19E-05	0.000012
rs3739707	LPAR1	73568	β	0.174	0.169	0.134	0.137
			$[95\%CI]$	$[0.076, 0.258]$	$[0.060, 0.259]$	$[0.040, 0.245]$	$[0.039, 0.249]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		65%	70%	75%	80%
			p-value	1.73E-04	8.29E-04	1.10E-02	1.00E-02
rs674424	ABCG4	73568	β	0.238	0.188	0.167	0.139
			$[95\%CI]$	$[0.139, 0.327]$	$[0.085, 0.281]$	$[0.077, 0.281]$	$[0.048, 0.257]$
			p-value	4.18E-07	1.88E-04	1.36E-03	9.21E-03
rs12225387	NEU3	60621	β	0.244	0.222	0.217	0.203
			$[95\%CI]$	$[0.144, 0.343]$	$[0.119, 0.321]$	$[0.113, 0.346]$	$[0.098, 0.318]$
			p-value	1.51E-06	1.57E-05	3.01E-04	0.000343
rs3812265	CNOT4	69776	β	0.216	0.237	0.182	0.255
			$[95\%CI]$	$[0.104, 0.302]$	$[0.126, 0.332]$	$[0.082, 0.294]$	$[0.152, 0.368]$
			p-value	2.19E-05	5.28E-06	0.000643	0.00000382
rs10208728	IHH	64409	β	0.273	0.271	0.215	0.233
			$[95\%CI]$	$[0.141, 0.411]$	$[0.119, 0.416]$	$[0.058, 0.376]$	$[0.079, 0.395]$
			p-value	0.0000568	2.44E-04	8.25E-03	3.63E-03
rs291979	GRK5	73568	β	0.207	0.223	0.2	0.134
			$[95\%CI]$	$[0.092, 0.302]$	$[0.121, 0.322]$	$[0.092, 0.294]$	$[0.050, 0.248]$
			p-value	1.16E-04	1.26E-05	9.81E-05	7.86E-03
rs2715553	RARA	69206	β	0.115	0.1	0.089	0.132
			$[95\%CI]$	$[0.028, 0.192]$	$[0.016, 0.188]$	$[0.006, 0.195]$	$[0.038, 0.228]$
			p-value	0.006	0.023	0.064	0.0059
rs2057291	GNAS	70100	β	0.201	0.201	0.192	0.131
			$[95\%CI]$	$[0.110, 0.283]$	$[0.115, 0.300]$	$[0.102, 0.292]$	$[0.039, 0.239]$
			p-value	5.29E-06	2.03E-05	6.52E-05	9.23E-03
rs4803520	GRIK5	69747	β	0.167	0.24	0.227	0.174
			$[95\%CI]$	$[0.023, 0.294]$	$[0.109, 0.364]$	$[0.063, 0.371]$	$[0.029, 0.319]$
			p-value	0.016	0.000178	0.00412	0.019
rs10736682	APLNR	73090	β	0.186	0.159	0.135	0.089
			$[95\%CI]$	$[0.110, 0.266]$	$[0.083, 0.246]$	$[0.042, 0.214]$	$[-0.003, 0.186]$
			p-value	3.01E-06	1.35E-04	1.85E-03	6.40E-02
rs2909430	TP53	73412	β	0.155	0.186	0.186	0.134
			$[95\%CI]$	$[0.033, 0.274]$	$[0.055, 0.314]$	$[0.042, 0.324]$	$[0.010, 0.270]$
			p-value	0.012	5.00E-03	1.10E-02	4.60E-02
rs12050767	CYP19A1	73563	β	0.153	0.133	0.08	0.054
			$[95\%CI]$	$[0.068, 0.226]$	$[0.050, 0.204]$	$[-0.001, 0.178]$	$[-0.032, 0.156]$
			p-value	1.21E-04	7.00E-04	8.30E-02	2.64E-01
rs602633	PSRC1	64375	β	0.146	0.146	0.171	0.178
			$[95\%CI]$	$[0.054, 0.252]$	$[0.048, 0.265]$	$[0.082, 0.299]$	$[0.043, 0.288]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		65%	70%	75%	80%
			p-value	3.92E-03	8.08E-03	0.00189	0.00518
rs1738475	HTR1D	64411	β	0.091	0.131	0.129	
			$[95\%CI]$	$[0.003, 0.180]$	$[0.033, 0.207]$	$[0.040, 0.217]$	
			p-value	4.50E-02	3.14E-03	4.49E-03	
rs17754	RFC1	73558	β	0.154	0.202	0.199	0.162
			$[95\%CI]$	$[0.076, 0.238]$	$[0.121, 0.287]$	$[0.102, 0.279]$	$[0.076, 0.254]$
			p-value	1.80E-04	2.23E-06	1.02E-05	3.19E-04
rs17541471	NPR3	69777	β	0.167	0.226	0.204	0.206
			$[95\%CI]$	$[0.082, 0.288]$	$[0.112, 0.334]$	$[0.105, 0.340]$	$[0.110, 0.338]$
			p-value	1.80E-03	5.01E-05	0.000607	0.00038
rs1342586	TGFB2	69745	β	0.1	0.13	0.119	0.104
			$[95\%CI]$	$[0.020, 0.216]$	$[0.025, 0.228]$	$[0.009, 0.218]$	$[-0.005, 0.212]$
			p-value	0.043	0.012	0.026	0.063
rs3814115	PCSK5	73537	β	0.138	0.153	0.178	0.201
			$[95\%CI]$	$[0.042, 0.223]$	$[0.067, 0.255]$	$[0.096, 0.284]$	$[0.096, 0.301]$
			p-value	2.99E-03	1.44E-03	0.000246	0.000135
rs1780616	LBP	73560	β	0.118	0.114	0.123	0.106
			$[95\%CI]$	$[0.035, 0.195]$	$[0.030, 0.201]$	$[0.015, 0.207]$	$[0.013, 0.200]$
			p-value	0.00345	0.00858	0.01	0.026
rs3736228	LRP5	73508	β	0.112	0.131	0.134	0.134
			$[95\%CI]$	$[-0.006, 0.218]$	$[0.004, 0.240]$	$[0.009, 0.253]$	$[0.020, 0.274]$
			p-value	5.00E-02	2.70E-02	3.20E-02	4.00E-02
rs212517	ECE1	71344	β	0.176	0.175	0.167	0.139
			$[95\%CI]$	$[0.093, 0.259]$	$[0.090, 0.258]$	$[0.071, 0.254]$	$[0.056, 0.245]$
			p-value	2.98E-05	4.95E-05	3.14E-04	0.00371
rs7359336	NFAT5	73094	β	0.098	0.133	0.143	0.135
			$[95\%CI]$	$[0.009, 0.167]$	$[0.039, 0.198]$	$[0.061, 0.245]$	$[0.043, 0.220]$
			p-value	1.50E-02	9.95E-04	2.16E-03	2.64E-03
rs2682552	XRCC1	73106	β	0.135	0.174		0.143
			$[95\%CI]$	$[0.036, 0.247]$	$[0.062, 0.273]$		$[0.055, 0.272]$
			p-value	1.20E-02	0.00116		0.00919
rs17085675	PCSK1	69775	β	0.132	0.117	0.082	0.033
			$[95\%CI]$	$[0.047, 0.214]$	$[0.019, 0.205]$	$[-0.022, 0.167]$	$[-0.073, 0.125]$
			p-value	0.00174	0.013	0.085	0.511
rs11102986	RXRA	71224	β	0.09	0.095	0.1	0.087
			$[95\%CI]$	$[-0.022, 0.191]$	$[-0.018, 0.203]$	$[-0.019, 0.216]$	$[-0.028, 0.206]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		65%	70%	75%	80%
			p-value	0.098	9.00E-02	0.094	0.146
rs12603813	PLCD3	71116	β	0.143	0.164	0.153	0.157
			$[95\%CI]$	$[0.056, 0.244]$	$[0.067, 0.263]$	$[0.067, 0.269]$	$[0.056, 0.269]$
			p-value	0.00266	0.000917	0.00305	0.00356
rs6219	IGF1	73562	β	0.299	0.307	0.315	0.363
			$[95\%CI]$	$[0.137, 0.436]$	$[0.176, 0.447]$	$[0.166, 0.446]$	$[0.202, 0.530]$
			p-value	8.34E-05	1.08E-05	9.60E-06	1.70E-05
rs2234693	ESR1	69772	β	0.1	0.128	0.109	0.091
			$[95\%CI]$	$[0.015, 0.184]$	$[0.035, 0.207]$	$[0.018, 0.202]$	$[0.004, 0.188]$
			p-value	0.02	3.19E-03	2.20E-02	5.40E-02
rs46522	UBE2Z	64204	β	0.043	0.062	0.071	0.067
			$[95\%CI]$	$[-0.044, 0.136]$	$[-0.036, 0.139]$	$[-0.022, 0.163]$	$[-0.027, 0.165]$
			p-value	3.46E-01	1.71E-01	1.33E-01	1.73E-01
rs891088	INSR	73562	β	0.088	0.126	0.134	0.154
			$[95\%CI]$	$[0.001, 0.193]$	$[0.030, 0.213]$	$[0.032, 0.221]$	$[0.048, 0.261]$
			p-value	7.20E-02	7.56E-03	0.00505	0.00469
rs9857730	VILL	73358	β	0.091	0.054	0.042	0.062
			$[95\%CI]$	$[-0.003, 0.189]$	$[-0.043, 0.160]$	$[-0.061, 0.139]$	$[-0.055, 0.171]$
			p-value	0.063	0.295	0.409	0.282
rs10185680	MFSD2B	73232	β	0.112	0.135	0.158	0.122
			$[95\%CI]$	$[0.034, 0.185]$	$[0.056, 0.223]$	$[0.068, 0.238]$	$[0.024, 0.209]$
			p-value	0.00326	1.39E-03	0.000259	0.00959
rs7163907	PTPN9	73234	β	-0.095	-0.126	-0.106	-0.064
			$[95\%CI]$	$[-0.198, -0.006]$	$[-0.211, -0.014]$	$[-0.204, -0.007]$	$[-0.179, 0.029]$
			p-value	5.10E-02	1.20E-02	0.034	0.224
rs2176167	NOP58	73508	β	0.064	0.100	0.081	0.075
			$[95\%CI]$	$[-0.014, 0.156]$	$[0.015, 0.187]$	$[-0.007, 0.180]$	$[-0.010, 0.178]$
			p-value	0.138	0.023	0.089	0.118
rs7557989	THADA	64417	β	0.088	0.053	0.070	0.067
			$[95\%CI]$	$[-0.008, 0.175]$	$[-0.047, 0.151]$	$[-0.030, 0.166]$	$[-0.037, 0.162]$
			p-value	0.058	0.298	0.164	0.182
rs510769	OPRM1	64413	β	0.136	0.156	0.141	0.120
			$[95\%CI]$	$[0.044, 0.237]$	$[0.049, 0.258]$	$[0.026, 0.240]$	$[0.014, 0.225]$
			p-value	0.005	0.003	0.010	0.025
rs7756224	NMBR	60624	β	0.138	0.129	0.145	0.131
			$[95\%CI]$	$[0.044, 0.226]$	$[0.034, 0.215]$	$[0.046, 0.258]$	$[0.022, 0.219]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		65%	70%	75%	80%
			p-value	0.003	0.005	0.007	0.008
rs2282537	POU2F3	64419	β	0.158	0.143	0.142	0.237
			$[95\%CI]$	$[0.038, 0.283]$	$[0.002, 0.263]$	$[0.015, 0.260]$	$[0.113, 0.370]$
			p-value	0.011	0.034	0.022	0.000
rs7853859	CENPP	69662	β	0.102	0.128	0.135	
			$[95\%CI]$	$[0.013, 0.179]$	$[0.037, 0.211]$	$[0.036, 0.230]$	
			p-value	0.015	0.004	0.005	
rs2229642	ITPR3	52677	β	0.161	0.131		0.143
			$[95\%CI]$	$[0.065, 0.259]$	$[0.037, 0.239]$		$[0.038, 0.264]$
			p-value	0.001	0.011		0.012
rs5015437	LMF1	64185	β	0.076	0.071	0.065	0.056
			$[95\%CI]$	$[-0.012, 0.180]$	$[-0.030, 0.169]$	$[-0.036, 0.156]$	$[-0.035, 0.152]$
			p-value	0.124	0.162	0.186	0.244
rs7963565	KNTC1	73568	β	0.108	0.134	0.140	0.164
			$[95\%CI]$	$[0.019, 0.191]$	$[0.050, 0.221]$	$[0.045, 0.242]$	$[0.066, 0.256]$
			p-value	0.014	0.002	0.00552	7.48E-04
rs696	NFKBIA	73461	β	0.131	0.136	0.148	0.125
			$[95\%CI]$	$[0.051, 0.215]$	$[0.053, 0.232]$	$[0.057, 0.235]$	$[0.031, 0.225]$
			p-value	1.50E-03	0.003	0.001	0.012
rs25656	NFATC1	48455	β	0.084	0.074	0.064	0.057
			$[95\%CI]$	$[-0.013, 0.191]$	$[-0.022, 0.194]$	$[-0.044, 0.188]$	$[-0.059, 0.182]$
			p-value	0.107	0.184	0.282	0.350
rs3100776	IHH	64334	β	0.311	0.424	0.423	0.447
			$[95\%CI]$	$[0.103, 0.580]$	$[0.164, 0.653]$	$[0.159, 0.655]$	$[0.202, 0.727]$
			p-value	0.010	7.23E-04	7.54E-04	8.43E-04
rs12145922	PKN2	69346	β	0.064	0.074	0.052	0.062
			$[95\%CI]$	$[-0.022, 0.152]$	$[-0.018, 0.157]$	$[-0.029, 0.146]$	$[-0.027, 0.158]$
			p-value	0.148	0.100	0.248	0.193
rs526134	USP37	60625	β	0.131	0.122	0.097	0.139
			$[95\%CI]$	$[0.031, 0.219]$	$[0.022, 0.217]$	$[0.008, 0.207]$	$[0.024, 0.237]$
			p-value	0.005	0.013	0.052	0.00953
rs7004280	RPS20	64415	β	0.340	0.281	0.155	0.200
			$[95\%CI]$	$[0.121, 0.571]$	$[0.047, 0.478]$	$[-0.033, 0.480]$	$[-0.041, 0.481]$
			p-value	0.00274	0.011	0.228	0.133
rs4808199	GATAD2A	64415	β	0.142	0.143	0.100	0.130
			$[95\%CI]$	$[0.027, 0.256]$	$[0.022, 0.262]$	$[-0.011, 0.225]$	$[0.022, 0.263]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		65%	70%	75%	80%
			p-value	0.015	0.020	0.095	0.035
rs3821009	PDE11A	69745	β	0.103	0.123	0.137	0.218
			$[95\%CI]$	$[-0.041, 0.265]$	$[-0.045, 0.288]$	$[-0.006, 0.313]$	$[0.045, 0.433]$
			p-value	0.180	0.149	0.090	0.027
rs2291617	METTL1	73099	β	0.125	0.148	0.103	0.067
			$[95\%CI]$	$[0.034, 0.208]$	$[0.053, 0.231]$	$[0.008, 0.198]$	$[-0.028, 0.165]$
			p-value	0.0046	0.00102	0.034	0.181
rs1051168	SEC11A	64399	β	0.064	0.064	0.071	0.120
			$[95\%CI]$	$[-0.029, 0.158]$	$[-0.044, 0.154]$	$[-0.018, 0.186]$	$[0.021, 0.209]$
			p-value	0.181	0.204	0.174	0.013
rs803932	ASTN2	64105	β	0.091	0.054	0.018	0.003
			$[95\%CI]$	$[-0.004, 0.174]$	$[-0.043, 0.146]$	$[-0.065, 0.111]$	$[-0.097, 0.111]$
			p-value	0.042	0.252	0.695	9.49E-01
rs2247870	GPR98	69775	β	0.111	0.108	0.092	0.073
			$[95\%CI]$	$[0.021, 0.183]$	$[0.018, 0.191]$	$[0.012, 0.192]$	$[-0.023, 0.157]$
			p-value	0.00724	0.016	0.046	0.113
rs12503378	NUDT6	73569	β	0.076	0.035	0.055	0.064
			$[95\%CI]$	$[-0.041, 0.172]$	$[-0.077, 0.158]$	$[-0.063, 0.168]$	$[-0.061, 0.180]$
			p-value	0.161	0.558	0.346	0.290
rs2145923	NPR2	69777	β	0.069	0.095	0.133	0.120
			$[95\%CI]$	$[-0.015, 0.184]$	$[-0.013, 0.214]$	$[0.008, 0.241]$	$[0.001, 0.230]$
			p-value	0.167	0.101	0.024	0.038
rs12603582	ITGB3	73541	β	0.022	0.035	0.035	0.012
			$[95\%CI]$	$[-0.077, 0.129]$	$[-0.073, 0.134]$	$[-0.075, 0.137]$	$[-0.098, 0.115]$
			p-value	0.677	0.517	0.519	0.825
rs4864548	CLOCK	73568	β	0.052	0.065	0.060	0.064
			$[95\%CI]$	$[-0.032, 0.134]$	$[-0.024, 0.151]$	$[-0.021, 0.162]$	$[-0.025, 0.166]$
			p-value	0.228	0.148	0.200	0.189
rs2735469	MRPL23	67563	β	0.098	0.137	0.154	0.151
			$[95\%CI]$	$[-0.020, 0.208]$	$[0.023, 0.256]$	$[0.020, 0.291]$	$[0.027, 0.274]$
			p-value	0.091	0.021	0.026	1.60E-02
rs1051431	MPHOSPH9	73561	β	0.073	0.08	0.091	0.067
			$[95\%CI]$	$[-0.035, 0.166]$	$[-0.020, 0.182]$	$[-0.015, 0.190]$	$[-0.032, 0.183]$
			p-value	0.158	0.117	0.079	0.221
rs41132	AP3B1	73441	β	0.057	0.072	0.093	0.072
			$[95\%CI]$	$[-0.034, 0.155]$	$[-0.039, 0.171]$	$[-0.020, 0.187]$	$[-0.020, 0.177]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		65%	70%	75%	80%
			p-value	0.246	0.174	0.079	0.155
rs6127698	MC3R	73486	β	0.043	0.006	0.000	0.026
			$[95\%CI]$	$[-0.040, 0.119]$	$[-0.072, 0.095]$	$[-0.084, 0.084]$	$[-0.059, 0.114]$
			p-value	0.287	0.894	1.000	0.555
rs1481892	ARNTL	73543	β	0.081	0.096	0.060	0.068
			$[95\%CI]$	$[-0.007, 0.169]$	$[-0.002, 0.179]$	$[-0.035, 0.154]$	$[-0.034, 0.151]$
			p-value	0.071	0.039	0.213	0.151
rs2633442	MKRN2	60624	β	0.030	0.067	0.020	0.009
			$[95\%CI]$	$[-0.055, 0.116]$	$[-0.034, 0.149]$	$[-0.064, 0.124]$	$[-0.082, 0.107]$
			p-value	0.489	0.153	0.675	0.852
rs1535	FADS1	73553	β	0.008	-0.005	-0.025	-0.04
			$[95\%CI]$	$[-0.074, 0.102]$	$[-0.098, 0.081]$	$[-0.120, 0.066]$	$[-0.144, 0.046]$
			p-value	0.861	0.915	0.602	0.412
rs10861148	HSP90B1	73557	β	0.041	-0.017	-0.017	0.007
			$[95\%CI]$	$[-0.088, 0.151]$	$[-0.157, 0.122]$	$[-0.160, 0.108]$	$[-0.158, 0.147]$
			p-value	0.500	0.811	0.799	0.932
GS-Height		73570	β	0.178	0.177	0.18	0.178
			$[95\%CI]$	$[0.171, 0.187]$	$[0.169, 0.185]$	$[0.170, 0.188]$	$[0.169, 0.187]$
			p-value	<2.2E-308	<2.2E-308	3.65e-314	7.79E-306
			VarianceExplained	1.935%	1.934%	1.909%	1.847%

Part 4 - 85% to 95% and OLS estimates

SNP	Gene	N		85%	90%	95%	OLS
rs1042725	HMGA2	73105	β	0.549	0.564	0.485	0.565
			$[95\%CI]$	$[0.440, 0.649]$	$[0.445, 0.676]$	$[0.345, 0.633]$	$[0.500, 0.631]$
			p-value	5.45E-25	6.94E-22	2.78E-11	6.56E-64
rs2853977	HCP5	44699	β	0.544	0.579	0.488	0.636
			$[95\%CI]$	$[0.416, 0.662]$	$[0.410, 0.729]$	$[0.316, 0.656]$	$[0.554, 0.717]$
			p-value	7.39E-18	1.03E-12	1.89E-08	1.01E-52
rs3782415	SOCS2	73568	β	0.445	0.479	0.487	0.464
			$[95\%CI]$	$[0.310, 0.556]$	$[0.341, 0.608]$	$[0.313, 0.666]$	$[0.383, 0.545]$
			p-value	1.22E-12	2.87E-12	6.04E-08	4.01E-29
rs780094	GCKR	73548	β	0.356	0.423	0.411	0.372
			$[95\%CI]$	$[0.266, 0.474]$	$[0.298, 0.546]$	$[0.257, 0.533]$	$[0.306, 0.439]$
			p-value	2.16E-11	2.15E-11	5.17E-09	8.74E-28
rs9892365	TBX2	73566	β	0.393	0.395	0.4	0.356
			$[95\%CI]$	$[0.286, 0.498]$	$[0.253, 0.505]$	$[0.245, 0.562]$	$[0.286, 0.426]$
			p-value	2.90E-13	7.53E-10	5.91E-07	2.28E-23
rs7137534	PDE3A	73567	β	0.377	0.395	0.399	0.352
			$[95\%CI]$	$[0.256, 0.479]$	$[0.260, 0.499]$	$[0.252, 0.572]$	$[0.282, 0.421]$
			p-value	2.98E-11	6.42E-11	7.66E-07	6.03E-23
rs1776897	HMGA1	64379	β	0.7	0.807	0.8	0.61
			$[95\%CI]$	$[0.514, 0.887]$	$[0.588, 1.057]$	$[0.518, 1.087]$	$[0.488, 0.732]$
			p-value	5.70E-14	5.87E-12	4.88E-08	1.15E-22
rs572169	GHSR	73554	β	0.335	0.347	0.345	0.351
			$[95\%CI]$	$[0.234, 0.445]$	$[0.231, 0.465]$	$[0.210, 0.508]$	$[0.280, 0.422]$
			p-value	4.89E-10	4.64E-09	4.86E-06	3.14E-22
rs2679178	NPPC	73567	β	0.537	0.521	0.489	0.527
			$[95\%CI]$	$[0.365, 0.719]$	$[0.328, 0.766]$	$[0.200, 0.697]$	$[0.410, 0.644]$
			p-value	4.87E-09	2.46E-06	8.46E-05	1.17E-18
rs2053156	GRB2	73536	β	0.429	0.415	0.389	0.371
			$[95\%CI]$	$[0.294, 0.566]$	$[0.233, 0.532]$	$[0.239, 0.601]$	$[0.286, 0.456]$
			p-value	6.06E-10	3.64E-08	0.0000246	1.15E-17
rs9930741	ERI2	73143	β	0.163	0.17	0.224	0.285
			$[95\%CI]$	$[0.065, 0.270]$	$[0.056, 0.294]$	$[0.083, 0.363]$	$[0.219, 0.352]$
			p-value	1.87E-03	5.16E-03	1.63E-03	3.92E-17
rs2854207	CSH2	67567	β	0.423	0.427	0.414	0.314
			$[95\%CI]$	$[0.312, 0.550]$	$[0.275, 0.571]$	$[0.251, 0.597]$	$[0.237, 0.392]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		85%	90%	95%	OLS
			p-value	1.75E-12	2.18E-08	2.27E-06	2.11E-15
rs4320932	IGF2	71204	β	0.302	0.414	0.293	0.336
			$[95\%CI]$	$[0.178, 0.422]$	$[0.220, 0.508]$	$[0.117, 0.466]$	$[0.252, 0.419]$
			p-value	1.24E-06	1.24E-08	9.98E-04	4.11E-15
rs752313	EZH1	69774	β	0.243	0.278	0.33	0.258
			$[95\%CI]$	$[0.144, 0.335]$	$[0.146, 0.386]$	$[0.138, 0.456]$	$[0.191, 0.326]$
			p-value	7.05E-07	6.99E-06	6.24E-05	4.78E-14
rs709939	SAMD4A	73569	β	0.276	0.285	0.367	0.248
			$[95\%CI]$	$[0.165, 0.363]$	$[0.135, 0.391]$	$[0.203, 0.493]$	$[0.181, 0.314]$
			p-value	3.86E-08	1.56E-05	0.000000797	2.37E-13
rs1036477	FBN1	73539	β	0.268	0.292	0.292	0.377
			$[95\%CI]$	$[0.091, 0.417]$	$[0.108, 0.490]$	$[0.032, 0.511]$	$[0.272, 0.483]$
			p-value	1.17E-03	2.69E-03	0.016	2.34E-12
rs158676	CDK5RAP1	73562	β	0.193	0.282	0.375	0.254
			$[95\%CI]$	$[0.101, 0.307]$	$[0.139, 0.402]$	$[0.222, 0.512]$	$[0.183, 0.325]$
			p-value	2.47E-04	2.32E-05	0.000000341	2.39E-12
rs1822469	PPP3R1	69710	β	0.212	0.151	0.256	0.243
			$[95\%CI]$	$[0.100, 0.311]$	$[0.029, 0.279]$	$[0.114, 0.405]$	$[0.175, 0.312]$
			p-value	0.0000769	0.02	0.000554	4.23E-12
rs258281	RAB26	67411	β	0.279	0.233	0.149	0.308
			$[95\%CI]$	$[0.150, 0.419]$	$[0.080, 0.383]$	$[-0.073, 0.328]$	$[0.220, 0.397]$
			p-value	5.47E-05	2.51E-03	0.143	1.06E-11
rs9366637	HFE	73557	β	0.407	0.298	0.218	0.47
			$[95\%CI]$	$[0.174, 0.612]$	$[0.018, 0.550]$	$[-0.116, 0.478]$	$[0.334, 0.606]$
			p-value	3.10E-04	2.80E-02	0.139	1.22E-11
rs551219	COL24A1	73076	β	0.137	0.13	0.103	0.249
			$[95\%CI]$	$[0.027, 0.247]$	$[0.005, 0.262]$	$[-0.060, 0.279]$	$[0.177, 0.321]$
			p-value	1.30E-02	0.046	0.229	1.30E-11
rs13072536	ITIH4	73519	β	0.3	0.288	0.184	0.265
			$[95\%CI]$	$[0.182, 0.429]$	$[0.146, 0.448]$	$[0.027, 0.347]$	$[0.188, 0.342]$
			p-value	2.13E-06	1.64E-04	0.025	1.71E-11
rs7522692	PIGC	64398	β	0.294	0.344	0.295	0.29
			$[95\%CI]$	$[0.167, 0.425]$	$[0.202, 0.479]$	$[0.090, 0.449]$	$[0.205, 0.375]$
			p-value	9.53E-06	1.06E-06	0.00134	2.24E-11
rs13076290	CTNNB1	73154	β	0.248	0.268	0.28	0.225
			$[95\%CI]$	$[0.139, 0.335]$	$[0.134, 0.363]$	$[0.103, 0.408]$	$[0.158, 0.291]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		85%	90%	95%	OLS
			p-value	8.82E-07	4.94E-06	2.80E-04	3.23E-11
rs1636255	GNA12	64371	β	0.252	0.277	0.352	0.261
			$[95\%CI]$	$[0.134, 0.365]$	$[0.117, 0.401]$	$[0.179, 0.492]$	$[0.184, 0.338]$
			p-value	1.80E-05	1.40E-04	1.16E-05	3.58E-11
rs3796529	REST	73119	β	0.22	0.19	0.29	0.277
			$[95\%CI]$	$[0.100, 0.343]$	$[0.050, 0.336]$	$[0.082, 0.455]$	$[0.194, 0.361]$
			p-value	3.81E-04	9.30E-03	0.00211	7.85E-11
rs1866146	POMC	73516	β	0.317	0.429	0.417	0.229
			$[95\%CI]$	$[0.217, 0.425]$	$[0.305, 0.555]$	$[0.234, 0.561]$	$[0.160, 0.299]$
			p-value	2.74E-09	2.40E-11	4.02E-07	8.47E-11
rs9844666	PCCB	73511	β	0.257	0.293	0.357	0.254
			$[95\%CI]$	$[0.139, 0.363]$	$[0.163, 0.426]$	$[0.193, 0.530]$	$[0.177, 0.331]$
			p-value	7.53E-06	1.41E-05	2.77E-05	9.23E-11
rs8071847	POLR2A	73565	β	0.313	0.281	0.226	0.264
			$[95\%CI]$	$[0.189, 0.428]$	$[0.127, 0.394]$	$[0.055, 0.397]$	$[0.184, 0.344]$
			p-value	2.98E-07	2.97E-05	9.53E-03	1.10E-10
rs3783937	FBLN5	73104	β	0.259	0.273	0.209	0.249
			$[95\%CI]$	$[0.137, 0.369]$	$[0.119, 0.397]$	$[0.048, 0.404]$	$[0.173, 0.326]$
			p-value	1.28E-05	1.22E-04	0.02	1.72E-10
rs11080149	NF1	73567	β	0.348	0.304	0.362	0.319
			$[95\%CI]$	$[0.211, 0.499]$	$[0.128, 0.457]$	$[0.143, 0.561]$	$[0.221, 0.418]$
			p-value	2.04E-06	0.000248	0.000645	2.48E-10
rs17472113	ZAR1	58807	β	0.235	0.288	0.212	0.27
			$[95\%CI]$	$[0.109, 0.371]$	$[0.142, 0.439]$	$[0.015, 0.381]$	$[0.186, 0.354]$
			p-value	4.23E-04	1.40E-04	2.30E-02	3.23E-10
rs490634	CISH	61124	β	0.3	0.263	0.395	0.346
			$[95\%CI]$	$[0.138, 0.445]$	$[0.065, 0.464]$	$[0.101, 0.617]$	$[0.238, 0.455]$
			p-value	1.33E-04	1.20E-02	2.61E-03	4.04E-10
rs17622208	SLC22A5	73531	β	0.231	0.268	0.25	0.209
			$[95\%CI]$	$[0.126, 0.333]$	$[0.122, 0.358]$	$[0.103, 0.378]$	$[0.143, 0.275]$
			p-value	9.77E-06	9.96E-06	0.00035	5.08E-10
rs2982712	ESR1	73566	β	0.276	0.345	0.284	0.209
			$[95\%CI]$	$[0.183, 0.378]$	$[0.221, 0.456]$	$[0.136, 0.404]$	$[0.143, 0.275]$
			p-value	1.69E-08	8.07E-09	0.0000315	6.50E-10
rs1950500	NFATC4	64408	β	0.209	0.3	0.273	0.24
			$[95\%CI]$	$[0.089, 0.338]$	$[0.157, 0.428]$	$[0.089, 0.415]$	$[0.163, 0.317]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		85%	90%	95%	OLS
			p-value	9.31E-04	0.0000128	0.00111	1.09E-09
rs1476387	PPIL6	73563	β	0.144	0.184	0.182	0.206
			$[95\%CI]$	$[0.034, 0.243]$	$[0.057, 0.298]$	$[0.039, 0.318]$	$[0.139, 0.272]$
			p-value	6.50E-03	2.90E-03	1.00E-02	1.32E-09
rs4946932	FOXO3	73565	β	0.22	0.254	0.278	0.22
			$[95\%CI]$	$[0.106, 0.323]$	$[0.116, 0.364]$	$[0.119, 0.429]$	$[0.148, 0.291]$
			p-value	8.18E-05	5.04E-05	0.000484	1.61E-09
rs1800783	NOS3	71277	β	0.294	0.262	0.235	0.214
			$[95\%CI]$	$[0.196, 0.399]$	$[0.139, 0.379]$	$[0.088, 0.386]$	$[0.144, 0.284]$
			p-value	9.35E-09	1.95E-05	1.84E-03	1.71E-09
rs6718902	STAT1	73557	β	0.2	0.174	0.146	0.233
			$[95\%CI]$	$[0.090, 0.329]$	$[0.053, 0.299]$	$[-0.032, 0.292]$	$[0.157, 0.309]$
			p-value	1.13E-03	5.39E-03	7.40E-02	1.98E-09
rs2425019	MMP24	69770	β		0.163	0.216	0.205
			$[95\%CI]$		$[0.044, 0.273]$	$[0.049, 0.361]$	$[0.138, 0.272]$
			p-value		4.83E-03	6.81E-03	2.55E-09
rs6731022	EIF2AK3	69767	β	0.278	0.257	0.263	0.22
			$[95\%CI]$	$[0.169, 0.391]$	$[0.125, 0.359]$	$[0.119, 0.406]$	$[0.147, 0.292]$
			p-value	5.17E-07	1.61E-05	2.81E-04	2.65E-09
rs12940055	MAP3K3	73566	β	0.322	0.229	0.102	0.311
			$[95\%CI]$	$[0.162, 0.502]$	$[0.052, 0.447]$	$[-0.130, 0.326]$	$[0.207, 0.416]$
			p-value	1.99E-04	2.30E-02	3.78E-01	5.15E-09
rs864745	JAZF1	60616	β	0.259	0.301	0.2	0.213
			$[95\%CI]$	$[0.149, 0.362]$	$[0.180, 0.427]$	$[0.048, 0.348]$	$[0.141, 0.284]$
			p-value	1.92E-06	1.54E-06	7.97E-03	5.21E-09
rs6487088	PDE3A	73563	β	0.229	0.291	0.228	0.246
			$[95\%CI]$	$[0.096, 0.343]$	$[0.128, 0.414]$	$[0.067, 0.425]$	$[0.163, 0.328]$
			p-value	2.65E-04	5.74E-05	0.012	5.63E-09
rs4973410	NCL	64408	β	0.205	0.184	0.22	0.205
			$[95\%CI]$	$[0.092, 0.300]$	$[0.064, 0.304]$	$[0.081, 0.392]$	$[0.136, 0.275]$
			p-value	1.21E-04	2.50E-03	5.31E-03	7.37E-09
rs451061	PRKCZ	71365	β	0.223	0.18	0.269	0.201
			$[95\%CI]$	$[0.113, 0.319]$	$[0.052, 0.301]$	$[0.115, 0.402]$	$[0.132, 0.270]$
			p-value	2.67E-05	4.74E-03	2.19E-04	9.78E-09
rs832575	MAP3K1	73539	β	0.294	0.389	0.358	0.287
			$[95\%CI]$	$[0.145, 0.452]$	$[0.179, 0.539]$	$[0.149, 0.544]$	$[0.187, 0.386]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		85%	90%	95%	OLS
			p-value	1.55E-04	0.0000239	0.000314	1.58E-08
rs4955526	EPHB1	64417	β	0.221	0.221	0.149	0.207
			$[95\%CI]$	$[0.126, 0.331]$	$[0.087, 0.357]$	$[-0.002, 0.322]$	$[0.135, 0.280]$
			p-value	2.66E-05	1.46E-03	0.071	1.94E-08
rs8038415	IGF1R	73516	β	0.168	0.152	0.145	0.189
			$[95\%CI]$	$[0.074, 0.270]$	$[0.034, 0.266]$	$[-0.005, 0.267]$	$[0.123, 0.255]$
			p-value	8.19E-04	1.10E-02	0.037	2.12E-08
rs7578199	HDLBP	73569	β	0.227	0.29	0.206	0.215
			$[95\%CI]$	$[0.114, 0.338]$	$[0.145, 0.422]$	$[0.055, 0.388]$	$[0.139, 0.291]$
			p-value	0.0000755	0.0000382	0.014	2.85E-08
rs7020782	PAPPA	73566	β	0.136	0.16	0.269	0.202
			$[95\%CI]$	$[0.024, 0.250]$	$[0.031, 0.297]$	$[0.130, 0.447]$	$[0.130, 0.274]$
			p-value	1.90E-02	0.018	0.000802	3.48E-08
rs2229712	RPS6KA1	48240	β	0.241	0.341	0.341	0.263
			$[95\%CI]$	$[0.100, 0.398]$	$[0.173, 0.471]$	$[0.104, 0.560]$	$[0.168, 0.357]$
			p-value	0.00146	0.00000637	3.10E-03	5.06E-08
rs7572476	BOK	71367	β	0.181	0.152	0.128	0.186
			$[95\%CI]$	$[0.061, 0.276]$	$[0.045, 0.277]$	$[-0.005, 0.294]$	$[0.119, 0.253]$
			p-value	8.89E-04	0.01	0.092	5.45E-08
rs2066807	PAN2	71822	β	0.211	0.24	0.26	0.232
			$[95\%CI]$	$[0.098, 0.377]$	$[0.116, 0.391]$	$[0.036, 0.488]$	$[0.147, 0.317]$
			p-value	0.00362	0.000481	0.023	8.76E-08
rs1516796	ACAN	69308	β	0.188	0.189	0.298	0.18
			$[95\%CI]$	$[0.091, 0.289]$	$[0.065, 0.295]$	$[0.146, 0.439]$	$[0.112, 0.247]$
			p-value	1.77E-04	0.00124	0.0000638	2.04E-07
rs6180	GHR	73552	β	0.136	0.151	0.124	0.175
			$[95\%CI]$	$[0.032, 0.232]$	$[0.044, 0.272]$	$[-0.005, 0.251]$	$[0.109, 0.240]$
			p-value	0.00687	0.00919	0.06	2.11E-07
rs8055190	LRRC36	64416	β	0.461	0.38	0.425	0.449
			$[95\%CI]$	$[0.161, 0.703]$	$[0.069, 0.727]$	$[0.069, 0.779]$	$[0.279, 0.620]$
			p-value	1.08E-03	2.20E-02	0.021	2.30E-07
rs17106235	FAF1	60489	β	0.402	0.411	0.424	0.316
			$[95\%CI]$	$[0.220, 0.593]$	$[0.198, 0.632]$	$[0.139, 0.647]$	$[0.195, 0.436]$
			p-value	3.70E-05	0.000209	0.0012	3.11E-07
rs3739707	LPAR1	73568	β	0.121	0.141	0.22	0.197
			$[95\%CI]$	$[0.009, 0.252]$	$[-0.011, 0.269]$	$[0.050, 0.359]$	$[0.121, 0.272]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		85%	90%	95%	OLS
			p-value	4.90E-02	4.80E-02	5.24E-03	3.22E-07
rs674424	ABCG4	73568	β	0.136	0.132	0.085	0.189
			$[95\%CI]$	$[0.013, 0.240]$	$[-0.025, 0.250]$	$[-0.092, 0.233]$	$[0.112, 0.265]$
			p-value	2.00E-02	6.30E-02	0.311	1.32E-06
rs12225387	NEU3	60621	β	0.196	0.202	0.192	0.199
			$[95\%CI]$	$[0.086, 0.319]$	$[0.043, 0.339]$	$[0.028, 0.365]$	$[0.118, 0.280]$
			p-value	0.000885	7.19E-03	0.025	1.39E-06
rs3812265	CNOT4	69776	β	0.26	0.286	0.275	0.191
			$[95\%CI]$	$[0.121, 0.368]$	$[0.137, 0.425]$	$[0.093, 0.437]$	$[0.113, 0.270]$
			p-value	0.0000371	0.0000944	1.54E-03	1.72E-06
rs10208728	IHH	64409	β	0.247	0.281	0.5	0.285
			$[95\%CI]$	$[0.042, 0.407]$	$[0.073, 0.475]$	$[0.271, 0.710]$	$[0.167, 0.402]$
			p-value	7.69E-03	5.67E-03	0.00000789	1.92E-06
rs291979	GRK5	73568	β	0.182	0.15	0.187	0.189
			$[95\%CI]$	$[0.062, 0.301]$	$[0.007, 0.296]$	$[0.017, 0.347]$	$[0.110, 0.267]$
			p-value	0.00289	0.039	0.025	2.37E-06
rs2715553	RARA	69206	β	0.16	0.111	0.147	0.156
			$[95\%CI]$	$[0.061, 0.263]$	$[-0.008, 0.232]$	$[-0.020, 0.277]$	$[0.088, 0.225]$
			p-value	0.00187	0.069	0.05	7.30E-06
rs2057291	GNAS	70100	β	0.147	0.118	0.121	0.159
			$[95\%CI]$	$[0.047, 0.266]$	$[-0.012, 0.242]$	$[-0.053, 0.263]$	$[0.088, 0.230]$
			p-value	0.00784	0.069	0.132	1.18E-05
rs4803520	GRIK5	69747	β	0.194	0.277	0.299	0.235
			$[95\%CI]$	$[0.036, 0.378]$	$[0.030, 0.435]$	$[0.078, 0.476]$	$[0.129, 0.341]$
			p-value	0.026	0.00791	0.00336	1.45E-05
rs10736682	APLNR	73090	β	0.068	0.071	0.081	0.143
			$[95\%CI]$	$[-0.027, 0.170]$	$[-0.060, 0.184]$	$[-0.080, 0.214]$	$[0.077, 0.210]$
			p-value	1.75E-01	0.26	0.275	2.31E-05
rs2909430	TP53	73412	β	0.125	0.132	0.305	0.209
			$[95\%CI]$	$[-0.035, 0.263]$	$[-0.060, 0.318]$	$[0.073, 0.482]$	$[0.110, 0.308]$
			p-value	0.097	0.172	0.00331	3.57E-05
rs12050767	CYP19A1	73563	β	0.144	0.142	0.145	0.139
			$[95\%CI]$	$[0.049, 0.241]$	$[0.010, 0.245]$	$[-0.019, 0.290]$	$[0.073, 0.205]$
			p-value	3.10E-03	1.80E-02	6.70E-02	3.62E-05
rs602633	PSRC1	64375	β	0.206	0.227	0.149	0.179
			$[95\%CI]$	$[0.070, 0.327]$	$[0.057, 0.383]$	$[-0.004, 0.360]$	$[0.094, 0.264]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		85%	90%	95%	OLS
			p-value	0.00152	0.00585	0.11	3.64E-05
rs1738475	HTR1D	64411	β	0.113	0.132	0.075	0.142
			$[95\%CI]$	$[0.005, 0.222]$	$[0.000, 0.263]$	$[-0.081, 0.226]$	$[0.072, 0.213]$
			p-value	4.10E-02	0.05	0.339	7.70E-05
rs17754	RFC1	73558	β	0.183	0.165	0.062	0.13
			$[95\%CI]$	$[0.093, 0.297]$	$[0.062, 0.286]$	$[-0.073, 0.242]$	$[0.063, 0.197]$
			p-value	4.30E-04	3.65E-03	4.43E-01	1.29E-04
rs17541471	NPR3	69777	β	0.197	0.197	0.158	0.165
			$[95\%CI]$	$[0.076, 0.330]$	$[0.026, 0.326]$	$[-0.054, 0.313]$	$[0.080, 0.250]$
			p-value	0.0022	0.0088	0.098	1.34E-04
rs1342586	TGFB2	69745	β	0.125	0.17	0.138	0.157
			$[95\%CI]$	$[-0.002, 0.254]$	$[0.021, 0.327]$	$[-0.052, 0.294]$	$[0.075, 0.238]$
			p-value	0.054	0.028	1.15E-01	1.69E-04
rs3814115	PCSK5	73537	β	0.193	0.177	0.154	0.133
			$[95\%CI]$	$[0.089, 0.307]$	$[0.052, 0.297]$	$[0.028, 0.323]$	$[0.063, 0.203]$
			p-value	0.000505	4.48E-03	4.00E-02	2.02E-04
rs1780616	LBP	73560	β	0.094	0.076	0.062	0.13
			$[95\%CI]$	$[-0.012, 0.197]$	$[-0.038, 0.211]$	$[-0.082, 0.203]$	$[0.061, 0.200]$
			p-value	0.076	0.232	0.387	2.28E-04
rs3736228	LRP5	73508	β	0.213	0.275	0.218	0.175
			$[95\%CI]$	$[0.049, 0.338]$	$[0.069, 0.425]$	$[0.012, 0.424]$	$[0.081, 0.269]$
			p-value	3.86E-03	2.76E-03	0.036	2.74E-04
rs212517	ECE1	71344	β	0.108	0.1	0.073	0.126
			$[95\%CI]$	$[-0.004, 0.213]$	$[-0.028, 0.227]$	$[-0.076, 0.213]$	$[0.058, 0.195]$
			p-value	0.05	0.129	0.314	2.94E-04
rs7359336	NFAT5	73094	β	0.137	0.055	0.115	0.12
			$[95\%CI]$	$[0.027, 0.231]$	$[-0.062, 0.191]$	$[-0.045, 0.249]$	$[0.053, 0.187]$
			p-value	7.60E-03	4.03E-01	1.27E-01	4.64E-04
rs2682552	XRCC1	73106	β	0.207	0.26	0.363	0.148
			$[95\%CI]$	$[0.066, 0.328]$	$[0.105, 0.408]$	$[0.175, 0.547]$	$[0.064, 0.232]$
			p-value	0.0025	0.000749	0.000122	5.35E-04
rs17085675	PCSK1	69775	β	0.005	0.006	0.037	0.131
			$[95\%CI]$	$[-0.097, 0.125]$	$[-0.118, 0.147]$	$[-0.134, 0.198]$	$[0.057, 0.206]$
			p-value	0.931	0.93	0.661	5.54E-04
rs11102986	RXRA	71224	β	0.046	0.082	0.115	0.151
			$[95\%CI]$	$[-0.077, 0.189]$	$[-0.063, 0.252]$	$[-0.094, 0.277]$	$[0.064, 0.238]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		85%	90%	95%	OLS
			p-value	0.503	0.301	0.219	6.49E-04
rs12603813	PLCD3	71116	β	0.159	0.199	0.227	0.135
			$[95\%CI]$	$[0.033, 0.271]$	$[0.052, 0.326]$	$[0.054, 0.387]$	$[0.057, 0.212]$
			p-value	0.00928	0.00447	0.00747	6.61E-04
rs6219	IGF1	73562	β	0.34	0.423	0.293	0.19
			$[95\%CI]$	$[0.190, 0.518]$	$[0.204, 0.580]$	$[0.062, 0.570]$	$[0.080, 0.299]$
			p-value	0.000044	0.0000125	0.026	7.01E-04
rs2234693	ESR1	69772	β	0.131	0.128	0.186	0.113
			$[95\%CI]$	$[0.015, 0.214]$	$[-0.005, 0.252]$	$[0.040, 0.338]$	$[0.045, 0.180]$
			p-value	0.00918	0.05	1.40E-02	1.03E-03
rs46522	UBE2Z	64204	β	0.055	0.004	0.15	0.115
			$[95\%CI]$	$[-0.057, 0.151]$	$[-0.106, 0.145]$	$[0.008, 0.300]$	$[0.046, 0.185]$
			p-value	0.303	0.945	0.044	1.19E-03
rs891088	INSR	73562	β	0.162	0.212	0.073	0.123
			$[95\%CI]$	$[0.050, 0.270]$	$[0.054, 0.317]$	$[-0.075, 0.260]$	$[0.049, 0.197]$
			p-value	0.00384	0.00154	0.392	1.20E-03
rs9857730	VILL	73358	β	0.097	0.156	0.266	0.132
			$[95\%CI]$	$[-0.041, 0.206]$	$[-0.001, 0.306]$	$[0.114, 0.441]$	$[0.052, 0.213]$
			p-value	0.13	0.046	0.00124	1.33E-03
rs10185680	MFSD2B	73232	β	0.067	0.044	0.073	0.108
			$[95\%CI]$	$[-0.041, 0.173]$	$[-0.055, 0.176]$	$[-0.063, 0.191]$	$[0.042, 0.173]$
			p-value	0.224	0.461	0.256	1.39E-03
rs7163907	PTPN9	73234	β	-0.117	-0.148	-0.1	-0.121
			$[95\%CI]$	$[-0.220, 0.005]$	$[-0.290, -0.017]$	$[-0.263, 0.095]$	$[-0.196, -0.046]$
			p-value	0.038	0.032	0.268	1.65E-03
rs2176167	NOP58	73508	β	0.107	0.113	0.134	0.109
			$[95\%CI]$	$[0.011, 0.214]$	$[0.000, 0.245]$	$[-0.021, 0.274]$	$[0.040, 0.179]$
			p-value	0.039	0.067	0.078	1.90E-03
rs7557989	THADA	64417	β	0.058	0.097	0.075	0.114
			$[95\%CI]$	$[-0.056, 0.167]$	$[-0.037, 0.228]$	$[-0.078, 0.228]$	$[0.040, 0.188]$
			p-value	0.313	0.149	0.345	0.003
rs510769	OPRM1	64413	β	0.145	0.175	0.149	0.123
			$[95\%CI]$	$[0.022, 0.266]$	$[0.034, 0.308]$	$[-0.038, 0.323]$	$[0.043, 0.204]$
			p-value	0.020	0.013	0.102	0.003
rs7756224	NMBR	60624	β	0.032	0.059	0.012	0.111
			$[95\%CI]$	$[-0.065, 0.146]$	$[-0.077, 0.181]$	$[-0.135, 0.182]$	$[0.038, 0.184]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		85%	90%	95%	OLS
			p-value	0.553	0.364	0.879	0.003
rs2282537	POU2F3	64419	β	0.286	0.188	0.200	0.152
			$[95\%CI]$	$[0.109, 0.414]$	$[0.005, 0.387]$	$[-0.013, 0.386]$	$[0.052, 0.252]$
			p-value	0.000	0.055	0.049	0.003
rs7853859	CENPP	69662	β	0.099	0.130	0.110	0.105
			$[95\%CI]$	$[-0.005, 0.209]$	$[-0.004, 0.254]$	$[-0.033, 0.262]$	$[0.036, 0.175]$
			p-value	0.072	0.049	0.135	0.003
rs2229642	ITPR3	52677	β	0.125	0.178	0.116	0.112
			$[95\%CI]$	$[-0.008, 0.219]$	$[0.035, 0.312]$	$[-0.070, 0.265]$	$[0.033, 0.190]$
			p-value	0.031	0.011	0.180	0.005
rs5015437	LMF1	64185	β	0.101	0.071	0.050	0.102
			$[95\%CI]$	$[-0.018, 0.201]$	$[-0.060, 0.196]$	$[-0.105, 0.203]$	$[0.030, 0.174]$
			p-value	0.069	0.278	0.523	0.006
rs7963565	KNTC1	73568	β	0.183	0.221	0.220	0.097
			$[95\%CI]$	$[0.066, 0.290]$	$[0.100, 0.338]$	$[0.072, 0.359]$	$[0.028, 0.167]$
			p-value	1.46E-03	2.72E-04	0.003	0.006
rs696	NFKBIA	73461	β	0.119	0.076	0.008	0.092
			$[95\%CI]$	$[0.003, 0.213]$	$[-0.043, 0.195]$	$[-0.130, 0.160]$	$[0.024, 0.160]$
			p-value	0.025	0.209	0.909	0.00815
rs25656	NFATC1	48455	β	0.039	0.052	0.163	0.114
			$[95\%CI]$	$[-0.086, 0.182]$	$[-0.083, 0.226]$	$[-0.017, 0.360]$	$[0.029, 0.199]$
			p-value	0.570	0.510	0.089	0.00879
rs3100776	IHH	64334	β	0.435	0.526	0.461	0.243
			$[95\%CI]$	$[0.131, 0.698]$	$[0.174, 0.804]$	$[0.052, 0.826]$	$[0.061, 0.426]$
			p-value	2.33E-03	8.42E-04	0.017	0.00886
rs12145922	PKN2	69346	β	0.132	0.143	0.147	0.088
			$[95\%CI]$	$[0.024, 0.229]$	$[0.022, 0.265]$	$[-0.004, 0.288]$	$[0.020, 0.156]$
			p-value	0.012	0.018	0.048	0.011
rs526134	USP37	60625	β	0.149	0.170	0.153	0.094
			$[95\%CI]$	$[0.028, 0.263]$	$[0.033, 0.305]$	$[0.016, 0.328]$	$[0.022, 0.167]$
			p-value	0.011	0.014	0.055	0.011
rs7004280	RPS20	64415	β	0.169	0.196	0.299	0.207
			$[95\%CI]$	$[-0.084, 0.421]$	$[-0.153, 0.596]$	$[-0.107, 0.671]$	$[0.018, 0.396]$
			p-value	0.188	0.307	0.130	0.032
rs4808199	GATAD2A	64415	β	0.166	0.224	0.198	0.098
			$[95\%CI]$	$[0.046, 0.316]$	$[0.059, 0.398]$	$[-0.021, 0.349]$	$[0.007, 0.189]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		85%	90%	95%	OLS
			p-value	1.80E-02	0.010	0.037	0.035
rs3821009	PDE11A	69745	β	0.284	0.356	0.411	0.129
			$[95\%CI]$	$[0.126, 0.464]$	$[0.130, 0.570]$	$[0.176, 0.746]$	$[0.006, 0.252]$
			p-value	8.95E-04	0.00128	0.00554	0.04
rs2291617	METTL1	73099	β	0.072	0.063	0.122	0.073
			$[95\%CI]$	$[-0.038, 0.176]$	$[-0.063, 0.198]$	$[-0.036, 0.280]$	$[0.003, 0.144]$
			p-value	0.185	0.344	0.125	0.042
rs1051168	SEC11A	64399	β	0.075	0.172	0.279	0.08
			$[95\%CI]$	$[-0.035, 0.188]$	$[0.029, 0.318]$	$[0.110, 0.425]$	$[0.002, 0.157]$
			p-value	0.192	2.00E-02	0.000418	0.043
rs803932	ASTN2	64105	β	-0.002	-0.016	-0.044	0.073
			$[95\%CI]$	$[-0.119, 0.107]$	$[-0.142, 0.116]$	$[-0.188, 0.108]$	$[-0.001, 0.147]$
			p-value	9.77E-01	0.806	0.567	0.054
rs2247870	GPR98	69775	β	0.000	0.002	-0.083	0.066
			$[95\%CI]$	$[-0.094, 0.121]$	$[-0.124, 0.113]$	$[-0.238, 0.085]$	$[-0.001, 0.134]$
			p-value	1.000	0.979	0.308	0.055
rs12503378	NUDT6	73569	β	0.070	0.144	0.049	0.077
			$[95\%CI]$	$[-0.060, 0.201]$	$[-0.015, 0.305]$	$[-0.154, 0.258]$	$[-0.011, 0.166]$
			p-value	0.291	0.079	0.642	0.085
rs2145923	NPR2	69777	β	0.143	0.128	0.098	0.076
			$[95\%CI]$	$[0.016, 0.258]$	$[-0.037, 0.288]$	$[-0.094, 0.287]$	$[-0.011, 0.163]$
			p-value	0.019	0.122	0.312	0.088
rs12603582	ITGB3	73541	β	0.000	-0.026	-0.131	0.060
			$[95\%CI]$	$[-0.119, 0.145]$	$[-0.174, 0.086]$	$[-0.303, 0.061]$	$[-0.019, 0.139]$
			p-value	1.000	0.699	0.161	0.137
rs4864548	CLOCK	73568	β	0.110	0.023	0.000	0.045
			$[95\%CI]$	$[0.001, 0.204]$	$[-0.079, 0.150]$	$[-0.136, 0.167]$	$[-0.022, 0.113]$
			p-value	0.031	0.694	1.000	0.190
rs2735469	MRPL23	67563	β	0.141	0.137	0.056	0.061
			$[95\%CI]$	$[0.004, 0.290]$	$[-0.042, 0.306]$	$[-0.162, 0.238]$	$[-0.033, 0.154]$
			p-value	0.049	0.117	0.587	0.204
rs1051431	MPHOSPH9	73561	β	0.143	0.236	0.242	0.038
			$[95\%CI]$	$[0.018, 0.261]$	$[0.081, 0.360]$	$[0.075, 0.430]$	$[-0.041, 0.117]$
			p-value	0.020	7.46E-04	0.007	0.341
rs41132	AP3B1	73441	β	0.069	0.130	0.085	0.037
			$[95\%CI]$	$[-0.042, 0.212]$	$[-0.004, 0.251]$	$[-0.099, 0.228]$	$[-0.040, 0.115]$

Continued on next page

Table A7 – *Continued from previous page*

SNP	Gene	N		85%	90%	95%	OLS
			p-value	0.279	0.045	0.312	0.344
rs6127698	MC3R	73486	β	0.001	0.023	0.074	0.030
			$[95\%CI]$	$[-0.086, 0.121]$	$[-0.073, 0.170]$	$[-0.096, 0.198]$	$[-0.035, 0.096]$
			p-value	0.983	0.709	0.324	0.368
rs1481892	ARNTL	73543	β	0.058	0.051	-0.053	0.033
			$[95\%CI]$	$[-0.054, 0.151]$	$[-0.074, 0.188]$	$[-0.204, 0.115]$	$[-0.039, 0.105]$
			p-value	0.268	0.445	0.512	0.370
rs2633442	MKRN2	60624	β	0.000	-0.028	-0.026	0.023
			$[95\%CI]$	$[-0.090, 0.123]$	$[-0.150, 0.107]$	$[-0.175, 0.119]$	$[-0.049, 0.095]$
			p-value	1.000	0.669	0.731	0.527
rs1535	FADS1	73553	β	-0.107	-0.115	0.032	-0.011
			$[95\%CI]$	$[-0.208, 0.011]$	$[-0.237, 0.024]$	$[-0.114, 0.173]$	$[-0.081, 0.059]$
			p-value	0.054	0.084	0.664	0.754
rs10861148	HSP90B1	73557	β	0.000	0.014	0.055	0.005
			$[95\%CI]$	$[-0.152, 0.165]$	$[-0.180, 0.214]$	$[-0.175, 0.293]$	$[-0.099, 0.110]$
			p-value	1.000	0.891	0.648	0.921
GS-Height		73570	β	0.175	0.18	0.179	0.176
			$[95\%CI]$	$[0.165, 0.186]$	$[0.169, 0.192]$	$[0.166, 0.194]$	$[0.169, 0.182]$
			p-value	2.64E-229	3.06E-219	2.45E-132	<2.2E-308
			VarianceExplained	1.805%	1.822%	1.794%	1.636%

Table A8: Quantifying the effect of height percentile on conditional quantile regression (CQR) estimates using meta-regression (MR). MR was used to model variability in the CQR estimates across height percentiles. Note that the percentiles were re-centered around the 50$^{\text{th}}$percentile so that the intercept from MR models corresponds to the main effect of the SNP at the median. ($*$) Denotes statistical significance at the Bonferroni-adjusted p-value ($p < 3.85 \times 10^{-4}$), RI_{50} is the re-centered intercept of the MR models, MR is the effect of height percentile on CQR estimates (cm per Effect Allele per Height Percentile), 95%CI are the 95% confidence intervals.

SNP	Gene	RI_{50}	$\beta_{MR}\left[95\%CI\right]$	p-value	
rs6219	IGF1	0.162	0.479$\left[0.229, 0.730\right]$	1.80E-04	$*$
rs1866146	POMC	0.235	0.282$\left[0.120, 0.443\right]$	6.17E-04	
rs551219	COL24A1	0.241	-0.277$\left[-0.444, -0.109\right]$	0.001	
rs9930741	ERI2	0.285	-0.203$\left[-0.359, -0.047\right]$	0.011	
rs3821009	PDE11A	0.096	0.336$\left[0.062, 0.610\right]$	0.016	
rs7963565	KNTC1	0.110	0.190$\left[0.029, 0.350\right]$	0.021	
rs3100776	IHH	0.304	0.501$\left[0.070, 0.933\right]$	0.023	
rs17085675	PCSK1	0.121	-0.199$\left[-0.373, -0.026\right]$	0.024	
rs2853977	HCP5	0.624	-0.218$\left[-0.409, -0.026\right]$	0.026	
rs1051431	MPHOSPH9	0.050	0.208$\left[0.024, 0.392\right]$	0.026	
rs258281	RAB26	0.305	-0.234$\left[-0.440, -0.029\right]$	0.026	
rs526134	USP37	0.091	0.188$\left[0.016, 0.360\right]$	0.032	
rs1476387	PPIL6	0.191	-0.163$\left[-0.318, -0.008\right]$	0.040	
rs11102986	RXRA	0.138	-0.203$\left[-0.404, -0.003\right]$	0.047	
rs7557989	THADA	0.118	-0.169$\left[-0.340, 0.001\right]$	0.051	
rs6718902	STAT1	0.251	-0.171$\left[-0.345, 0.003\right]$	0.054	
rs1036477	FBN1	0.375	-0.240$\left[-0.485, 0.006\right]$	0.055	
rs12603582	ITGB3	0.057	-0.181$\left[-0.370, 0.008\right]$	0.060	
rs803932	ASTN2	0.076	-0.162$\left[-0.334, 0.009\right]$	0.064	
rs9892365	TBX2	0.357	0.153$\left[-0.010, 0.316\right]$	0.067	
rs1776897	HMGA1	0.595	0.264$\left[-0.023, 0.551\right]$	0.072	
rs3814115	PCSK5	0.127	0.150$\left[-0.013, 0.313\right]$	0.072	
rs1051168	SEC11A	0.087	0.144$\left[-0.028, 0.316\right]$	0.100	
rs2854207	CSH2	0.296	0.153$\left[-0.031, 0.337\right]$	0.103	
rs41132	AP3B1	0.022	0.150$\left[-0.031, 0.331\right]$	0.105	
rs1738475	HTR1D	0.148	-0.136$\left[-0.304, 0.032\right]$	0.112	
rs2682552	XRCC1	0.133	0.152$\left[-0.037, 0.342\right]$	0.115	
rs3796529	REST	0.275	-0.151$\left[-0.341, 0.039\right]$	0.120	
rs3812265	CNOT4	0.197	0.145$\left[-0.040, 0.329\right]$	0.125	
rs46522	UBE2Z	0.122	-0.122$\left[-0.284, 0.040\right]$	0.140	
rs4808199	GATAD2A	0.099	0.155$\left[-0.052, 0.362\right]$	0.142	
rs6180	GHR	0.188	-0.111$\left[-0.259, 0.037\right]$	0.143	
rs17754	RFC1	0.134	0.105$\left[-0.048, 0.258\right]$	0.177	

Continued on next page

Table A8 – *Continued from previous page*

SNP	Gene	RI_{50}	β_{MR}	$95\%CI$	p-value
rs2735469	MRPL23	0.074	0.147	$-0.074, 0.368$	0.192
rs4864548	CLOCK	0.036	0.104	$-0.055, 0.262$	0.200
rs7137534	PDE3A	0.362	0.106	$-0.059, 0.272$	0.209
rs2282537	POU2F3	0.154	0.145	$-0.083, 0.374$	0.213
rs709939	SAMD4A	0.245	0.100	$-0.057, 0.257$	0.213
rs2715553	RARA	0.166	-0.101	$-0.261, 0.059$	0.215
rs2291617	METTL1	0.081	0.104	$-0.062, 0.270$	0.221
rs9857730	VILL	0.115	-0.112	$-0.294, 0.070$	0.227
rs2229642	ITPR3	0.103	0.109	$-0.071, 0.290$	0.236
rs6731022	EIF2AK3	0.225	0.100	$-0.066, 0.266$	0.237
rs3739707	LPAR1	0.191	-0.107	$-0.285, 0.071$	0.240
rs8038415	IGF1R	0.201	-0.092	$-0.247, 0.063$	0.245
rs25656	NFATC1	0.115	-0.120	$-0.323, 0.083$	0.246
rs17106235	FAF1	0.340	0.168	$-0.120, 0.456$	0.253
rs2909430	TP53	0.213	-0.130	$-0.360, 0.100$	0.266
rs490634	CISH	0.374	-0.133	$-0.386, 0.120$	0.304
rs12603813	PLCD3	0.144	0.089	$-0.090, 0.268$	0.329
rs17541471	NPR3	0.169	0.099	$-0.101, 0.299$	0.333
rs1636255	GNA12	0.256	0.088	$-0.092, 0.267$	0.338
rs7756224	NMBR	0.107	-0.080	$-0.250, 0.090$	0.358
rs1822469	PPP3R1	0.258	-0.075	$-0.236, 0.087$	0.364
rs7163907	PTPN9	-0.111	0.080	$-0.096, 0.256$	0.371
rs10736682	APLNR	0.136	-0.070	$-0.224, 0.085$	0.377
rs1042725	HMGA2	0.569	-0.069	$-0.224, 0.085$	0.379
rs7004280	RPS20	0.229	0.180	$-0.235, 0.594$	0.396
rs674424	ABCG4	0.198	-0.077	$-0.256, 0.102$	0.398
rs17622208	SLC22A5	0.203	0.067	$-0.090, 0.224$	0.401
rs2247870	GPR98	0.092	-0.065	$-0.228, 0.098$	0.434
rs1342586	TGFB2	0.128	-0.070	$-0.263, 0.122$	0.474
rs12145922	PKN2	0.076	0.056	$-0.100, 0.212$	0.484
rs1516796	ACAN	0.173	0.056	$-0.103, 0.215$	0.486
rs752313	EZH1	0.262	-0.057	$-0.218, 0.103$	0.486
rs10208728	IHH	0.300	0.090	$-0.178, 0.358$	0.509
rs1481892	ARNTL	0.041	0.056	$-0.111, 0.223$	0.514
rs2145923	NPR2	0.074	0.065	$-0.132, 0.261$	0.519
rs572169	GHSR	0.350	-0.054	$-0.221, 0.112$	0.521
rs2176167	NOP58	0.114	-0.050	$-0.209, 0.108$	0.534
rs1780616	LBP	0.121	-0.052	$-0.217, 0.113$	0.537
rs2982712	ESR1	0.206	0.049	$-0.108, 0.205$	0.541
rs696	NFKBIA	0.094	0.049	$-0.110, 0.208$	0.546
rs2679178	NPPC	0.536	0.087	$-0.195, 0.368$	0.546
rs2066807	PAN2	0.231	0.062	$-0.140, 0.265$	0.546

Continued on next page

Table A8 – *Continued from previous page*

SNP	Gene	RI_{50}	β_{MR}	$95\%CI$	p-value
rs451061	PRKCZ	0.207	-0.049	$-0.206, 0.109$	0.546
rs891088	INSR	0.121	0.053	$-0.120, 0.227$	0.548
rs5015437	LMF1	0.104	-0.048	$-0.214, 0.118$	0.572
rs7020782	PAPPA	0.203	-0.048	$-0.218, 0.122$	0.577
rs4946932	FOXO3	0.227	0.045	$-0.120, 0.210$	0.593
rs4320932	IGF2	0.347	-0.051	$-0.250, 0.147$	0.612
rs864745	JAZF1	0.204	0.043	$-0.124, 0.209$	0.615
rs1800783	NOS3	0.212	0.041	$-0.120, 0.201$	0.619
rs7359336	NFAT5	0.128	-0.037	$-0.193, 0.118$	0.636
rs4803520	GRIK5	0.243	-0.058	$-0.302, 0.185$	0.638
rs6127698	MC3R	0.045	0.037	$-0.117, 0.191$	0.638
rs510769	OPRM1	0.119	0.044	$-0.145, 0.233$	0.651
rs9844666	PCCB	0.251	-0.038	$-0.217, 0.141$	0.677
rs10861148	HSP90B1	0.017	0.049	$-0.192, 0.289$	0.690
rs212517	ECE1	0.137	0.032	$-0.127, 0.191$	0.694
rs13076290	CTNNB1	0.233	-0.029	$-0.183, 0.126$	0.717
rs1950500	NFATC4	0.241	0.032	$-0.145, 0.208$	0.724
rs13072536	ITIH4	0.252	-0.032	$-0.214, 0.149$	0.725
rs17472113	ZAR1	0.291	0.036	$-0.165, 0.236$	0.727
rs602633	PSRC1	0.170	-0.031	$-0.227, 0.166$	0.760
rs1535	FADS1	0.005	-0.025	$-0.190, 0.139$	0.762
rs780094	GCKR	0.368	-0.021	$-0.177, 0.135$	0.790
rs12503378	NUDT6	0.075	0.028	$-0.182, 0.238$	0.792
rs12940055	MAP3K3	0.283	-0.030	$-0.275, 0.215$	0.808
rs2425019	MMP24	0.217	-0.019	$-0.179, 0.141$	0.815
rs2053156	GRB2	0.351	0.023	$-0.172, 0.217$	0.821
rs6487088	PDE3A	0.247	-0.022	$-0.216, 0.173$	0.827
rs9366637	HFE	0.433	0.035	$-0.285, 0.355$	0.829
rs2633442	MKRN2	0.005	-0.018	$-0.184, 0.147$	0.831
rs158676	CDK5RAP1	0.237	0.017	$-0.147, 0.182$	0.835
rs832575	MAP3K1	0.280	0.024	$-0.212, 0.260$	0.843
rs7522692	PIGC	0.298	0.019	$-0.182, 0.221$	0.852
rs2057291	GNAS	0.176	-0.014	$-0.182, 0.154$	0.868
rs8071847	POLR2A	0.250	-0.016	$-0.203, 0.171$	0.868
rs7853859	CENPP	0.106	-0.014	$-0.177, 0.150$	0.871
rs2234693	ESR1	0.114	-0.013	$-0.174, 0.148$	0.871
rs7578199	HDLBP	0.200	-0.014	$-0.196, 0.169$	0.883
rs3782415	SOCS2	0.455	-0.013	$-0.205, 0.178$	0.892
rs3736228	LRP5	0.139	-0.014	$-0.236, 0.209$	0.904
rs4955526	EPHB1	0.221	0.010	$-0.164, 0.183$	0.914
rs10185680	MFSD2B	0.107	0.007	$-0.146, 0.160$	0.928
rs4973410	NCL	0.201	-0.007	$-0.169, 0.155$	0.935

Continued on next page

Table A8 – *Continued from previous page*

SNP	Gene	RI_{50}	β_{MR} $95\%CI$	p-value
rs291979	GRK5	0.167	0.006 $[-0.175, 0.188]$	0.945
rs11080149	NF1	0.306	-0.008 $[-0.235, 0.219]$	0.945
rs3783937	FBLN5	0.245	-0.006 $[-0.185, 0.173]$	0.947
rs8055190	LRRC36	0.472	0.014 $[-0.388, 0.415]$	0.947
rs12050767	CYP19A1	0.143	0.003 $[-0.151, 0.156]$	0.973
rs12225387	NEU3	0.202	0.002 $[-0.184, 0.188]$	0.981
rs2229712	RPS6KA1	0.270	-0.002 $[-0.223, 0.218]$	0.983
rs7572476	BOK	0.190	0.002 $[-0.155, 0.159]$	0.984

Table A9: Sensitivity analysis. Conditional quantile regression (CQR) models for 37 BMI/obesity SNPs were conducted before except that models were fitted with adjustment for diabetic status or age-linear. Meta-regression (MR) analysis was applied to examine the association between CQR estimates and the BMI percentile as above. In addition, CQR estimates were obtained every 10th percentile rather than every 5th percentile of BMI. The results from Table 2 are included for comparison. $(*)$ Denotes statistical significance at the Bonferroni-adjusted p-value ($p < 1.32 \times 10^{-3}$), RI_{50} is the re-centered intercept of the MR models, β_{MR} is the effect of BMI quantile on CQR estimates (kg/m^2 per Effect Allele per BMI Percentile), 95%CI are the 95% confidence intervals. These conditions had little effect on the outcome of MR analysis, which supports the robustness of the main findings. Note that 3 SNPs with significantly increasing effects across the sample BMI distribution showed nominal effects when diabetic status adjustments were applied including, MAP2K5 (rs997295, FTO (rs6499653) and NT5C2 (rs3824755); while CDKAL1 (rs9356744) showed significantly increasing effects across the sample BMI distribution only when CQR models were fitted with diabetic status adjustment.

Part 1 - Original Model

SNP	Gene	RI_{50}	β_{MR}	$[95\%CI]$	p-value	
rs1421085	FTO	0.473	0.495	$[0.370, 0.620]$	8.69E-15	$*$
rs6235	PCSK1	0.078	0.320	$[0.180, 0.459]$	7.11E-06	$*$
rs7903146	TCF7L2	0.144	0.303	$[0.169, 0.437]$	9.60E-06	$*$
rs11873305	MC4R	0.344	0.603	$[0.311, 0.895]$	5.08E-05	$*$
rs12617233	FANCL	0.129	0.261	$[0.134, 0.387]$	5.30E-05	$*$
rs11672660	GIPR	0.227	0.294	$[0.141, 0.447]$	1.64E-04	$*$
rs997295	MAP2K5	0.131	0.228	$[0.103, 0.352]$	3.25E-04	$*$
rs6499653	FTO	0.121	0.253	$[0.108, 0.398]$	6.23E-04	$*$
rs3824755	NT5C2	0.222	0.362	$[0.151, 0.574]$	7.90E-04	$*$
rs7553158	TNNI3K	0.099	0.196	$[0.071, 0.322]$	0.002	
rs10767664	BDNF	0.247	0.217	$[0.064, 0.370]$	0.006	
rs4788099	SH2B1	0.151	0.194	$[0.057, 0.332]$	0.006	
rs17066846	MC4R	0.124	0.215	$[0.063, 0.367]$	0.006	
rs9356744	CDKAL1	0.063	0.186	$[0.050, 0.322]$	0.007	
rs6453133	HMGCR	0.130	0.177	$[0.040, 0.314]$	0.011	
rs2819347	LMOD1	0.111	0.137	$[0.004, 0.269]$	0.044	
rs2075650	TOMM40	0.283	0.161	$[-0.019, 0.341]$	0.079	
rs4946932	FOXO3	0.106	0.120	$[-0.016, 0.256]$	0.084	
rs2984618	TAL1	0.069	0.108	$[-0.019, 0.235]$	0.095	
rs980828	NOS1AP	0.024	0.095	$[-0.030, 0.220]$	0.135	
rs1788826	NPC1	0.109	0.094	$[-0.036, 0.224]$	0.156	
rs11570094	SPI1	0.103	0.096	$[-0.039, 0.231]$	0.163	
rs7988412	MTIF3	0.088	0.109	$[-0.062, 0.280]$	0.212	
rs2283228	KCNQ1	0.003	0.147	$[-0.094, 0.388]$	0.232	

Continued on next page

Table A9 – *Continued from previous page*

SNP	Gene	RI_{50}	β_{MR} $[95\%CI]$	p-value	
rs739564	IQCK	0.122	0.100 $[-0.065, 0.265]$	0.234	
rs526134	USP37	0.062	0.079 $[-0.055, 0.212]$	0.247	
rs2272903	TFAP2B	0.145	0.113 $[-0.084, 0.310]$	0.261	
rs2836754	ETS2	0.086	0.073 $[-0.060, 0.206]$	0.280	
rs2535633	ITIH4	0.016	0.068 $[-0.059, 0.194]$	0.296	
rs11208662	LEPR	0.142	0.111 $[-0.105, 0.327]$	0.314	
rs6232	PCSK1	0.075	0.133 $[-0.137, 0.404]$	0.334	
rs749767	KAT8	0.048	0.058 $[-0.075, 0.191]$	0.390	
rs1561288	ADCY3	0.027	-0.037 $[-0.185, 0.112]$	0.627	
rs10144353	PRKCH	0.043	0.049 $[-0.171, 0.269]$	0.662	
rs1211166	NTRK2	0.029	-0.027 $[-0.179, 0.126]$	0.731	
rs17001561	SCARB2	0.068	-0.020 $[-0.194, 0.154]$	0.824	
rs1780050	NEXN	0.045	0.010 $[-0.117, 0.136]$	0.883	
GS-BMI		0.112	0.151 $[0.128, 0.175]$	7.03E-37	$*$

Part 2 - Linear Age Adjusted Model

SNP	Gene	RI_{50}	β_{MR} $[95\%CI]$	p-value	
rs1421085	FTO	0.477	0.489 $[0.362, 0.617]$	5.61E-14	$*$
rs6235	PCSK1	0.081	0.292 $[0.151, 0.432]$	4.61E-05	$*$
rs7903146	TCF7L2	0.138	0.296 $[0.164, 0.429]$	1.14E-05	$*$
rs11873305	MC4R	0.343	0.613 $[0.300, 0.927]$	1.26E-04	$*$
rs12617233	FANCL	0.128	0.255 $[0.126, 0.384]$	1.08E-04	$*$
rs11672660	GIPR	0.236	0.275 $[0.117, 0.432]$	6.23E-04	$*$
rs997295	MAP2K5	0.145	0.210 $[0.084, 0.337]$	1.13E-03	$*$
rs6499653	FTO	0.117	0.270 $[0.127, 0.414]$	2.13E-04	$*$
rs3824755	NT5C2	0.219	0.331 $[0.123, 0.538]$	0.002	
rs7553158	TNNI3K	0.087	0.219 $[0.095, 0.344]$	5.68E-04	$*$
rs10767664	BDNF	0.238	0.217 $[0.065, 0.370]$	0.005	
rs4788099	SH2B1	0.164	0.202 $[0.065, 0.338]$	0.004	
rs17066846	MC4R	0.136	0.211 $[0.051, 0.371]$	0.010	
rs9356744	CDKAL1	0.075	0.232 $[0.098, 0.365]$	6.58E-04	$*$
rs6453133	HMGCR	0.134	0.198 $[0.061, 0.334]$	0.005	
rs2819347	LMOD1	0.100	0.111 $[-0.023, 0.244]$	0.104	
rs2075650	TOMM40	0.195	-0.003 $[-0.181, 0.175]$	0.972	
rs4946932	FOXO3	0.106	0.092 $[-0.042, 0.226]$	0.178	
rs2984618	TAL1	0.064	0.075 $[-0.055, 0.205]$	0.259	
rs980828	NOS1AP	0.024	0.098 $[-0.028, 0.224]$	0.128	
rs1788826	NPC1	0.108	0.069 $[-0.062, 0.200]$	0.303	

Continued on next page

Table A9 – *Continued from previous page*

SNP	Gene	RI_{50}	β_{MR}	$95\%CI$	p-value	
rs11570094	SPI1	0.093	0.057	$[-0.078, 0.192]$	0.411	
rs7988412	MTIF3	0.093	0.080	$[-0.091, 0.252]$	0.358	
rs2283228	KCNQ1	-0.033	0.094	$[-0.148, 0.336]$	0.445	
rs739564	IQCK	0.111	0.093	$[-0.070, 0.257]$	0.263	
rs526134	USP37	0.054	0.072	$[-0.067, 0.212]$	0.310	
rs2272903	TFAP2B	0.132	0.106	$[-0.092, 0.304]$	0.295	
rs2836754	ETS2	0.107	0.106	$[-0.028, 0.239]$	0.121	
rs2535633	ITIH4	0.011	0.019	$[-0.108, 0.147]$	0.768	
rs11208662	LEPR	0.125	0.085	$[-0.131, 0.300]$	0.442	
rs6232	PCSK1	0.061	0.136	$[-0.122, 0.394]$	0.300	
rs749767	KAT8	0.059	0.070	$[-0.061, 0.202]$	0.294	
rs1561288	ADCY3	0.021	-0.040	$[-0.190, 0.110]$	0.599	
rs10144353	PRKCH	0.052	0.045	$[-0.174, 0.265]$	0.685	
rs1211166	NTRK2	0.039	0.008	$[-0.149, 0.165]$	0.919	
rs17001561	SCARB2	0.058	-0.078	$[-0.263, 0.108]$	0.410	
rs1780050	NEXN	0.038	0.019	$[-0.111, 0.148]$	0.776	
GS-BMI		0.110	0.148	$[0.124, 0.172]$	1.60E-32	$*$

Part 3 - Diabetes Adjusted Model

SNP	Gene	RI_{50}	β_{MR}	$95\%CI$	p-value	
rs1421085	FTO	0.432	0.397	$[0.271, 0.523]$	6.57E-10	$*$
rs6235	PCSK1	0.082	0.289	$[0.153, 0.425]$	3.11E-05	$*$
rs7903146	TCF7L2	0.221	0.408	$[0.276, 0.539]$	1.25E-09	$*$
rs11873305	MC4R	0.315	0.515	$[0.209, 0.821]$	9.62E-04	$*$
rs12617233	FANCL	0.115	0.239	$[0.115, 0.363]$	1.53E-04	$*$
rs11672660	GIPR	0.234	0.288	$[0.136, 0.440]$	2.04E-04	$*$
rs997295	MAP2K5	0.113	0.185	$[0.060, 0.309]$	0.004	
rs6499653	FTO	0.109	0.227	$[0.083, 0.372]$	0.002	
rs3824755	NT5C2	0.205	0.304	$[0.107, 0.500]$	0.002	
rs7553158	TNNI3K	0.097	0.183	$[0.058, 0.307]$	0.004	
rs10767664	BDNF	0.240	0.215	$[0.065, 0.365]$	0.005	
rs4788099	SH2B1	0.148	0.188	$[0.053, 0.322]$	0.006	
rs17066846	MC4R	0.112	0.164	$[0.010, 0.319]$	0.037	
rs9356744	CDKAL1	0.094	0.235	$[0.101, 0.368]$	5.54E-04	$*$
rs6453133	HMGCR	0.109	0.166	$[0.033, 0.299]$	0.014	
rs2819347	LMOD1	0.091	0.099	$[-0.031, 0.230]$	0.136	
rs2075650	TOMM40	0.200	0.014	$[-0.167, 0.194]$	0.882	
rs4946932	FOXO3	0.085	0.080	$[-0.056, 0.216]$	0.250	

Continued on next page

Table A9 – *Continued from previous page*

SNP	Gene	RI_{50}	β_{MR}	$95\%CI$	p-value	
rs2984618	TAL1	0.060	0.080	$[-0.044, 0.205]$	0.205	
rs980828	NOS1AP	0.031	0.120	$[-0.003, 0.242]$	0.056	
rs1788826	NPC1	0.090	0.080	$[-0.047, 0.206]$	0.218	
rs11570094	SPI1	0.085	0.050	$[-0.079, 0.180]$	0.448	
rs7988412	MTIF3	0.079	0.057	$[-0.117, 0.230]$	0.523	
rs2283228	KCNQ1	-0.019	0.158	$[-0.086, 0.402]$	0.205	
rs739564	IQCK	0.100	0.058	$[-0.104, 0.219]$	0.484	
rs526134	USP37	0.072	0.056	$[-0.078, 0.191]$	0.411	
rs2272903	TFAP2B	0.110	0.078	$[-0.124, 0.280]$	0.447	
rs2836754	ETS2	0.080	0.088	$[-0.046, 0.222]$	0.200	
rs2535633	ITIH4	0.025	0.053	$[-0.071, 0.177]$	0.404	
rs11208662	LEPR	0.134	0.157	$[-0.053, 0.367]$	0.143	
rs6232	PCSK1	0.081	0.138	$[-0.144, 0.421]$	0.338	
rs749767	KAT8	0.068	0.088	$[-0.044, 0.220]$	0.193	
rs1561288	ADCY3	0.022	-0.034	$[-0.177, 0.110]$	0.646	
rs10144353	PRKCH	0.029	0.065	$[-0.164, 0.294]$	0.578	
rs1211166	NTRK2	0.030	0.002	$[-0.150, 0.153]$	0.982	
rs17001561	SCARB2	0.079	-0.033	$[-0.208, 0.141]$	0.709	
rs1780050	NEXN	0.042	0.000	$[-0.125, 0.125]$	0.997	
GS-BMI		0.111	0.136	$[0.112, 0.160]$	1.12E-29	$*$

Part 4 - Every 10^{th} Perecentile (5^{th} to 95^{th})

SNP	Gene	RI_{50}	β_{MR}	$95\%CI$	p-value	
rs1421085	FTO	0.467	0.495	$[0.365, 0.624]$	8.00E-14	$*$
rs6235	PCSK1	0.084	0.336	$[0.194, 0.478]$	3.49E-06	$*$
rs7903146	TCF7L2	0.147	0.314	$[0.177, 0.452]$	7.12E-06	$*$
rs11873305	MC4R	0.384	0.629	$[0.332, 0.926]$	3.26E-05	$*$
rs12617233	FANCL	0.132	0.270	$[0.141, 0.400]$	4.34E-05	$*$
rs11672660	GIPR	0.229	0.316	$[0.161, 0.471]$	6.24E-05	$*$
rs997295	MAP2K5	0.125	0.210	$[0.082, 0.338]$	0.001	
rs6499653	FTO	0.123	0.269	$[0.119, 0.419]$	4.43E-04	$*$
rs3824755	NT5C2	0.218	0.356	$[0.139, 0.573]$	1.30E-03	$*$
rs7553158	TNNI3K	0.093	0.189	$[0.060, 0.318]$	0.004	
rs10767664	BDNF	0.248	0.216	$[0.056, 0.375]$	0.008	
rs4788099	SH2B1	0.168	0.219	$[0.077, 0.361]$	0.002	
rs17066846	MC4R	0.117	0.209	$[0.051, 0.368]$	0.009	
rs9356744	CDKAL1	0.070	0.202	$[0.060, 0.345]$	0.005	
rs6453133	HMGCR	0.128	0.160	$[0.021, 0.299]$	0.024	

Continued on next page

Table A9 – *Continued from previous page*

SNP	Gene	RI_{50}	β_{MR}	$95\%CI$	p-value	
rs2819347	LMOD1	0.102	0.134	$[-0.003, 0.270]$	0.055	
rs2075650	TOMM40	0.275	0.161	$[-0.021, 0.343]$	0.082	
rs4946932	FOXO3	0.091	0.077	$[-0.066, 0.219]$	0.292	
rs2984618	TAL1	0.066	0.107	$[-0.022, 0.237]$	0.104	
rs980828	NOS1AP	0.025	0.092	$[-0.036, 0.219]$	0.158	
rs1788826	NPC1	0.105	0.088	$[-0.046, 0.223]$	0.198	
rs11570094	SPI1	0.098	0.108	$[-0.032, 0.248]$	0.131	
rs7988412	MTIF3	0.112	0.124	$[-0.057, 0.305]$	0.178	
rs2283228	KCNQ1	-0.023	0.080	$[-0.176, 0.336]$	0.540	
rs739564	IQCK	0.121	0.100	$[-0.068, 0.269]$	0.244	
rs526134	USP37	0.059	0.094	$[-0.042, 0.231]$	0.177	
rs2272903	TFAP2B	0.132	0.082	$[-0.118, 0.282]$	0.422	
rs2836754	ETS2	0.087	0.061	$[-0.076, 0.197]$	0.384	
rs2535633	ITIH4	0.023	0.069	$[-0.060, 0.198]$	0.297	
rs11208662	LEPR	0.129	0.123	$[-0.101, 0.347]$	0.283	
rs6232	PCSK1	0.070	0.151	$[-0.131, 0.434]$	0.294	
rs749767	KAT8	0.060	0.086	$[-0.051, 0.222]$	0.217	
rs1561288	ADCY3	0.035	-0.041	$[-0.194, 0.112]$	0.596	
rs10144353	PRKCH	0.051	0.098	$[-0.132, 0.329]$	0.403	
rs1211166	NTRK2	0.042	0.008	$[-0.152, 0.168]$	0.920	
rs17001561	SCARB2	0.078	-0.038	$[-0.220, 0.145]$	0.686	
rs1780050	NEXN	0.045	0.008	$[-0.121, 0.136]$	0.907	
GS-BMI		0.111	0.150	$[0.126, 0.174]$	5.79E-34	$*$

Table A10: BMI was divided into BMI categories, and the effects of SNPs on the risk of overweight, obesity class I, class II and class III relative to normal weight (Controls) were tested using logistic regression. Models were adjusted for age, age-squared, sex and study. OR is the odds ratio and 95%CI are the 95% confidence intervals.

SNP	Gene	RI_{50}	β_{MR}	$95\%CI$	p-value
Normal Weight vs Overweight					
rs1421085	FTO	27437/21507	1.096	$[1.067, 1.125]$	1.32E-11
rs2075650	TOMM40	27273/21339	1.102	$[1.061, 1.144]$	4.01E-07
rs10767664	BDNF	27247/21329	1.061	$[1.028, 1.095]$	2.62E-04
rs4788099	SH2B1	23391/19340	1.034	$[1.005, 1.064]$	0.021
rs4946932	FOXO3	26055/20518	1.034	$[1.004, 1.064]$	0.025
rs11672660	GIPR	26538/20708	1.038	$[1.005, 1.071]$	0.025
rs2819347	LMOD1	27432/21508	1.030	$[1.002, 1.059]$	0.037

Continued on next page

Table A10 – *Continued from previous page*

SNP	Gene	Case/Control	OR	95%CI	p-value
rs12617233	FANCL	27438/21507	1.028	1.001, 1.056	0.045
rs6453133	HMGCR	27394/21474	1.026	0.997, 1.056	0.074
rs17001561	SCARB2	24172/20030	1.033	0.995, 1.073	0.086
rs526134	USP37	23431/19389	1.025	0.996, 1.054	0.087
rs10144353	PRKCH	24005/19863	0.959	0.914, 1.007	0.090
rs11570094	SPI1	27419/21503	1.024	0.996, 1.054	0.097
rs749767	KAT8	25891/20351	1.022	0.995, 1.051	0.113
rs3824755	NT5C2	27434/21508	1.034	0.988, 1.082	0.149
rs11208662	LEPR	27422/21495	1.033	0.988, 1.081	0.153
rs17066846	MC4R	27388/21474	1.024	0.991, 1.059	0.162
rs739564	IQCK	26653/20816	1.025	0.990, 1.061	0.165
rs1788826	NPC1	27435/21504	1.019	0.991, 1.047	0.186
rs7988412	MTIF3	22626/18870	1.023	0.986, 1.061	0.219
rs2984618	TAL1	27417/21483	1.016	0.990, 1.044	0.228
rs7903146	TCF7L2	27437/21508	1.016	0.988, 1.046	0.260
rs9356744	CDKAL1	25837/20317	0.984	0.956, 1.012	0.260
rs6235	PCSK1	27418/21498	0.984	0.955, 1.013	0.274
rs2272903	TFAP2B	27437/21505	1.022	0.980, 1.066	0.306
rs7553158	TNNI3K	27437/21508	1.013	0.987, 1.040	0.335
rs2283228	KCNQ1	26658/20845	0.977	0.926, 1.030	0.381
rs997295	MAP2K5	27429/21505	1.011	0.985, 1.038	0.404
rs2836754	ETS2	24165/20021	1.010	0.982, 1.039	0.470
rs2535633	ITIH4	27422/21491	0.990	0.964, 1.017	0.477
rs1780050	NEXN	27435/21507	1.008	0.982, 1.035	0.539
rs1561288	ADCY3	27436/21506	1.007	0.977, 1.039	0.654
rs1211166	NTRK2	27431/21504	0.994	0.962, 1.026	0.696
rs6232	PCSK1	27435/21505	0.993	0.936, 1.054	0.826
rs6499653	FTO	27303/21389	0.997	0.967, 1.027	0.833
rs980828	NOS1AP	27431/21506	0.998	0.972, 1.024	0.860
rs11873305	MC4R	27437/21507	0.995	0.931, 1.063	0.881
GS-BMI		27434/21507	1.019	1.014, 1.024	2.86E-13
Normal Weight vs Obesity Class I					
rs1421085	FTO	15821/21507	1.167	1.132, 1.204	8.07E-23
rs2075650	TOMM40	15746/21339	1.120	1.071, 1.171	5.33E-07
rs10767664	BDNF	15732/21329	1.093	1.053, 1.134	2.73E-06
rs2819347	LMOD1	15819/21508	1.070	1.036, 1.105	4.24E-05
rs11672660	GIPR	15248/20708	1.068	1.028, 1.109	6.76E-04
rs2836754	ETS2	13258/20021	1.053	1.019, 1.089	2.28E-03
rs6453133	HMGCR	15803/21474	1.053	1.019, 1.089	2.31E-03
rs11873305	MC4R	15820/21507	1.127	1.040, 1.220	3.40E-03
rs4788099	SH2B1	12831/19340	1.049	1.014, 1.085	5.94E-03
rs3824755	NT5C2	15820/21508	1.075	1.020, 1.133	6.95E-03

Continued on next page

Table A10 – *Continued from previous page*

SNP	Gene	Case/Control	OR	95%CI	p-value
rs11570094	SPI1	15815/21503	1.046	1.012, 1.081	8.20E-03
rs749767	KAT8	14899/20351	1.043	1.010, 1.078	0.010
rs7553158	TNNI3K	15820/21508	1.040	1.008, 1.072	0.012
rs997295	MAP2K5	15818/21505	1.039	1.007, 1.071	0.016
rs1788826	NPC1	15819/21504	1.039	1.007, 1.073	0.018
rs739564	IQCK	15387/20816	1.049	1.008, 1.093	0.020
rs11208662	LEPR	15805/21495	1.061	1.007, 1.117	0.027
rs17066846	MC4R	15806/21474	1.044	1.004, 1.085	0.029
rs7988412	MTIF3	12341/18870	1.049	1.004, 1.095	0.030
rs10144353	PRKCH	13188/19863	1.061	1.004, 1.122	0.036
rs6499653	FTO	15762/21389	1.037	1.002, 1.074	0.040
rs17001561	SCARB2	13263/20030	1.046	1.001, 1.093	0.046
rs4946932	FOXO3	14974/20518	1.035	1.000, 1.070	0.049
rs12617233	FANCL	15820/21507	1.030	0.998, 1.062	0.068
rs2984618	TAL1	15814/21483	1.027	0.995, 1.059	0.096
rs1780050	NEXN	15820/21507	1.021	0.990, 1.053	0.191
rs7903146	TCF7L2	15821/21508	1.022	0.989, 1.056	0.201
rs1561288	ADCY3	15819/21506	1.022	0.986, 1.060	0.233
rs526134	USP37	12760/19389	1.019	0.985, 1.054	0.270
rs2272903	TFAP2B	15821/21505	1.019	0.970, 1.070	0.449
rs6235	PCSK1	15812/21498	1.013	0.979, 1.048	0.469
rs980828	NOS1AP	15820/21506	1.010	0.980, 1.042	0.508
rs1211166	NTRK2	15818/21504	1.010	0.972, 1.050	0.600
rs2535633	ITIH4	15816/21491	1.006	0.975, 1.037	0.724
rs9356744	CDKAL1	14868/20317	1.003	0.970, 1.037	0.856
rs6232	PCSK1	15820/21505	0.998	0.931, 1.069	0.946
rs2283228	KCNQ1	15308/20845	1.001	0.941, 1.065	0.977
GS-BMI		15818/21507	1.041	1.035, 1.047	5.80E-42
Normal Weight vs Obesity Class II					
rs1421085	FTO	6545/21507	1.293	1.240, 1.347	4.36E-34
rs11672660	GIPR	6273/20708	1.141	1.083, 1.202	7.45E-07
rs10767664	BDNF	6497/21329	1.122	1.067, 1.181	8.71E-06
rs4788099	SH2B1	5243/19340	1.109	1.059, 1.161	9.33E-06
rs2075650	TOMM40	6510/21339	1.123	1.058, 1.193	1.44E-04
rs7903146	TCF7L2	6545/21508	1.090	1.042, 1.140	1.76E-04
rs1788826	NPC1	6543/21504	1.083	1.037, 1.130	2.69E-04
rs2819347	LMOD1	6544/21508	1.082	1.036, 1.130	3.49E-04
rs2272903	TFAP2B	6545/21505	1.115	1.042, 1.193	1.60E-03
rs6453133	HMGCR	6542/21474	1.072	1.025, 1.121	2.48E-03
rs17066846	MC4R	6538/21474	1.078	1.024, 1.135	4.10E-03
rs3824755	NT5C2	6545/21508	1.105	1.030, 1.185	5.05E-03
rs7553158	TNNI3K	6545/21508	1.059	1.016, 1.103	6.53E-03

Continued on next page

Table A10 – *Continued from previous page*

SNP	Gene	Case/Control	OR	95%CI	p-value
rs11873305	MC4R	6545/21507	1.163	1.043, 1.299	7.11E-03
rs997295	MAP2K5	6542/21505	1.055	1.012, 1.100	0.011
rs11208662	LEPR	6540/21495	1.089	1.015, 1.167	0.016
rs739564	IQCK	6373/20816	1.065	1.008, 1.125	0.025
rs12617233	FANCL	6545/21507	1.046	1.003, 1.091	0.036
rs2836754	ETS2	5409/20021	1.050	1.003, 1.099	0.036
rs4946932	FOXO3	6192/20518	1.049	1.002, 1.098	0.042
rs526134	USP37	5174/19389	1.045	0.999, 1.095	0.058
rs2984618	TAL1	6540/21483	1.041	0.999, 1.085	0.058
rs9356744	CDKAL1	6155/20317	1.034	0.988, 1.083	0.145
rs6232	PCSK1	6545/21505	1.065	0.971, 1.166	0.180
rs749767	KAT8	6157/20351	1.029	0.985, 1.075	0.198
rs7988412	MTIF3	5024/18870	1.038	0.978, 1.101	0.221
rs980828	NOS1AP	6545/21506	1.025	0.984, 1.068	0.235
rs1780050	NEXN	6542/21507	1.024	0.983, 1.068	0.259
rs11570094	SPI1	6544/21503	1.025	0.980, 1.072	0.275
rs6499653	FTO	6528/21389	1.026	0.979, 1.075	0.281
rs6235	PCSK1	6541/21498	1.019	0.973, 1.067	0.428
rs2535633	ITIH4	6543/21491	0.989	0.948, 1.031	0.589
rs2283228	KCNQ1	6304/20845	0.982	0.903, 1.068	0.675
rs1211166	NTRK2	6543/21504	0.994	0.944, 1.046	0.813
rs10144353	PRKCH	5377/19863	1.009	0.934, 1.090	0.813
rs1561288	ADCY3	6545/21506	1.006	0.958, 1.056	0.819
rs17001561	SCARB2	5412/20030	1.001	0.942, 1.063	0.980
GS-BMI		6545/21507	1.058	1.050, 1.066	1.33E-44
Normal Weight vs Obesity Class III					
rs1421085	FTO	3919/21507	1.340	1.273, 1.410	3.18E-29
rs10767664	BDNF	3898/21329	1.147	1.076, 1.223	2.87E-05
rs4788099	SH2B1	3119/19340	1.120	1.058, 1.186	1.04E-04
rs12617233	FANCL	3920/21507	1.108	1.051, 1.167	1.34E-04
rs11672660	GIPR	3802/20708	1.123	1.053, 1.198	4.24E-04
rs7903146	TCF7L2	3919/21508	1.097	1.038, 1.161	1.19E-03
rs6499653	FTO	3912/21389	1.095	1.034, 1.160	1.90E-03
rs4946932	FOXO3	3700/20518	1.092	1.031, 1.157	2.94E-03
rs7553158	TNNI3K	3920/21508	1.077	1.023, 1.133	4.65E-03
rs2272903	TFAP2B	3920/21505	1.128	1.037, 1.229	5.12E-03
rs2984618	TAL1	3919/21483	1.074	1.020, 1.130	7.02E-03
rs997295	MAP2K5	3920/21505	1.073	1.020, 1.130	7.03E-03
rs1788826	NPC1	3919/21504	1.075	1.020, 1.134	7.42E-03
rs11570094	SPI1	3919/21503	1.078	1.020, 1.139	7.77E-03
rs739564	IQCK	3836/20816	1.083	1.011, 1.160	0.023
rs3824755	NT5C2	3920/21508	1.102	1.009, 1.202	0.029

Continued on next page

Table A10 – *Continued from previous page*

SNP	Gene	Case/Control	OR	95%CI	p-value
rs2836754	ETS2	3201/20021	1.064	1.005, 1.127	0.035
rs6235	PCSK1	3914/21498	1.063	1.003, 1.125	0.038
rs6453133	HMGCR	3915/21474	1.061	1.003, 1.122	0.039
rs11873305	MC4R	3920/21507	1.149	1.001, 1.323	0.050
rs17066846	MC4R	3914/21474	1.059	0.993, 1.129	0.079
rs11208662	LEPR	3915/21495	1.069	0.980, 1.164	0.130
rs2075650	TOMM40	3898/21339	1.057	0.982, 1.138	0.139
rs526134	USP37	3105/19389	1.041	0.983, 1.102	0.172
rs1780050	NEXN	3920/21507	1.035	0.983, 1.090	0.191
rs1211166	NTRK2	3917/21504	1.043	0.978, 1.112	0.205
rs2819347	LMOD1	3920/21508	1.034	0.980, 1.092	0.220
rs980828	NOS1AP	3920/21506	1.032	0.981, 1.085	0.230
rs10144353	PRKCH	3180/19863	0.942	0.853, 1.040	0.240
rs9356744	CDKAL1	3686/20317	1.031	0.974, 1.091	0.298
rs7988412	MTIF3	2960/18870	1.039	0.964, 1.119	0.315
rs2283228	KCNQ1	3818/20845	0.950	0.855, 1.054	0.341
rs749767	KAT8	3678/20351	1.027	0.972, 1.084	0.341
rs2535633	ITIH4	3917/21491	1.023	0.971, 1.078	0.383
rs1561288	ADCY3	3920/21506	1.021	0.962, 1.085	0.493
rs17001561	SCARB2	3202/20030	1.025	0.951, 1.104	0.517
rs6232	PCSK1	3920/21505	1.026	0.914, 1.150	0.660
GS-BMI		3920/21507	1.069	1.059, 1.080	1.25E-40

Appendix B

The differences between CQR and UQR

The terms "conditional" and "unconditional" refer to the type of quantiles modeled by each of the two methods. CQR models quantiles of the response variable in the form of a conditional distribution (i.e. $Q_{Y|X_1,X_2}(\tau \mid X_1 = x_1, X_2 = x_2)$, where X_1 and X_2 are two explanatory variables). The coefficients of explanatory variables in CQR models are conditional effects. They are average effects on population quantiles given information on all other explanatory variables. There is a growing consensus in the literature that many researchers have misused CQR by misinterpreting coefficients [278]. Unfortunately, the law of iterative expectations do not apply for conditional quantile functions. Hence, integrating out other conditioning explanatory variables is necessary to obtain an interpretable marginal effect. However, the difficulty shifts from coefficient interpretations to integration methods that may not work (i.e. sparse data) and their computational overheads. Examples of integration approaches are provided by Melly B. and Powell D. [279, 280]. The marginalization of CQR coefficients remains an active research area [281, 282]. Other methods for marginalizing quantile regression estimates include UQR that was introduced by Fripo 2009 [253]. UQR is based on the concept of re-centered influence functions (RIF) that transform the response into a new variable having asymptotic mean and variance statistics equal to that of the sample quantile. Hence, the transformed variable can be modeled as a function of explanatory variables using OLS regression where coefficients are interpreted as marginal effects (i.e. $E\big[RIF(Y,\tau) \mid X_1 = x_1, X_2 = x_2\big] = \beta_0 + \beta_1 x_1 + \beta_2 x_2$) [253]. In short, UQR approximates and models unconditional quantile of the response variable post removal of 'contamination' by the explanatory variables [253]. It is straightforward and is computationally inexpensive

given closed form solutions.

Overall, both CQR and UQR models provide similar unadjusted estimates using univariable models where both estimates are marginal effects. This is because CQR estimates are not conditioned on other explanatory variables that need to be integrated out. The difference between CQR and UQR lies in adjusted estimates using multvariable regression models with two or more explanatory variables.

Motivational Example

Quantile regression estimates of genetic variants across the sample distribution capture the location, scale, and shape of interacting variables. To show this, let's consider simulating the interacting variable, X, from a skew-normal distribution, $SN(\zeta, \omega, \alpha)$, with location ($\zeta$), scale ($\omega$), and shape ($\alpha$) parameters. The mean and variance of X are given as

$$\begin{aligned} \mu_x &= \zeta + \omega\delta\sqrt{2/\pi} \\ \sigma_x^2 &= \omega^2\big(1 - 2\delta^2/\pi\big) \end{aligned} \tag{5.1}$$

where

$$\delta = \alpha/\sqrt{1+\alpha^2} \tag{5.2}$$

The partial residual, ϵ_x, follows the same distribution family of X but with a mean of zero. That is, a skew-normal distribution with a location parameter, $\zeta = -\omega\delta\sqrt{2/\pi}$, such that $\mu_{\epsilon_x} = 0$. We simulated independent and identically distributed (i.i.d.) samples from X, and G, with error ϵ. The response , Y, is generated as given in equation 4.1 with coefficients $\beta_1 = 1$, $\beta_2 = 0$, $\beta_3 = 1, 0$, or -1 under different scale and shape parameters for the interacting variable X. Interaction effects with the same direction as the marginal effects are called synergistic interaction effects (e.g. $\beta_1 = 1$ and $\beta_3 = 1$), where as interactions with opposite direction of effects are called antagonistic interaction effects (e.g. $\beta_1 = 1$ and $\beta_3 = -1$). The sample size was set to $n = 10000$. G and ϵ are simulated from a binomial with minor allele frequency (MAF) of 0.5 and a standard normal distribution, respectively. Lastly, CQR and UQR were fitted and compared to the truth given in equation 4.5. Note that CQR and UQR produce similar estimates for univeriable models (Supplementary Material). Figure B1 shows CQR and UQR estimates under an unadjusted interacting variable X with different scale and shape parameters. An increase in the scale parameter of X results in a larger slope of $\beta(\tau)$ with τ, while the increasing the skewness shifts $\beta(\tau)$ vertically corresponding to the direction of the interaction effect as $\omega\delta\sqrt{2/\pi}$. Note that in the case of perfect antagonistic interactions where $\beta_1 = -\beta_3$, the resulting QR estimates correspond to a shift in interacting variable ($\omega\delta\sqrt{2/\pi}$). This can be easily seen if we consider the effect of such interactions on the heterosdastic model in equation 4.4 where y_i becomes $\big(\beta_0 + \beta_1\mu_x\big) + \big(\beta_2 + \beta_3\mu_x\big)g_i + \epsilon_i$.

Hence, QR curves characterize the distribution of interacting variables(s), where we note that skewness results in vertical shifts similar to marginal effects.

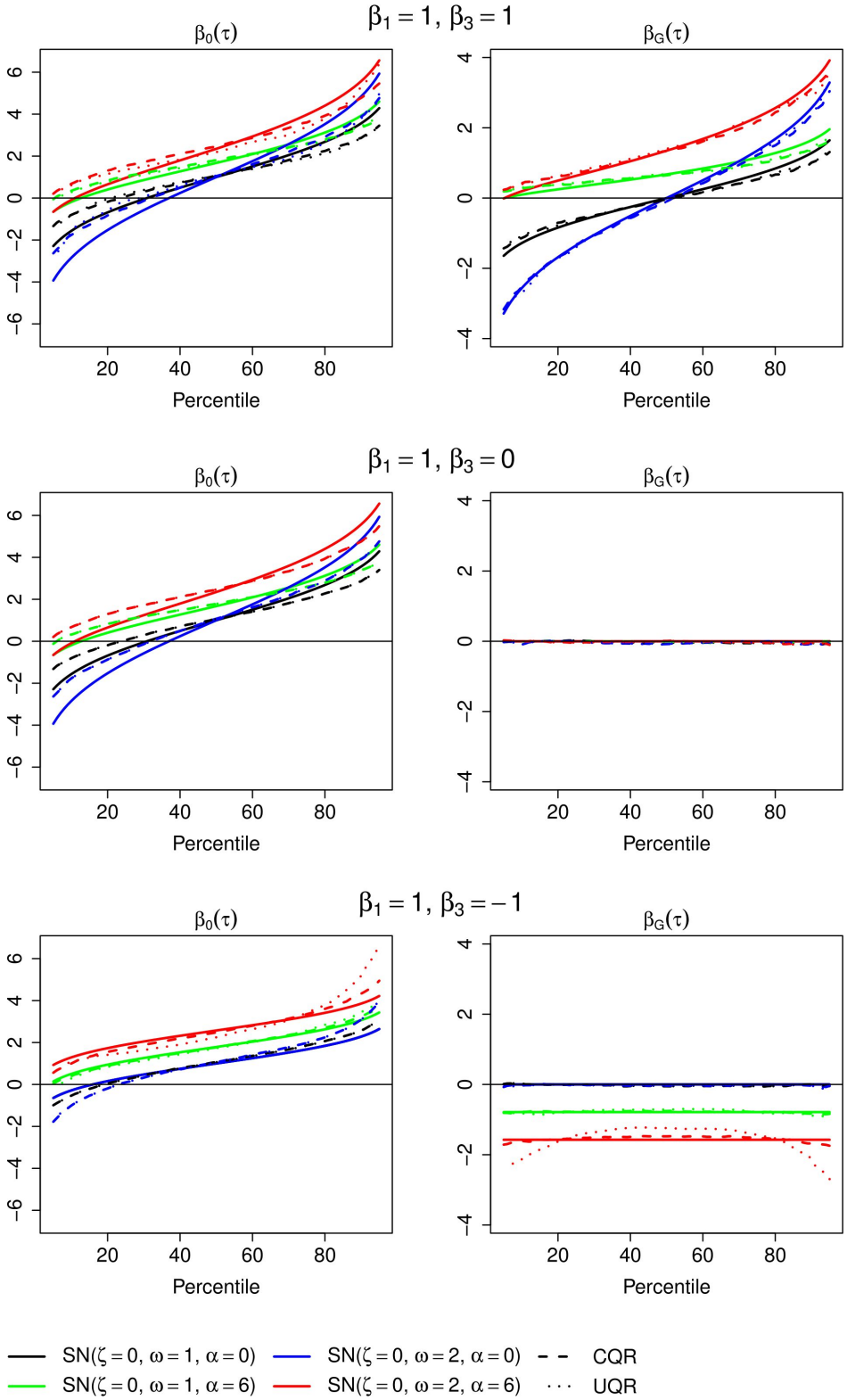


Figure B1: Genetic effects across percentiles for interacting variables with different scale and shape parameters.

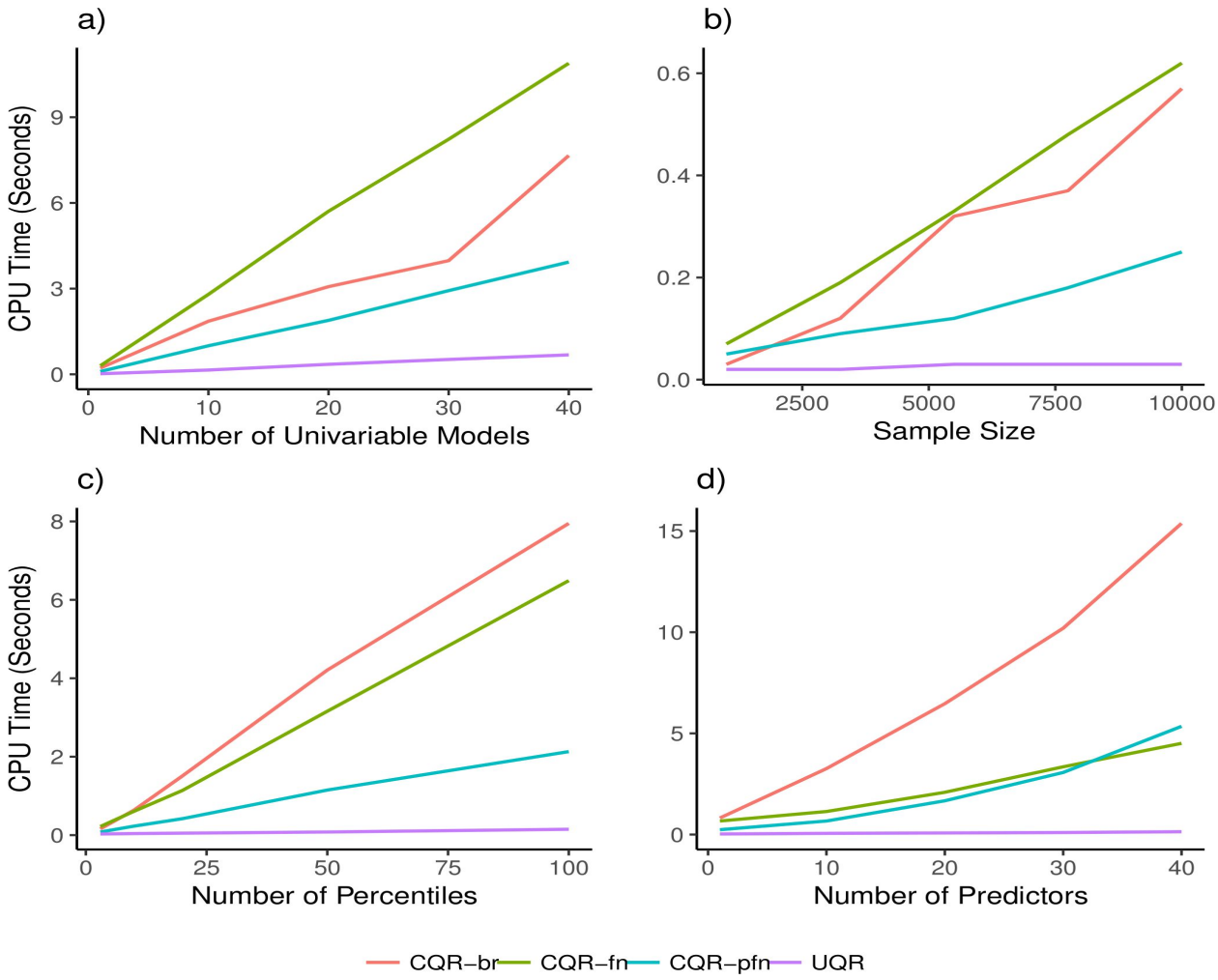


Figure B2: Computational Efficiency of UQR over CQR. The computational time for UQR scales well with all number of snps, sample size, percentiles, and covariates compared to CQR. CPU time for UQR included time required for RIF transformation.

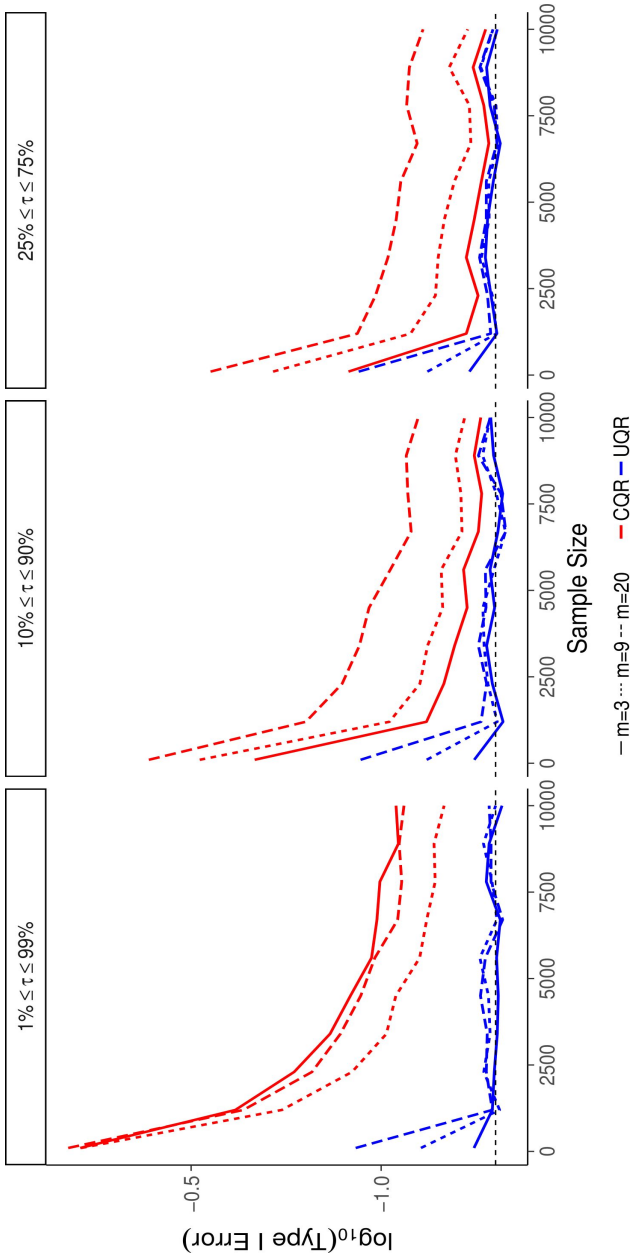


Figure B3: Type I error rate for the number and range of percentiles by sample size. UQR estimates (blue) are more asymptoticly efficient than CQR estimates (red)

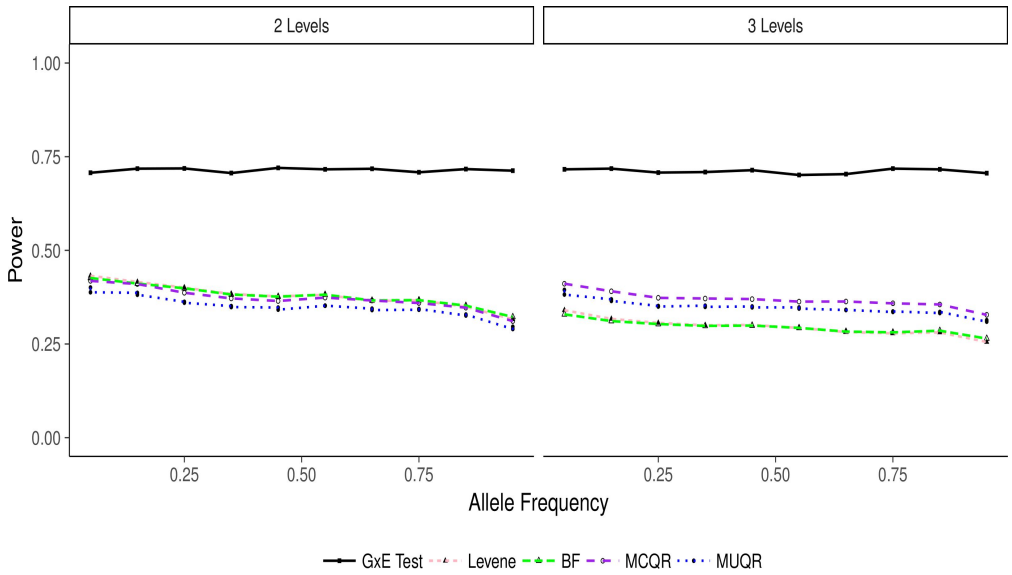


Figure B4: Power of detecting interaction effects for 4 and 5 genotype group levels.

Multiple Interactions

The formulation of heteroscedasticity due to unadjusted interactions as given in equation 4.5 can be generalized further for a set of k independent interacting variables in matrix form as

$$Q_Y(\tau \mid G = g) = A\beta(\tau) \tag{5.3}$$

where A is the design matrix $\begin{pmatrix} \mathbf{1} & G' \end{pmatrix}$ and

$$\beta(\tau) = \begin{pmatrix} \beta_0 + \sum_{j=1}^{k} \beta_{x_j}\mu_{x_j} \\ \beta_g + \sum_{j=1}^{k} \beta_{int_j}\mu_{x_j} \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^{k} \beta_{x_j} Q_{\epsilon_j}(\tau) + Q_{\epsilon_{k+1}}(\tau) \\ \sum_{j=1}^{k} \beta_{int_j} Q_{\epsilon_j}(\tau) \end{pmatrix} \tag{5.4}$$

Here, β_g and β_{x_j} are the marginal effects of the genetic variant and k interacting variables, while β_{int_j} are their respective interaction coefficients. The cumulative two-way interactions of k variables results in a linear heteroscedastic function $\mathbf{1}'\boldsymbol{\gamma}$ where γ has elements $\sigma_j(g) = \beta_{x_j} + \beta_{int_j} g$.

We can further break down the independence assumption between interacting variables given a joint CDF $F_{\boldsymbol{X}}$ with their corresponding mean vector and variance-covariance matrix $\boldsymbol{\Sigma_X}$. The multivariate density of their corresponding errors $\boldsymbol{\epsilon_X}$, $F_{\boldsymbol{\epsilon_X}}$, captures the same shape and scale of $F_{\boldsymbol{X}}$ with a mean vector of zero and variance-covariance matrix $\boldsymbol{\Sigma_X}$. In this case, $\beta_G(\tau)$ changes with the multivariate quantile function $Q_{\boldsymbol{\epsilon_X}}$ [283]. This formulation highlights that QR estimates of $\beta_G(\tau)$ represent contour lines at $\boldsymbol{\tau}$ given the joint density of k interacting variables weighted by degree of interaction effects. In this regard, the modeling of QR estimates across phenotype distributions can be useful for identifying variants with potential interactions.

Meta-regression of QR Estimates Under no Interactions

This section aims to compute the MR parameter estimates to verify how the compares with the original linear model of two-way interactions given in equation 4.1. Let's define A and Σ_G as the design matrix and the cross-distribution variance-covariance matrix for $\beta_G(\tau)$ estimates. The closed solution for MR coefficients are given as:

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_M &= \left(A^{'}\boldsymbol{\Sigma}_G^{-1}A\right)A^{'}\boldsymbol{\Sigma}_G^{-1}\boldsymbol{\beta_G(\tau)} \\ &= \begin{pmatrix}\widehat{\beta}_G & \widehat{\beta}_\tau\end{pmatrix}\end{aligned} \tag{5.5}$$

where $\widehat{\beta}_G$ and $\widehat{\beta}_\tau$ correspond to the variant's marginal and slope for percentiles respectively. Under the null hypothesis of no interactions($H_0 : \beta_3 = 0$), $\beta_G(\tau)^{'} = \begin{pmatrix}\beta_g & \cdots & \beta_g\end{pmatrix}$. The design matrix of MR, A is given by

$$A = \begin{pmatrix}1 & \tau_1 \\ \vdots & \vdots \\ 1 & \tau_m\end{pmatrix} \tag{5.6}$$

Let's further denote the inverse matrix of Σ_G as

$$\Sigma^{-1} = \begin{pmatrix}w_{11} & \cdots & w_{1m} \\ \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mm}\end{pmatrix} \tag{5.7}$$

The initial matrix computations as parts can be given as:

$$\begin{aligned}A^{'}\Sigma^{-1}A &= \begin{pmatrix}\sum_{i=1}^{m} w_{i1} & \cdots & \sum_{i=1}^{mm} w_{im} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{m} \tau_i w_{i1} & \cdots & \sum_{i=1}^{mm} \tau_i w_{im}\end{pmatrix}\begin{pmatrix}1 & \tau_1 \\ \vdots & \vdots \\ 1 & \tau_m\end{pmatrix} \\ &= \begin{pmatrix}\sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij} & \sum_{i=1}^{m}\tau_i\sum_{j=1}^{m} w_{ij} \\ \sum_{i=1}^{m}\sum_{j=1}^{m}\tau_i w_{ij} & \sum_{i=1}^{m}\tau_i\sum_{j=1}^{m}\tau_i w_{ij}\end{pmatrix}\end{aligned} \tag{5.8}$$

Note that

$$(A^{'}\Sigma^{-1}A)^{-1} = \frac{1}{C}\begin{pmatrix} \sum_{i=1}^{m}\tau_i \sum_{j=1}^{m}\tau_i w_{ij} & -\sum_{i=1}^{m}\tau_i \sum_{j=1}^{m} w_{ij} \\ -\sum_{i=1}^{m}\sum_{j=1}^{m}\tau_i w_{ij} & \sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij} \end{pmatrix} \tag{5.9}$$

where $C = \big(\sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij}\big)\big(\sum_{i=1}^{m}\tau_i\sum_{j=1}^{m}\tau_i w_{ij}\big) - \big(\sum_{i=1}^{m}\tau_i\sum_{j=1}^{m} w_{ij}\big)\big(\sum_{i=1}^{m}\sum_{j=1}^{m}\tau_i w_{ij}\big)$

Furthermore,

$$\begin{aligned} A^{'}\Sigma^{-1}A\beta(\tau) &= \begin{pmatrix} \sum_{i=1}^{m} w_{i1} & \cdots & \sum_{i=1}^{mm} w_{im} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{m}\tau_i w_{i1} & \cdots & \sum_{i=1}^{mm}\tau_i w_{im} \end{pmatrix}\begin{pmatrix} \beta_g \\ \vdots \\ \beta_g \end{pmatrix} \\ &= \begin{pmatrix} \beta_g \sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij} \\ \beta_g \sum_{i=1}^{m}\sum_{j=1}^{m}\tau_i w_{ij} \end{pmatrix} \end{aligned} \tag{5.10}$$

Hence,

$$\begin{aligned} \widehat{\beta}_M &= (A^{'}\Sigma^{-1}A)^{-1}A^{'}\Sigma^{-1}A\beta(\tau) \\ &= \frac{1}{C}\begin{pmatrix} \sum_{i=1}^{m}\tau_i\sum_{j=1}^{m}\tau_i w_{ij} & -\sum_{i=1}^{m}\tau_i\sum_{j=1}^{m} w_{ij} \\ -\sum_{i=1}^{m}\sum_{j=1}^{m}\tau_i w_{ij} & \sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij} \end{pmatrix}\begin{pmatrix} \beta_g \sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij} \\ \beta_g \sum_{i=1}^{m}\sum_{j=1}^{m}\tau_i w_{ij} \end{pmatrix} \\ &= \frac{1}{C}\begin{pmatrix} \beta_g\big[\big(\big(\sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij}\big)\big(\sum_{i=1}^{m}\tau_i\sum_{j=1}^{m}\tau_i w_{ij}\big) - \big(\sum_{i=1}^{m}\tau_i\sum_{j=1}^{m} w_{ij}\big)\big(\sum_{i=1}^{m}\sum_{j=1}^{m}\tau_i w_{ij}\big)\big] \\ \beta_g\big[-\big(\sum_{i=1}^{m}\sum_{j=1}^{m}\tau_i w_{ij}\big)\big(\sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij}\big) + \big(\sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij}\big)\big(\sum_{i=1}^{m}\sum_{j=1}^{m}\tau_i w_{ij}\big)\big] \end{pmatrix} \\ &= \frac{1}{C}\begin{pmatrix} \beta_g C \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \beta_g \\ 0 \end{pmatrix} \end{aligned} \tag{5.11}$$

Bibliography

[1] M. Ng, T. Fleming, M. Robinson, B. Thomson, N. Graetz, C. Margono, E. C. Mullany, S. Biryukov, C. Abbafati, S. F. Abera, *et al.*, "Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the global burden of disease study 2013," *The lancet*, vol. 384, no. 9945, pp. 766–781, 2014.

[2] B. H. Hidaka, "Depression as a disease of modernity: explanations for increasing prevalence," *Journal of affective disorders*, vol. 140, no. 3, pp. 205–214, 2012.

[3] N. R. F. Collaboration *et al.*, "Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4· 4 million participants," *The Lancet*, vol. 387, no. 10027, pp. 1513–1530, 2016.

[4] K. T. Mills, J. D. Bundy, T. N. Kelly, J. E. Reed, P. M. Kearney, K. Reynolds, J. Chen, and J. He, "Global disparities of hypertension prevalence and control," *Circulation*, vol. 134, no. 6, pp. 441–450, 2016.

[5] A. Aguilar, "Hypertension: Global blood pressure trends," *Nature Reviews Nephrology*, 2016.

[6] F. Farzadfar, M. M. Finucane, G. Danaei, P. M. Pelizzari, M. J. Cowan, C. J. Paciorek, G. M. Singh, J. K. Lin, G. A. Stevens, L. M. Riley, *et al.*, "National, regional, and global trends in serum total cholesterol since 1980: systematic analysis of health examination surveys and epidemiological studies with 321 country-years and 3· 0 million participants," *The Lancet*, vol. 377, no. 9765, pp. 578–586, 2011.

[7] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman, "Global cancer statistics," *CA: a cancer journal for clinicians*, vol. 61, no. 2, pp. 69–90, 2011.

[8] J. M. Gohlke, R. Thomas, Y. Zhang, M. C. Rosenstein, A. P. Davis, C. Murphy, K. G. Becker, C. J. Mattingly, and C. J. Portier, "Genetic and environmental pathways to complex diseases," *BMC Systems Biology*, vol. 3, no. 1, p. 46, 2009.

[9] L. Hood and M. Flores, "A personal view on systems medicine and the emergence of proactive p4 medicine: predictive, preventive, personalized and participatory," *New Biotechnology*, vol. 29, no. 6, pp. 613–624, 2012.

[10] B. Maher, "Personal genomes: The case of the missing heritability," *Nature News*, vol. 456, no. 7218, pp. 18–21, 2008.

[11] M. J. Khoury, M. L. Gwinn, R. E. Glasgow, and B. S. Kramer, "A population approach to precision medicine," *American Journal of Preventive Medicine*, vol. 42, no. 6, pp. 639–645, 2012.

[12] G. Taubes and C. C. Mann, "Epidemiology faces its limits," *Science*, vol. 269, no. 5221, p. 164, 1995.

[13] R. J. Loos and E. E. Schadt, "This i believe: gaining new insights through integrating old data," *Frontiers in Genetics*, vol. 3, 2012.

[14] E. E. Schadt and J. L. Björkegren, "New: network-enabled wisdom in biology, medicine, and health care," *Science Translational Medicine*, vol. 4, no. 115, pp. 115rv1–115rv1, 2012.

[15] E. E. Schadt, "Molecular networks as sensors and drivers of common human diseases," *Nature*, vol. 461, no. 7261, pp. 218–223, 2009.

[16] M. Tremblay-Servier, J. Tremblay, P. Hamet, C. Lenfant, M. Kalia, J. E. Manson, B. S. McEwen, L. Getz, R. J. Wurtman, T. B. VanItallie, *et al.*, "The medicine of tomorrow," *Personalized Medicine*, vol. 62, no. 1 Suppl 1, 2013.

[17] B.-J. Hardy, B. Séguin, F. Goodsaid, G. Jimenez-Sanchez, P. A. Singer, and A. S. Daar, "The next steps for genomic medicine: challenges and opportunities for the developing world," *Nature Reviews Genetics*, vol. 9, pp. S23–S27, 2008.

[18] E. R. Mardis, "The $1,000 genome, the 100,000$ analysis?," *Genome Medicine*, vol. 2, no. 11, p. 84, 2010.

[19] Y. Yuan, H. Failmezger, O. M. Rueda, H. R. Ali, S. Gräf, S.-F. Chin, R. F. Schwarz, C. Curtis, M. J. Dunning, H. Bardwell, *et al.*, "Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling," *Science Translational Medicine*, vol. 4, no. 157, pp. 157ra143–157ra143, 2012.

[20] V. Kumar, Y. Gu, S. Basu, A. Berglund, S. A. Eschrich, M. B. Schabath, K. Forster, H. J. Aerts, A. Dekker, D. Fenstermacher, *et al.*, "Radiomics: the process and the challenges," *Magnetic Resonance Imaging*, vol. 30, no. 9, pp. 1234–1248, 2012.

[21] A. Brügmann, M. Eld, G. Lelkaitis, S. Nielsen, M. Grunkin, J. D. Hansen, N. T. Foged, and M. Vyberg, "Digital image analysis of membrane connectivity is a robust measure of her2 immunostains," *Breast Cancer Research and Treatment*, vol. 132, no. 1, pp. 41–49, 2012.

[22] P. Gottret and G. Schieber, "Health transitions, disease burdens, and health expenditure patterns," *Health Financing Revisited: A Practitioners Guide: The International Bank for Reconstruction and Development*, pp. 23–39, 2006.

[23] C. Lu, M. T. Schneider, P. Gubbins, K. Leach-Kemon, D. Jamison, and C. J. Murray, "Public financing of health in developing countries: a cross-national systematic analysis," *The Lancet*, vol. 375, no. 9723, pp. 1375–1387, 2010.

[24] A. Li and D. Meyre, "Jumping on the train of personalized medicine: a primer for non-geneticist clinicians: Part 2. fundamental concepts in genetic epidemiology," *Current Psychiatry Reviews*, vol. 10, no. 2, pp. 101–117, 2014.

[25] A. Li and D. Meyre, "Jumping on the train of personalized medicine a primer for

non- geneticist clinicians part 1. fundamental concepts in molecular genetics," *Current Psychiatry Reviews*, vol. 10, no. 4, pp. 91–100, 2014.

[26] A. Li and D. Meyre, "Jumping on the train of personalized medicine a primer for non-geneticist clinicians part 3. clinical applications in the personalized medicine area," *Current Psychiatry Reviews*, vol. 10, no. 4, pp. 118–130, 2014.

[27] L. Hood, "Systems biology and p4 medicine: past, present, and future," *Rambam Maimonides Medical Journal*, vol. 4, no. 2, 2013.

[28] G. Vecchio, M. Fenech, P. P. Pompa, and N. H. Voelcker, "Lab-on-a-chip-based high-throughput screening of the genotoxicity of engineered nanomaterials," *Small*, vol. 10, no. 13, pp. 2721–2734, 2014.

[29] E. E. Schadt, "The changing privacy landscape in the era of big data," *Molecular Systems Biology*, vol. 8, no. 1, p. 612, 2012.

[30] K. A. Phillips, J. A. Sakowski, J. Trosman, M. P. Douglas, S.-Y. Liang, and P. Neumann, "The economic value of personalized medicine tests: what we know and what we need to know," *Genetics in Medicine*, vol. 16, no. 3, pp. 251–257, 2013.

[31] N. Hekim, Y. Coşkun, A. Sınav, A. H. Abou-Zeid, M. Ağırbaşlı, S. O. Akintola, Ş. Aynacıoğlu, M. Bayram, N. L. Bragazzi, C. Dandara, *et al.*, "Translating biotechnology to knowledge-based innovation, peace, and development? deploy a science peace corpsan open letter to world leaders," *Omics: A Journal of Integrative Biology*, vol. 18, no. 7, pp. 415–420, 2014.

[32] V. Özdemir, K. F. Badr, E. S. Dove, L. Endrenyi, C. J. Geraci, P. J. Hotez, D. Milius, M. Neves-Pereira, T. Pang, C. N. Rotimi, *et al.*, "Crowd-funded micro-grants for genomics and big data: an actionable idea connecting small (artisan) science, infrastructure science, and citizen philanthropy," *Omics: A Journal of Integrative Biology*, vol. 17, no. 4, pp. 161–172, 2013.

[33] E. S. Dove and V. Özdemir, "All the postgenomic world is a stage: the actors and

narrators required for translating pharmacogenomics into public health," *Personalized Medicine*, vol. 10, no. 3, pp. 213–216, 2013.

[34] L. Mbuagbaw, M. L. Van Der Kop, R. T. Lester, H. Thirumurthy, C. Pop-Eleches, C. Ye, M. Smieja, L. Dolovich, E. J. Mills, and L. Thabane, "Mobile phone text messages for improving adherence to antiretroviral therapy (art): an individual patient data meta-analysis of randomised trials," *BMJ Open*, vol. 3, no. 12, p. e003950, 2013.

[35] E. Ostrom, "Coping with tragedies of the commons," *Annual review of political science*, vol. 2, no. 1, pp. 493–535, 1999.

[36] E. Ostrom, *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press, 1990.

[37] R. De Vries, "How can we help? from "sociology in" to "sociology of" bioethics," *J Law Med Ethics*, vol. 32, no. 2, pp. 279–92, 191, 2004.

[38] E. S. Dove and V. Ozdemir, "The epiknowledge of socially responsible innovation," *EMBO Rep*, vol. 15, no. 5, pp. 462–3, 2014.

[39]

[40] J. E. McDermott, J. Wang, H. Mitchell, B. J. Webb-Robertson, R. Hafen, J. Ramey, and K. D. Rodland, "Challenges in biomarker discovery: Combining expert insights with statistical analysis of complex omics data," *Expert Opin Med Diagn*, vol. 7, no. 1, pp. 37–51, 2013.

[41] V. N. Kristensen, O. C. Lingjaerde, H. G. Russnes, H. K. Vollan, A. Frigessi, and A. L. Borresen-Dale, "Principles and methods of integrative genomic analyses in cancer," *Nat Rev Cancer*, vol. 14, no. 5, pp. 299–313, 2014.

[42] J. Shendure and E. Lieberman Aiden, "The expanding scope of dna sequencing," *Nat Biotechnol*, vol. 30, no. 11, pp. 1084–94, 2012.

[43] A. Pal and M. I. McCarthy, "The genetics of type 2 diabetes and its clinical relevance," *Clin Genet*, vol. 83, no. 4, pp. 297–306, 2013.

[44] M. B. Scholz, C. C. Lo, and P. S. Chain, "Next generation sequencing and bioinformatic bottlenecks: the current state of metagenomic data analysis," *Curr Opin Biotechnol*, vol. 23, no. 1, pp. 9–15, 2012.

[45] B. Berger, J. Peng, and M. Singh, "Computational solutions for omics data," *Nature Reviews Genetics*, vol. 14, no. 5, pp. 333–346, 2013.

[46] D. Gomez-Cabrero, I. Abugessaisa, D. Maier, A. Teschendorff, M. Merkenschlager, A. Gisel, E. Ballestar, E. Bongcam-Rudloff, A. Conesa, and J. Tegnr, "Data integration in the era of omics: current and future challenges," *BMC Systems Biology*, vol. 8, no. Suppl 2, 2014.

[47] L. M. McShane, M. M. Cavenagh, T. G. Lively, D. A. Eberhard, W. L. Bigbee, P. M. Williams, J. P. Mesirov, M.-Y. C. Polley, K. Y. Kim, J. V. Tricoli, J. M. G. Taylor, D. J. Shuman, R. M. Simon, J. H. Doroshow, and B. A. Conley, "Criteria for the use of omics-based predictors in clinical trials: explanation and elaboration," *BMC Medicine*, vol. 11, no. 1, 2013.

[48] N. J. Brown, D. A. MacDonald, M. P. Samanta, H. L. Friedman, and J. C. Coyne, "A critical reanalysis of the relationship between genomics and well-being," *Proc Natl Acad Sci U S A*, vol. 111, no. 35, pp. 12705–9, 2014.

[49] G. Wilson, D. A. Aruliah, C. T. Brown, N. P. Chue Hong, M. Davis, R. T. Guy, S. H. D. Haddock, K. D. Huff, I. M. Mitchell, M. D. Plumbley, B. Waugh, E. P. White, and P. Wilson, "Best practices for scientific computing," *PLoS Biol*, vol. 12, no. 1, 2014.

[50]

[51]

[52] E. Marshall, "Human genome 10th anniversary. waiting for the revolution," *Science*, vol. 331, no. 6017, pp. 526–9, 2011.

[53] A. Cesario, C. Auffray, P. Russo, and L. Hood, "P4 medicine needs p4 education," *Curr Pharm Des*, 2014.

[54] M. C. Schatz, B. Langmead, and S. L. Salzberg, "Cloud computing and the dna data race," *Nature Biotechnology*, vol. 28, no. 7, pp. 691–693, 2010.

[55] E. E. Schadt, M. D. Linderman, J. Sorenson, L. Lee, and G. P. Nolan, "Cloud and heterogeneous computing solutions exist today for the emerging big data problems in biology," *Nature Reviews Genetics*, vol. 12, no. 3, p. 224, 2011.

[56] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A view of cloud computing," *Commun. ACM*, vol. 53, no. 4, pp. 50–58, 2010.

[57] V. Marx, "Biology: The big challenges of big data," *Nature*, vol. 498, no. 7453, pp. 255–260, 2013.

[58] S. Hiltemann, H. Mei, M. de Hollander, I. Palli, P. van der Spek, G. Jenster, and A. Stubbs, "Cgtag: complete genomics toolkit and annotation in a cloud-based galaxy," *Gigascience*, vol. 3, no. 1, p. 1, 2014.

[59] B. Liu, R. K. Madduri, B. Sotomayor, K. Chard, L. Lacinski, U. J. Dave, J. Li, C. Liu, and I. T. Foster, "Cloud-based bioinformatics workflow platform for large-scale next-generation sequencing analyses," *Journal of Biomedical Informatics*, vol. 49, pp. 119–33, 2014.

[60] G. Zheng, H. Li, C. Wang, Q. Sheng, H. Fan, S. Yang, B. Liu, J. Dai, R. Zeng, and L. Xie, "A platform to standardize, store, and visualize proteomics experimental data," *Acta Biochim Biophys Sin (Shanghai)*, vol. 41, no. 4, pp. 273–9, 2009.

[61] H. Jo, J. Jeong, M. Lee, and D. H. Choi, "Exploiting gpus in virtual machine for biocloud," *BioMed Research International*, vol. 2013, 2013.

[62] L. S. Yung, C. Yang, X. Wan, and W. Yu, "Gboost: a gpu-based tool for detecting gene-gene interactions in genome-wide case control studies," *Bioinformatics*, vol. 27, no. 9, pp. 1309–10, 2011.

[63] S. A. Manavski and G. Valle, "Cuda compatible gpu cards as efficient hardware accelerators for smith-waterman sequence alignment," *BMC Bioinformatics*, vol. 9, no. Suppl 2, 2008.

[64] D. G. McArt, P. Bankhead, P. D. Dunne, M. Salto-Tellez, P. Hamilton, and S. D. Zhang, "cudamap: a gpu accelerated program for gene expression connectivity mapping," *BMC Bioinformatics*, vol. 14, p. 305, 2013.

[65] M. C. Schatz, C. Trapnell, A. L. Delcher, and A. Varshney, "High-throughput sequence alignment using graphics processing units," *BMC Bioinformatics*, vol. 8, no. 1, 2007.

[66]

[67] J. Fadista and C. Bendixen, "Genomic position mapping discrepancies of commercial snp chips," *PLoS One*, vol. 7, no. 2, 2012.

[68] Z. Merali, "Computational science: ...error," *Nature News*, vol. 467, no. 7317, pp. 775–777, 2010.

[69] S. Robiou-du Pont, A. Li, S. Christie, Z. N. Sohani, and D. Meyre, "Should we have blind faith in bioinformatics software? illustrations from the snap web-based tool," *PLoS One*, vol. 10, no. 3, p. e0118925, 2015.

[70] M. A. Khan, L. M. Soto-Jimenez, T. Howe, A. Streit, A. Sosinsky, and C. D. Stern, "Computational tools and resources for prediction and analysis of gene regulatory regions in the chick genome," *Genesis*, vol. 51, no. 5, pp. 311–24, 2013.

[71] A. P. Heath, M. Greenway, R. Powell, J. Spring, R. Suarez, D. Hanley, C. Bandlamudi, M. E. McNerney, K. P. White, and R. L. Grossman, "Bionimbus: a cloud for managing, analyzing and sharing large genomics datasets," *J Am Med Inform Assoc*, 2014.

[72] R. C. Gentleman, V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A. J. Rossini, G. Sawitzki, C. Smith,

G. Smyth, L. Tierney, J. Y. Yang, and J. Zhang, "Bioconductor: open software development for computational biology and bioinformatics," *Genome Biol*, vol. 5, no. 10, p. R80, 2004.

[73] R. Saito, M. E. Smoot, K. Ono, J. Ruscheinski, P.-L. Wang, S. Lotia, A. R. Pico, G. D. Bader, and T. Ideker, "A travel guide to cytoscape plugins," *Nat Methods*, vol. 9, no. 11, pp. 1069–1076, 2012.

[74] L. Dai, X. Gao, Y. Guo, J. Xiao, and Z. Zhang, "Bioinformatics clouds for big data manipulation," *Biology Direct*, vol. 7, no. 1, 2012.

[75] J. D. Tenenbaum, S. A. Sansone, and M. Haendel, "A sea of standards for omics data: sink or swim?," *J Am Med Inform Assoc*, vol. 21, no. 2, pp. 200–3, 2014.

[76] A. Oberst, C. P. Dillon, R. Weinlich, L. L. McCormick, P. Fitzgerald, C. Pop, R. Hakem, G. S. Salvesen, and D. R. Green, "Catalytic activity of the caspase-8-flip(l) complex inhibits ripk3-dependent necrosis," *Nature*, vol. 471, no. 7338, pp. 363–7, 2011.

[77] R. Clarke, H. W. Ressom, A. Wang, J. Xuan, M. C. Liu, E. A. Gehan, and Y. Wang, "The properties of high-dimensional data spaces: implications for exploring gene and protein expression data," *Nat Rev Cancer*, vol. 8, no. 1, pp. 37–49, 2008.

[78] W. S. Noble, "How does multiple testing correction work?," *Nat Biotechnol*, vol. 27, no. 12, pp. 1135–7, 2009.

[79] S. Dudoit and M. J. v. d. Laan, *Multiple Testing Procedures with Applications to Genomics*. Springer Science & Business Media, 2007.

[80] J. Miller, Rupert G., *Simultaneous Statistical Inference*. Springer New York, 2011.

[81] P. H. Westfall and J. F. Troendle, "Multiple testing with minimal assumptions," *Biom J*, vol. 50, no. 5, pp. 745–55, 2008.

[82] P. H. Westfall, *Resampling-Based Multiple Testing: Examples and Methods for P-Value Adjustment*. John Wiley & Sons, 1993.

[83] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 289–300, 1995.

[84] E. Parkhomenko, D. Tritchler, and J. Beyene, "Genome-wide sparse canonical correlation of gene expression with genotypes," *BMC Proc*, vol. 1 Suppl 1, p. S119, 2007.

[85] F. Yao, J. Coquery, and K. A. Le Cao, "Independent principal component analysis for biologically meaningful dimension reduction of large biological data sets," *BMC Bioinformatics*, vol. 13, p. 24, 2012.

[86] K. A. Le Cao, I. Gonzalez, and S. Dejean, "integromics: an r package to unravel relationships between two omics datasets," *Bioinformatics*, vol. 25, no. 21, pp. 2855–6, 2009.

[87] Y. Fan and C. Y. Tang, "Tuning parameter selection in high dimensional penalized likelihood," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 75, no. 3, pp. 531–552, 2013.

[88] H. Park, F. Sakaori, and S. Konishi, "Robust sparse regression and tuning parameter selection via the efficient bootstrap information criteria," *Journal of Statistical Computation and Simulation*, vol. 84, no. 7, pp. 1596–1607, 2013.

[89] P. Bhlmann and S. v. d. Geer, *Statistics for High-Dimensional Data: Methods, Theory and Applications*. Springer Science & Business Media, 2011.

[90] C.-H. Zhang and J. Huang, "The sparsity and bias of the lasso selection in high-dimensional linear regression," *The Annals of Statistics*, vol. 36, no. 4, pp. 1567–1594, 2008.

[91] S. Sass, F. Buettner, N. S. Mueller, and F. J. Theis, "A modular framework for gene set analysis integrating multilevel omics data," *Nucleic Acids Res*, vol. 41, no. 21, pp. 9622–9633, 2013.

[92]

[93] S. Isci, H. Dogan, C. Ozturk, and H. H. Otu, "Bayesian network prior: Network analysis of biological data using external knowledge," *Bioinformatics*, 2013.

[94] P. Reshetova, A. K. Smilde, A. H. C. v. Kampen, and J. A. Westerhuis, "Use of prior knowledge for the analysis of high-throughput transcriptomics and metabolomics data," *BMC Systems Biology*, vol. 8, no. Suppl 2, 2014.

[95] S. Doldec and D. Chessel, "Co-inertia analysis: an alternative method for studying speciesenvironment relationships," *Freshwater Biology*, vol. 31, no. 3, pp. 277–294, 1994.

[96] A. Fagan, A. C. Culhane, and D. G. Higgins, "A multivariate analysis approach to the integration of proteomic and gene expression data," *Proteomics*, vol. 7, no. 13, pp. 2162–71, 2007.

[97] A. C. Culhane, G. Perriere, and D. G. Higgins, "Cross-platform comparison and visualisation of gene expression data using co-inertia analysis," *BMC Bioinformatics*, vol. 4, p. 59, 2003.

[98] C. Meng, B. Kuster, A. C. Culhane, and A. M. Gholami, "A multivariate approach to the integration of multi-omics datasets," *BMC Bioinformatics*, vol. 15, p. 162, 2014.

[99] O. Alter, P. O. Brown, and D. Botstein, "Generalized singular value decomposition for comparative analysis of genome-scale expression data sets of two different organisms," *Proc Natl Acad Sci U S A*, vol. 100, no. 6, pp. 3351–6, 2003.

[100] J. A. Hartigan, "Direct clustering of a data matrix," *Journal of the American Statistical Association*, vol. 67, no. 337, pp. 123–129, 1972.

[101] Y. Cheng and G. M. Church, "Biclustering of expression data," *Proc Int Conf Intell Syst Mol Biol*, vol. 8, pp. 93–103, 2000.

[102] O. A. Tomescu, D. Mattanovich, and G. G. Thallinger, "Integrative omics analysis. a study based on plasmodium falciparum mrna and protein data," *BMC Syst Biol*, vol. 8 Suppl 2, p. S4, 2014.

[103] J. S. Hamid, C. M. T. Greenwood, and J. Beyene, "Weighted kernel fisher discriminant analysis for integrating heterogeneous data," *Computational Statistics & Data Analysis*, vol. 56, no. 6, pp. 2031–2040, 2012.

[104] S. Haider and R. Pal, "Integrated analysis of transcriptomic and proteomic data," *Curr Genomics*, vol. 14, no. 2, pp. 91–110, 2013.

[105] G. Chen, T. G. Gharib, C. C. Huang, J. M. Taylor, D. E. Misek, S. L. Kardia, T. J. Giordano, M. D. Iannettoni, M. B. Orringer, S. M. Hanash, and D. G. Beer, "Discordant protein and mrna expression in lung adenocarcinomas," *Mol Cell Proteomics*, vol. 1, no. 4, pp. 304–13, 2002.

[106] S. P. Gygi, Y. Rochon, B. R. Franza, and R. Aebersold, "Correlation between protein and mrna abundance in yeast," *Mol Cell Biol*, vol. 19, no. 3, pp. 1720–30, 1999.

[107] E. S. Yeung, "Genome-wide correlation between mrna and protein in a single cell," *Angew Chem Int Ed Engl*, vol. 50, no. 3, pp. 583–5, 2011.

[108] T. Van den Bulcke, K. Lemmens, Y. Van de Peer, and K. Marchal, "Inferring transcriptional networks by mining omics data," *Current Bioinformatics*, vol. 1, no. 3, pp. 301–313, 2006.

[109] D. Hwang, J. J. Smith, D. M. Leslie, A. D. Weston, A. G. Rust, S. Ramsey, P. de Atauri, A. F. Siegel, H. Bolouri, J. D. Aitchison, and L. Hood, "A data integration methodology for systems biology: experimental verification," *Proc Natl Acad Sci U S A*, vol. 102, no. 48, pp. 17302–7, 2005.

[110] R. Nagarajan, M. Scutari, and S. Lbre, *Bayesian Networks in R: with Applications in Systems Biology*. Springer Science & Business Media, 2013.

[111] N. Friedman, M. Linial, I. Nachman, and D. Pe'er, "Using bayesian networks to analyze expression data," *J Comput Biol*, vol. 7, no. 3-4, pp. 601–20, 2000.

[112] S. Huang, J. Li, J. Ye, A. Fleisher, K. Chen, T. Wu, and E. Reiman, "A sparse structure learning algorithm for gaussian bayesian network identification from high-dimensional data," *IEEE Trans Pattern Anal Mach Intell*, vol. 35, no. 6, pp. 1328–42, 2013.

[113] M. M. Hoffman, O. J. Buske, J. Wang, Z. Weng, J. A. Bilmes, and W. S. Noble, "Unsupervised pattern discovery in human chromatin structure through genomic segmentation," *Nat Methods*, vol. 9, no. 5, pp. 473–6, 2012.

[114] J. D. Allen, Y. Xie, M. Chen, L. Girard, and G. Xiao, "Comparing statistical methods for constructing large scale gene networks," *PLoS One*, vol. 7, no. 1, 2012.

[115] P. Hu, C. M. Greenwood, and J. Beyene, "Integrative analysis of multiple gene expression profiles with quality-adjusted effect size models," *BMC Bioinformatics*, vol. 6, p. 128, 2005.

[116] S. Yoo, T. Huang, J. D. Campbell, E. Lee, Z. Tu, M. W. Geraci, C. A. Powell, E. E. Schadt, A. Spira, and J. Zhu, "Modmatcher: Multi-omics data matcher for integrative genomic analysis," *PLoS Comput Biol*, vol. 10, no. 8, p. e1003790, 2014.

[117] X. F. Wang, "Joint generalized models for multidimensional outcomes: a case study of neuroscience data from multimodalities," *Biom J*, vol. 54, no. 2, pp. 264–80, 2012.

[118] N. K. Batmanghelich, A. V. Dalca, M. R. Sabuncu, and G. Polina, "Joint modeling of imaging and genetics," *Inf Process Med Imaging*, vol. 23, pp. 766–77, 2013.

[119] P. F. O'Reilly, C. J. Hoggart, Y. Pomyen, F. C. Calboli, P. Elliott, M. R. Jarvelin, and L. J. Coin, "Multiphen: joint model of multiple phenotypes can increase discovery in gwas," *PLoS One*, vol. 7, no. 5, p. e34861, 2012.

[120] J. H. Chu, C. P. Hersh, P. J. Castaldi, M. H. Cho, B. A. Raby, N. Laird, R. Bowler, S. Rennard, J. Loscalzo, J. Quackenbush, and E. K. Silverman, "Analyzing networks of phenotypes in complex diseases: methodology and applications in copd," *BMC Syst Biol*, vol. 8, p. 78, 2014.

[121] S. Grosdidier, A. Ferrer, R. Faner, J. Pinero, J. Roca, B. Cosio, A. Agusti, J. Gea, F. Sanz, and L. I. Furlong, "Network medicine analysis of copd multimorbidities," *Respir Res*, vol. 15, no. 1, p. 111, 2014.

[122] A. L. Barabasi, N. Gulbahce, and J. Loscalzo, "Network medicine: a network-based approach to human disease," *Nature Reviews Genetics*, vol. 12, no. 1, pp. 56–68, 2011.

[123] V. Ozdemir, E. Kolker, P. J. Hotez, S. Mohin, B. Prainsack, B. Wynne, E. Vayena, Y. Coskun, T. Dereli, F. Huzair, A. Borda-Rodriguez, N. L. Bragazzi, J. Faris, R. Ramesar, A. Wonkam, C. Dandara, B. Nair, A. Llerena, K. Kilic, R. Jain, P. J. Reddy, K. Gollapalli, S. Srivastava, and I. Kickbusch, "Ready to put metadata on the post-2015 development agenda? linking data publications to responsible innovation and science diplomacy," *OMICS*, vol. 18, no. 1, pp. 1–9, 2014.

[124] M. Snyder, G. Mias, L. Stanberry, and E. Kolker, "Metadata checklist for the integrated personal omics study: proteomics and metabolomics experiments," *OMICS*, vol. 18, no. 1, pp. 81–5, 2014.

[125] E. Kolker, V. Ozdemir, L. Martens, W. Hancock, G. Anderson, N. Anderson, S. Aynacioglu, A. Baranova, S. R. Campagna, R. Chen, J. Choiniere, S. P. Dearth, W. C. Feng, L. Ferguson, G. Fox, D. Frishman, R. Grossman, A. Heath, R. Higdon, M. H. Hutz, I. Janko, L. Jiang, S. Joshi, A. Kel, J. W. Kemnitz, I. S. Kohane, N. Kolker, D. Lancet, E. Lee, W. Li, A. Lisitsa, A. Llerena, C. Macnealy-Koch, J. C. Marshall, P. Masuzzo, A. May, G. Mias, M. Monroe, E. Montague, S. Mooney, A. Nesvizhskii, S. Noronha, G. Omenn, H. Rajasimha, P. Ramamoorthy, J. Sheehan, L. Smarr, C. V. Smith, T. Smith, M. Snyder, S. Rapole, S. Srivastava, L. Stanberry, E. Stewart, S. Toppo, P. Uetz, K. Verheggen, B. H. Voy, L. Warnich, S. W. Wilhelm, and G. Yandl, "Toward more transparent and reproducible omics studies through a common metadata checklist and data publications," *OMICS*, vol. 18, no. 1, pp. 10–4, 2014.

[126] J. P. Ioannidis and M. J. Khoury, "Improving validation practices in "omics" research," *Science*, vol. 334, no. 6060, pp. 1230–2, 2011.

[127] D. J. Hand, "Deconstructing statistical questions," *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, vol. 157, no. 3, pp. 317–356, 1994.

[128] Y. C. Wang, K. McPherson, T. Marsh, S. L. Gortmaker, and M. Brown, "Health and

economic burden of the projected obesity trends in the usa and the uk," *The Lancet*, vol. 378, no. 9793, pp. 815–825, 2011.

[129] A. Must, J. Spadano, E. H. Coakley, A. E. Field, G. Colditz, and W. H. Dietz, "The disease burden associated with overweight and obesity," *JAMA*, vol. 282, no. 16, pp. 1523–1529, 1999.

[130] J. O. Hill and J. C. Peters, "Environmental contributions to the obesity epidemic," *Science*, vol. 280, no. 5368, pp. 1371–1374, 1998.

[131] A. Misra and L. Khurana, "Obesity and the metabolic syndrome in developing countries," *The Journal of Clinical Endocrinology & Metabolism*, vol. 93, no. 11_supplement_1, pp. s9–s30, 2008.

[132] A. J. Stunkard, J. R. Harris, N. L. Pedersen, and G. E. McClearn, "The body-mass index of twins who have been reared apart," *New England Journal of Medicine*, vol. 322, no. 21, pp. 1483–1487, 1990.

[133] J. Wardle, S. Carnell, C. M. Haworth, and R. Plomin, "Evidence for a strong genetic influence on childhood adiposity despite the force of the obesogenic environment–," *The American journal of clinical nutrition*, vol. 87, no. 2, pp. 398–404, 2008.

[134] M. Pigeyre, F. T. Yazdi, Y. Kaur, and D. Meyre, "Recent progress in genetics, epigenetics and metagenomics unveils the pathophysiology of human obesity," *Clinical science*, vol. 130, no. 12, pp. 943–986, 2016.

[135] D. Gilbert-Diamond and J. H. Moore, "Analysis of gene-gene interactions," *Current protocols in human genetics*, pp. 1–14, 2011.

[136] H. J. Cordell, "Detecting gene–gene interactions that underlie human diseases," *Nature Reviews Genetics*, vol. 10, no. 6, p. 392, 2009.

[137] R. De, S. S. Verma, F. Drenos, E. R. Holzinger, M. V. Holmes, M. A. Hall, D. R. Crosslin, D. S. Carrell, H. Hakonarson, G. Jarvik, *et al.*, "Identifying gene-gene interactions that are highly associated with body mass index using quantitative multifactor dimensionality reduction (qmdr)," *BioData mining*, vol. 8, no. 1, p. 41, 2015.

[138] H. Reddon, J.-L. Guéant, and D. Meyre, "The importance of gene–environment interactions in human obesity," *Clinical Science*, vol. 130, no. 18, pp. 1571–1597, 2016.

[139] B. Rokholm, K. Silventoinen, L. Ängquist, A. Skytthe, K. O. Kyvik, and T. I. Sørensen, "Increased genetic variance of bmi with a higher prevalence of obesity," *PloS one*, vol. 6, no. 6, p. e20816, 2011.

[140] B. Rokholm, K. Silventoinen, P. Tynelius, M. Gamborg, T. I. Sørensen, and F. Rasmussen, "Increasing genetic variance of body mass index during the swedish obesity epidemic," *PloS one*, vol. 6, no. 11, p. e27135, 2011.

[141] S. Walter, I. Mejía-Guevara, K. Estrada, S. Y. Liu, and M. M. Glymour, "Association of a genetic risk score with body mass index across different birth cohorts," *JAMA*, vol. 316, no. 1, pp. 63–69, 2016.

[142] T. O. Kilpeläinen, L. Qi, S. Brage, S. J. Sharp, E. Sonestedt, E. Demerath, T. Ahmad, S. Mora, M. Kaakinen, C. H. Sandholt, *et al.*, "Physical activity attenuates the influence of fto variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children," *PLoS Medicine*, vol. 8, no. 11, p. e1001116, 2011.

[143] T. Ahmad, I.-M. Lee, G. Paré, D. I. Chasman, L. Rose, P. M. Ridker, and S. Mora, "Lifestyle interaction with fat mass and obesity-associated (fto) genotype and risk of obesity in apparently healthy us women," *Diabetes Care*, vol. 34, no. 3, pp. 675–680, 2011.

[144] S. Ahmad, G. Rukh, T. V. Varga, A. Ali, A. Kurbasic, D. Shungin, U. Ericson, R. W. Koivula, A. Y. Chu, L. M. Rose, *et al.*, "Gene× physical activity interactions in obesity: combined analysis of 111,421 individuals of european ancestry," *PLoS Genetics*, vol. 9, no. 7, p. e1003607, 2013.

[145] C. H. Andreasen, K. L. Stender-Petersen, M. S. Mogensen, S. S. Torekov, L. Wegner, G. Andersen, A. L. Nielsen, A. Albrechtsen, K. Borch-Johnsen, S. S. Rasmussen, *et al.*, "Low physical activity accentuates the effect of the fto rs9939609 polymorphism on body fat accumulation," *Diabetes*, vol. 57, no. 1, pp. 95–101, 2008.

[146] B. Xi, C. Wang, L. Wu, M. Zhang, Y. Shen, X. Zhao, X. Wang, and J. Mi, "Influence of physical inactivity on associations between single nucleotide polymorphisms and genetic predisposition to childhood obesity," *American Journal of Epidemiology*, vol. 173, no. 11, pp. 1256–1262, 2011.

[147] A. I. Young, F. Wauthier, and P. Donnelly, "Multiple novel gene-by-environment interactions modify the effect of fto variants on body mass index," *Nature Communications*, vol. 7, p. 12724, 2016.

[148] E. W. Demerath, A. C. Choh, W. Johnson, J. E. Curran, M. Lee, C. Bellis, T. D. Dyer, S. A. Czerwinski, J. Blangero, and B. Towne, "The positive association of obesity variants with adulthood adiposity strengthens over an 80-year period: a gene-by-birth year interaction," *Human Heredity*, vol. 75, no. 2-4, pp. 175–185, 2013.

[149] M. R. Robinson, G. Hemani, C. Medina-Gomez, M. Mezzavilla, T. Esko, K. Shakhbazov, J. E. Powell, A. Vinkhuyzen, S. I. Berndt, S. Gustafsson, *et al.*, "Population genetic differentiation of height and body mass index across europe," *Nature Genetics*, vol. 47, no. 11, p. 1357, 2015.

[150] R. Koenker and K. F. Hallock, "Quantile regression," *Journal of Economic Perspectives*, vol. 15, no. 4, pp. 143–156, 2001.

[151] R. Koenker, "Quantile regression cambridge univ," 2005.

[152] S. G. Thompson and J. Higgins, "How should meta-regression analyses be undertaken and interpreted?," *Statistics in Medicine*, vol. 21, no. 11, pp. 1559–1573, 2002.

[153] M. Borenstein, L. V. Hedges, J. Higgins, and H. R. Rothstein, "Meta-regression," *Introduction to meta-analysis*, pp. 187–203, 2009.

[154] A. D. Association *et al.*, "Diagnosis and classification of diabetes mellitus," *Diabetes Care*, vol. 33, no. Suppl 1, p. S62, 2010.

[155]

[156] S. Anand, G. Dagenais, V. Mohan, R. Diaz, J. Probstfield, R. Freeman, J. Shaw, F. Lanas, A. Avezum, A. Budaj, *et al.*, "Glucose levels are associated with cardiovascular disease and death in an international cohort of normal glycaemic and dysglycaemic men and women: the epidream cohort study," *European Journal of Preventive Cardiology*, vol. 19, no. 4, pp. 755–764, 2012.

[157] K. Musunuru, G. Lettre, T. Young, D. N. Farlow, J. P. Pirruccello, K. G. Ejebe, B. J. Keating, Q. Yang, M.-H. Chen, N. Lapchyk, *et al.*, "Candidate gene association resource (care) clinical perspective: Design, methods, and proof of concept," *Circulation: Genomic and Precision Medicine*, vol. 3, no. 3, pp. 267–275, 2010.

[158] B. J. Keating, S. Tischfield, S. S. Murray, T. Bhangale, T. S. Price, J. T. Glessner, L. Galver, J. C. Barrett, S. F. Grant, D. N. Farlow, *et al.*, "Concept, design and implementation of a cardiovascular gene-centric 50 k snp array for large-scale genomic association studies," *PloS one*, vol. 3, no. 10, p. e3583, 2008.

[159] M. E. Weale, "Quality control for genome-wide association studies," in *Genetic Variation*, pp. 341–372, Springer, 2010.

[160] C. A. Anderson, F. H. Pettersson, G. M. Clarke, L. R. Cardon, A. P. Morris, and K. T. Zondervan, "Data quality control in genetic case-control association studies," *Nature Protocols*, vol. 5, no. 9, p. 1564, 2010.

[161] S. G. Pillai, D. Ge, G. Zhu, X. Kong, K. V. Shianna, A. C. Need, S. Feng, C. P. Hersh, P. Bakke, A. Gulsvik, *et al.*, "A genome-wide association study in chronic obstructive pulmonary disease (copd): identification of two major susceptibility loci," *PLoS Genetics*, vol. 5, no. 3, p. e1000421, 2009.

[162] G. Zhu, L. Warren, J. Aponte, A. Gulsvik, P. Bakke, I. C. G. N. I. Investigators, W. H. Anderson, D. A. Lomas, E. K. Silverman, and S. G. Pillai, "The serpine2 gene is associated with chronic obstructive pulmonary disease in two large populations," *American journal of respiratory and critical care medicine*, vol. 176, no. 2, pp. 167–173, 2007.

[163] E. K. Speliotes, C. J. Willer, S. I. Berndt, K. L. Monda, G. Thorleifsson, A. U. Jackson, H. L. Allen, C. M. Lindgren, J. Luan, R. Mägi, *et al.*, "Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index," *Nature genetics*, vol. 42, no. 11, p. 937, 2010.

[164] A. E. Locke, B. Kahali, S. I. Berndt, A. E. Justice, T. H. Pers, F. R. Day, C. Powell, S. Vedantam, M. L. Buchkovich, J. Yang, *et al.*, "Genetic studies of body mass index yield new insights for obesity biology," *Nature*, vol. 518, no. 7538, p. 197, 2015.

[165] A. R. Wood, T. Esko, J. Yang, S. Vedantam, T. H. Pers, S. Gustafsson, A. Y. Chu, K. Estrada, J. Luan, Z. Kutalik, *et al.*, "Defining the role of common variation in the genomic and biological architecture of adult human height," *Nature Genetics*, vol. 46, no. 11, p. 1173, 2014.

[166] J. MacArthur, E. Bowler, M. Cerezo, L. Gil, P. Hall, E. Hastings, H. Junkins, A. McMahon, A. Milano, J. Morales, *et al.*, "The new nhgri-ebi catalog of published genome-wide association studies (gwas catalog)," *Nucleic Acids Research*, vol. 45, no. D1, pp. D896–D901, 2016.

[167] . G. P. Consortium *et al.*, "A global reference for human genetic variation," *Nature*, vol. 526, no. 7571, p. 68, 2015.

[168] J. Yang, T. Ferreira, A. P. Morris, S. E. Medland, P. A. Madden, A. C. Heath, N. G. Martin, G. W. Montgomery, M. N. Weedon, R. J. Loos, *et al.*, "Conditional and joint multiple-snp analysis of gwas summary statistics identifies additional variants influencing complex traits," *Nature Genetics*, vol. 44, no. 4, p. 369, 2012.

[169] M. Benzinou, J. W. Creemers, H. Choquet, S. Lobbens, C. Dina, E. Durand, A. Guerardel, P. Boutin, B. Jouret, B. Heude, *et al.*, "Common nonsynonymous variants in pcsk1 confer risk of obesity," *Nature Genetics*, vol. 40, no. 8, p. 943, 2008.

[170] G. Thorleifsson, G. B. Walters, D. F. Gudbjartsson, V. Steinthorsdottir, P. Sulem, A. Helgadottir, U. Styrkarsdottir, S. Gretarsdottir, S. Thorlacius, I. Jonsdottir, *et al.*,

"Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity," *Nature Genetics*, vol. 41, no. 1, p. 18, 2009.

[171] M. Rivera, S. Cohen-Woods, K. Kapur, G. Breen, M. Ng, A. W. Butler, N. Craddock, M. Gill, A. Korszun, W. Maier, *et al.*, "Depressive disorder moderates the effect of the fto gene on body mass index," *Molecular Psychiatry*, vol. 17, no. 6, p. 604, 2012.

[172] S. S. Verma, M. De Andrade, G. Tromp, H. Kuivaniemi, E. Pugh, B. Namjou-Khales, S. Mukherjee, G. P. Jarvik, L. C. Kottyan, A. Burt, *et al.*, "Imputation and quality control steps for combining multiple genome-wide datasets," *Frontiers in Genetics*, vol. 5, p. 370, 2014.

[173] A. C. J. Janssens, R. Moonesinghe, Q. Yang, E. W. Steyerberg, C. M. van Duijn, and M. J. Khoury, "The impact of genotype frequencies on the clinical validity of genomic profiling for predicting common chronic diseases," *Genetics in Medicine*, vol. 9, no. 8, p. 528, 2007.

[174] F. Dudbridge, "Power and predictive accuracy of polygenic risk scores," *PLoS Genetics*, vol. 9, no. 3, p. e1003348, 2013.

[175] R. Saxena, B. F. Voight, V. Lyssenko, N. P. Burtt, P. I. de Bakker, H. Chen, J. J. Roix, S. Kathiresan, J. N. Hirschhorn, M. J. Daly, *et al.*, "Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels," *Science*, vol. 316, no. 5829, pp. 1331–1336, 2007.

[176] R. Saxena, C. C. Elbers, Y. Guo, I. Peter, T. R. Gaunt, J. L. Mega, M. B. Lanktree, A. Tare, B. A. Castillo, Y. R. Li, *et al.*, "Large-scale gene-centric meta-analysis across 39 studies identifies type 2 diabetes loci," *The American Journal of Human Genetics*, vol. 90, no. 3, pp. 410–425, 2012.

[177] R. Sladek, G. Rocheleau, J. Rung, C. Dina, L. Shen, D. Serre, P. Boutin, D. Vincent, A. Belisle, S. Hadjadj, *et al.*, "A genome-wide association study identifies novel risk loci for type 2 diabetes," *Nature*, vol. 445, no. 7130, p. 881, 2007.

[178] Y. S. Aulchenko, S. Ripatti, I. Lindqvist, D. Boomsma, I. M. Heid, P. P. Pramstaller, B. W. Penninx, A. C. J. Janssens, J. F. Wilson, T. Spector, *et al.*, "Loci influencing lipid levels and coronary heart disease risk in 16 european population cohorts," *Nature Genetics*, vol. 41, no. 1, p. 47, 2009.

[179] C. N. Spracklen, P. Chen, Y. J. Kim, X. Wang, H. Cai, S. Li, J. Long, Y. Wu, Y. X. Wang, F. Takeuchi, *et al.*, "Association analyses of east asian individuals and trans-ancestry analyses with european individuals reveal new loci associated with cholesterol and triglyceride levels," *Human Molecular Genetics*, vol. 26, no. 9, pp. 1770–1784, 2017.

[180] R. J. Strawbridge, J. Dupuis, I. Prokopenko, A. Barker, E. Ahlqvist, D. Rybin, J. R. Petrie, M. E. Travers, N. Bouatia-Naji, A. S. Dimas, *et al.*, "Genome-wide association identifies nine common variants associated with fasting proinsulin levels and provides new insights into the pathophysiology of type 2 diabetes," *Diabetes*, vol. 60, no. 10, pp. 2624–2634, 2011.

[181] J.-Y. Hwang, X. Sim, Y. Wu, J. Liang, Y. Tabara, C. Hu, K. Hara, C. H. Tam, Q. Cai, Q. Zhao, *et al.*, "Genome-wide association meta-analysis identifies novel variants associated with fasting plasma glucose in east asians," *Diabetes*, p. DB_140563, 2014.

[182] M. C. Ng, D. Shriner, B. H. Chen, J. Li, W.-M. Chen, X. Guo, J. Liu, S. J. Bielinski, L. R. Yanek, M. A. Nalls, *et al.*, "Meta-analysis of genome-wide association studies in african americans provides insights into the genetic architecture of type 2 diabetes," *PLoS Genetics*, vol. 10, no. 8, p. e1004517, 2014.

[183] G. Paré, N. R. Cook, P. M. Ridker, and D. I. Chasman, "On the use of variance per genotype as a tool to identify quantitative trait interaction effects: a report from the women's genome health study," *PLoS Genetics*, vol. 6, no. 6, p. e1000981, 2010.

[184] X. He and F. Hu, "Markov chain marginal bootstrap," *Journal of the American Statistical Association*, vol. 97, no. 459, pp. 783–795, 2002.

[185] M. Kocherginsky, X. He, and Y. Mu, "Practical confidence intervals for regression

quantiles," *Journal of Computational and Graphical Statistics*, vol. 14, no. 1, pp. 41–55, 2005.

[186] R. Koenker and G. Bassett Jr, "Robust tests for heteroscedasticity based on regression quantiles," *Econometrica: Journal of the Econometric Society*, pp. 43–61, 1982.

[187] R. Koenker and J. A. Machado, "Goodness of fit and related inference processes for quantile regression," *Journal of the American Statistical Association*, vol. 94, no. 448, pp. 1296–1310, 1999.

[188] Y. J. Sung, K. Schwander, D. K. Arnett, S. L. Kardia, T. Rankinen, C. Bouchard, E. Boerwinkle, S. C. Hunt, and D. C. Rao, "An empirical comparison of meta-analysis and mega-analysis of individual participant data for identifying gene-environment interactions," *Genetic Epidemiology*, vol. 38, no. 4, pp. 369–378, 2014.

[189] R. D. Riley, P. C. Lambert, and G. Abo-Zaid, "Meta-analysis of individual participant data: rationale, conduct, and reporting," *Bmj*, vol. 340, p. c221, 2010.

[190] B. J. Borah and A. Basu, "Highlighting differences between conditional and unconditional quantile regression approaches through an application to assess medication adherence," *Health Economics*, vol. 22, no. 9, pp. 1052–1070, 2013.

[191] M. B. Lanktree, Y. Guo, M. Murtaza, J. T. Glessner, S. D. Bailey, N. C. Onland-Moret, G. Lettre, H. Ongen, R. Rajagopalan, T. Johnson, *et al.*, "Meta-analysis of dense genecentric association studies reveals common and uncommon variants associated with height," *The American Journal of Human Genetics*, vol. 88, no. 1, pp. 6–18, 2011.

[192] Y. Guo, M. B. Lanktree, K. C. Taylor, H. Hakonarson, L. A. Lange, B. J. Keating, and I. K. S. array BMI Consortium, "Gene-centric meta-analyses of 108 912 individuals confirm known body mass index loci and reveal three novel signals," *Human Molecular Genetics*, vol. 22, no. 1, pp. 184–201, 2012.

[193] R. C. Team *et al.*, "R: A language and environment for statistical computing," 2013.

[194] S. M. Purcell and C. C. Chang, "Plink v1.90b3.42 64-bit," 20 Sep 2016.

[195] C. C. Chang, C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, and J. J. Lee, "Second-generation plink: rising to the challenge of larger and richer datasets," *Gigascience*, vol. 4, no. 1, p. 7, 2015.

[196] S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. De Bakker, M. J. Daly, *et al.*, "Plink: a tool set for whole-genome association and population-based linkage analyses," *The American Journal of Human Genetics*, vol. 81, no. 3, pp. 559–575, 2007.

[197] J. E. Wigginton, D. J. Cutler, and G. R. Abecasis, "A note on exact tests of hardy-weinberg equilibrium," *The American Journal of Human Genetics*, vol. 76, no. 5, pp. 887–893, 2005.

[198] D. Taliun, J. Gamper, and C. Pattaro, "Efficient haplotype block recognition of very long and dense genetic sequences," *BMC Bioinformatics*, vol. 15, no. 1, p. 10, 2014.

[199] J. Yang, S. H. Lee, M. E. Goddard, and P. M. Visscher, "Gcta: a tool for genome-wide complex trait analysis," *The American Journal of Human Genetics*, vol. 88, no. 1, pp. 76–82, 2011.

[200] J. Graffelman and V. Moreno, "The mid p-value in exact tests for hardy-weinberg equilibrium," *Statistical Applications in Genetics and Molecular Biology*, vol. 12, no. 4, pp. 433–448, 2013.

[201] T. R. Gaunt, S. Rodríguez, and I. N. Day, "Cubic exact solutions for the estimation of pairwise haplotype frequencies: implications for linkage disequilibrium analyses and a web tool'cubex'," *BMC Bioinformatics*, vol. 8, no. 1, p. 428, 2007.

[202] R. Koenker, "quantreg: Quantile regression. r package version 5.05," *Vienna: R Foundation*, 2013.

[203] W. Viechtbauer *et al.*, "Conducting meta-analyses in r with the metafor package," *J Stat Softw*, vol. 36, no. 3, pp. 1–48, 2010.

[204] H. W. Borchers, "Pracma: practical numerical math functions," *R package version*, vol. 1, no. 3, 2015.

[205] R. Analytics and S. Weston, "doparallel: Foreach parallel adaptor for the parallel package," *R package version*, vol. 1, no. 8, 2014.

[206] M. Dowle, T. Short, S. Lianoglou, R. Saporta, A. Srinivasan, and E. Antonyan, "data. table: Extension of data. frame," 2014.

[207] G. R. Warnes, B. Bolker, G. Gorjanc, G. Grothendieck, A. Korosec, T. Lumley, D. MacQueen, A. Magnusson, J. Rogers, *et al.*, "gdata: Various r programming tools for data manipulation," *R package version*, vol. 2, no. 3, p. 35, 2014.

[208] A. Abadi, A. Alyass, S. R. du Pont, B. Bolker, P. Singh, V. Mohan, R. Diaz, J. C. Engert, H. C. Gerstein, S. S. Anand, *et al.*, "Penetrance of polygenic obesity susceptibility loci across the body mass index distribution: an update on scaling effects.," *BioRxiv*, p. 225128, 2017.

[209] H. Yaghootkar, M. P. Bancks, S. E. Jones, A. McDaid, R. Beaumont, L. Donnelly, A. R. Wood, A. Campbell, J. Tyrrell, L. J. Hocking, *et al.*, "Quantifying the extent to which index event biases influence large genetic association studies," *Human molecular genetics*, vol. 26, no. 5, pp. 1018–1030, 2017.

[210] K. Silventoinen, S. Sammalisto, M. Perola, D. I. Boomsma, B. K. Cornes, C. Davis, L. Dunkel, M. De Lange, J. R. Harris, J. V. Hjelmborg, *et al.*, "Heritability of adult body height: a comparative study of twin cohorts in eight countries," *Twin Research and Human Genetics*, vol. 6, no. 5, pp. 399–408, 2003.

[211] G. Hemani, J. Yang, A. Vinkhuyzen, J. E. Powell, G. Willemsen, J.-J. Hottenga, A. Abdellaoui, M. Mangino, A. M. Valdes, S. E. Medland, *et al.*, "Inference of the genetic architecture underlying bmi and height with the use of 20,240 sibling pairs," *The American Journal of Human Genetics*, vol. 93, no. 5, pp. 865–875, 2013.

[212] P. T. Williams, "Quantile-specific penetrance of genes affecting lipoproteins, adiposity and height," *PloS One*, vol. 7, no. 1, p. e28764, 2012.

[213] A. Beyerlein, R. von Kries, A. R. Ness, and K. K. Ong, "Genetic markers of obesity risk: stronger associations with body composition in overweight compared to normal-weight children," *Plos One*, vol. 6, no. 4, p. e19057, 2011.

[214] J. A. Mitchell, H. Hakonarson, T. R. Rebbeck, and S. F. Grant, "Obesity-susceptibility loci and the tails of the pediatric bmi distribution," *Obesity*, vol. 21, no. 6, pp. 1256–1260, 2013.

[215] S. I. Berndt, S. Gustafsson, R. Mägi, A. Ganna, E. Wheeler, M. F. Feitosa, A. E. Justice, K. L. Monda, D. C. Croteau-Chonka, F. R. Day, *et al.*, "Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture," *Nature Genetics*, vol. 45, no. 5, p. 501, 2013.

[216] Y. Wei, X. Song, M. Liu, I. Ionita-Laza, and J. Reibman, "Quantile regression in the secondary analysis of case–control data," *Journal of the American Statistical Association*, vol. 111, no. 513, pp. 344–354, 2016.

[217] J. Yang, A. Bakshi, Z. Zhu, G. Hemani, A. A. Vinkhuyzen, S. H. Lee, M. R. Robinson, J. R. Perry, I. M. Nolte, J. V. van Vliet-Ostaptchouk, *et al.*, "Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index," *Nature Genetics*, vol. 47, no. 10, p. 1114, 2015.

[218] H. Reddon, H. C. Gerstein, J. C. Engert, V. Mohan, J. Bosch, D. Desai, S. D. Bailey, R. Diaz, S. Yusuf, S. S. Anand, *et al.*, "Physical activity and genetic predisposition to obesity in a multiethnic longitudinal study," *Scientific Reports*, vol. 6, p. 18672, 2016.

[219] D. Corella, D. K. Arnett, K. L. Tucker, E. K. Kabagambe, M. Tsai, L. D. Parnell, C.-Q. Lai, Y.-C. Lee, D. Warodomwichit, P. N. Hopkins, *et al.*, "A high intake of saturated fatty acids strengthens the association between the fat mass and obesity-associated gene and bmi–," *The Journal of Nutrition*, vol. 141, no. 12, pp. 2219–2225, 2011.

[220] T. Lappalainen, J. Lindström, J. Paananen, J. G. Eriksson, L. Karhunen, J. Tuomilehto, and M. Uusitupa, "Association of the fat mass and obesity-associated (fto) gene

variant (rs9939609) with dietary intake in the finnish diabetes prevention study," *British Journal of Nutrition*, vol. 108, no. 10, pp. 1859–1865, 2012.

[221] Q. Qi, T. O. Kilpeläinen, M. K. Downer, T. Tanaka, C. E. Smith, I. Sluijs, E. Sonestedt, A. Y. Chu, F. Renström, X. Lin, *et al.*, "Fto genetic variants, dietary intake and body mass index: insights from 177 330 individuals," *Human Molecular Genetics*, vol. 23, no. 25, pp. 6961–6972, 2014.

[222] J. Mattei, Q. Qi, F. B. Hu, F. M. Sacks, and L. Qi, "Tcf7l2 genetic variants modulate the effect of dietary fat intake on changes in body composition during a weight-loss intervention–," *The American Journal of Clinical Nutrition*, vol. 96, no. 5, pp. 1129–1136, 2012.

[223] J. Tyrrell, A. R. Wood, R. M. Ames, H. Yaghootkar, R. N. Beaumont, S. E. Jones, M. A. Tuke, K. S. Ruth, R. M. Freathy, G. Davey Smith, *et al.*, "Gene–obesogenic environment interactions in the uk biobank study," *International Journal of Epidemiology*, vol. 46, no. 2, pp. 559–575, 2017.

[224] G. L. G. Consortium *et al.*, "Discovery and refinement of loci associated with lipid levels," *Nature Genetics*, vol. 45, no. 11, pp. 1274–1283, 2013.

[225] D. Thomas, "Gene–environment-wide association studies: emerging approaches," *Nature Reviews Genetics*, vol. 11, no. 4, pp. 259–272, 2010.

[226] J. Flint, "Gwas," *Current Biology*, vol. 23, no. 7, pp. R265–R266, 2013.

[227] T. A. Manolio and F. S. Collins, "Genes, environment, health, and disease: facing up to complexity," *Human Heredity*, vol. 63, no. 2, pp. 63–66, 2007.

[228] C. Stryjecki, A. Alyass, and D. Meyre, "Ethnic and population differences in the genetic predisposition to human obesity," *Obesity Reviews*, 2017.

[229] P. C. Phillips, "Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems," *Nature Reviews Genetics*, vol. 9, no. 11, p. 855, 2008.

[230] J. Marchini, P. Donnelly, and L. R. Cardon, "Genome-wide strategies for detecting multiple loci that influence complex diseases," *Nature Genetics*, vol. 37, no. 4, pp. 413–417, 2005.

[231] M. I. McCarthy, G. R. Abecasis, L. R. Cardon, D. B. Goldstein, J. Little, J. P. Ioannidis, and J. N. Hirschhorn, "Genome-wide association studies for complex traits: consensus, uncertainty and challenges," *Nature Reviews Genetics*, vol. 9, no. 5, pp. 356–369, 2008.

[232] A. Dempfle, A. Scherag, R. Hein, L. Beckmann, J. Chang-Claude, and H. Schäfer, "Gene–environment interactions for complex traits: definitions, methodological requirements and challenges," *European Journal of Human Genetics*, vol. 16, no. 10, pp. 1164–1172, 2008.

[233] M. J. Khoury and S. Wacholder, "Invited commentary: from genome-wide association studies to gene—environment—wide interaction studies—challenges and opportunities," *American Journal of Epidemiology*, vol. 169, no. 2, pp. 227–230, 2008.

[234] N. Risch, R. Herrell, T. Lehner, K.-Y. Liang, L. Eaves, J. Hoh, A. Griem, M. Kovacs, J. Ott, and K. R. Merikangas, "Interaction between the serotonin transporter gene (5-httlpr), stressful life events, and risk of depression: a meta-analysis," *JAMA*, vol. 301, no. 23, pp. 2462–2471, 2009.

[235] E. E. Eichler, J. Flint, G. Gibson, A. Kong, S. M. Leal, J. H. Moore, and J. H. Nadeau, "Missing heritability and strategies for finding the underlying causes of complex disease," *Nature Reviews Genetics*, vol. 11, no. 6, pp. 446–450, 2010.

[236] T. A. Manolio, F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorff, D. J. Hunter, M. I. McCarthy, E. M. Ramos, L. R. Cardon, A. Chakravarti, *et al.*, "Finding the missing heritability of complex diseases," *Nature*, vol. 461, no. 7265, pp. 747–753, 2009.

[237] O. Zuk, E. Hechter, S. R. Sunyaev, and E. S. Lander, "The mystery of missing heritability: Genetic interactions create phantom heritability," *Proceedings of the National Academy of Sciences*, vol. 109, no. 4, pp. 1193–1198, 2012.

[238] L. S. Aiken, S. G. West, and R. R. Reno, *Multiple regression: Testing and interpreting interactions*. Sage, 1991.

[239] J. Luan, M. Wong, N. Day, and N. Wareham, "Sample size determination for studies of gene-environment interaction," *International Journal of Epidemiology*, vol. 30, no. 5, pp. 1035–1040, 2001.

[240] W. J. Gauderman, "Sample size requirements for association studies of gene-gene interaction," *American Journal of Epidemiology*, vol. 155, no. 5, pp. 478–484, 2002.

[241] H. Aschard, "A perspective on interaction effects in genetic association studies," *Genetic Epidemiology*, vol. 40, no. 8, pp. 678–688, 2016.

[242] W. W. Piegorsch, C. R. Weinberg, and J. A. Taylor, "Non-hierarchical logistic models and case-only designs for assessing susceptibility in population-based case-control studies," *Statistics in Medicine*, vol. 13, no. 2, pp. 153–162, 1994.

[243] J. Y. Dai, C. Kooperberg, M. Leblanc, and R. L. Prentice, "Two-stage testing procedures with independent filtering for genome-wide gene-environment interaction," *Biometrika*, vol. 99, no. 4, pp. 929–944, 2012.

[244] A. C. Grandjean, "Dietary intake data collection: challenges and limitations," *Nutrition Reviews*, vol. 70, no. suppl_2, pp. S101–S104, 2012.

[245] R. A. Scott, A. Y. Chu, N. Grarup, A. K. Manning, M.-F. Hivert, D. Shungin, A. Tönjes, A. Yesupriya, D. Barnes, N. Bouatia-Naji, *et al.*, "No interactions between previously associated 2-hour glucose gene variants and physical activity or bmi on 2-hour glucose levels," *Diabetes*, vol. 61, no. 5, pp. 1291–1296, 2012.

[246] R. C. Culverhouse, N. L. Saccone, A. C. Horton, Y. Ma, K. J. Anstey, T. Banaschewski,

M. Burmeister, S. Cohen-Woods, B. Etain, H. L. Fisher, *et al.*, "Collaborative meta-analysis finds no evidence of a strong interaction between stress and 5-httlpr genotype contributing to the development of depression," *Molecular Psychiatry*, 2017.

[247] M. V. Struchalin, A. Dehghan, J. C. Witteman, C. van Duijn, and Y. S. Aulchenko, "Variance heterogeneity analysis for detection of potentially interacting genetic loci: method and its limitations," *BMC Genetics*, vol. 11, no. 1, p. 92, 2010.

[248] X. Sun, R. Elston, N. Morris, and X. Zhu, "What is the significance of difference in phenotypic variability across snp genotypes?," *The American Journal of Human Genetics*, vol. 93, no. 2, pp. 390–397, 2013.

[249] W. J. Gauderman, P. Zhang, J. L. Morrison, and J. P. Lewinger, "Finding novel genes by testing g×e interactions in a genome-wide association study," *Genetic Epidemiology*, vol. 37, no. 6, pp. 603–613, 2013.

[250] T. Vorapongsathorn, S. Taejaroenkul, and C. Viwatwongkasem, "A comparison of type i error and power of bartlett's test, levene's test and cochran's test under violation of assumptions," *Songklanakarin J. Sci. Technol*, vol. 26, no. 4, pp. 537–547, 2004.

[251] R. Koenker and G. Bassett Jr, "Regression quantiles," *Econometrica: Journal of the Econometric Society*, pp. 33–50, 1978.

[252] A. Hagemann, "Cluster-robust bootstrap inference in quantile regression models," *Journal of the American Statistical Association*, vol. 112, no. 517, pp. 446–456, 2017.

[253] S. Firpo, N. M. Fortin, and T. Lemieux, "Unconditional quantile regressions," *Econometrica*, vol. 77, no. 3, pp. 953–973, 2009.

[254] A. C. Harvey, "Estimating regression models with multiplicative heteroscedasticity," *Econometrica: Journal of the Econometric Society*, pp. 461–465, 1976.

[255] L. G. Godfrey, "Testing for multiplicative heteroskedasticity," *Journal of Econometrics*, vol. 8, no. 2, pp. 227–236, 1978.

[256] J. L. Powell, "Least absolute deviations estimation for the censored regression model," *Journal of Econometrics*, vol. 25, no. 3, pp. 303–325, 1984.

[257] T.-H. Kim and H. White, "Estimation, inference, and specification testing for possibly misspecified quantile regression," in *Maximum likelihood estimation of misspecified models: twenty years later*, pp. 107–132, Emerald Group Publishing Limited, 2003.

[258] V. Chernozhukov, "Extremal quantile regression," *Annals of Statistics*, pp. 806–839, 2005.

[259] V. Chernozhukov, I. Fernández-Val, and T. Kaji, "Extremal quantile regression: An overview," *arXiv preprint arXiv:1612.06850*, 2016.

[260] Q. Wang, Q. Lu, and H. Zhao, "A review of study designs and statistical methods for genomic epidemiology studies using next generation sequencing," *Frontiers in Genetics*, vol. 6, 2015.

[261] L. Weiss, *On the asymptotic joint normality of quantiles from a multivariate distribution*. Mathematics Research Center, United States Army, University of Wisconsin, 1963.

[262] S. Lee and P. Sa, "Testing the variance of skewed distributions," *Communications in Statistics-Simulation and Computation*, vol. 27, no. 3, pp. 807–822, 1998.

[263] C. G. Small, "A survey of multidimensional medians," *International Statistical Review/Revue Internationale de Statistique*, pp. 263–277, 1990.

[264] R. Serfling, "Quantile functions for multivariate analysis: approaches and applications," *Statistica Neerlandica*, vol. 56, no. 2, pp. 214–232, 2002.

[265] N. Altman and M. Krzywinski, "Points of significance: Association, correlation and causation.," *Nature Methods*, vol. 12, no. 10, 2015.

[266] H. Levene *et al.*, "Robust tests for equality of variances," *Contributions to probability and statistics*, vol. 1, pp. 278–292, 1960.

[267] M. B. Brown and A. B. Forsythe, "Robust tests for the equality of variances," *Journal of the American Statistical Association*, vol. 69, no. 346, pp. 364–367, 1974.

[268] E. V. Khmaladze, "Martingale approach in the theory of goodness-of-fit tests," *Theory of Probability & Its Applications*, vol. 26, no. 2, pp. 240–257, 1982.

[269] R. Koenker and Z. Xiao, "Inference on the quantile regression process," *Econometrica*, vol. 70, no. 4, pp. 1583–1612, 2002.

[270] H. Aschard, N. Zaitlen, R. M. Tamimi, S. Lindström, and P. Kraft, "A nonparametric test to detect quantitative trait loci where the phenotypic distribution differs by genotypes," *Genetic Epidemiology*, vol. 37, no. 4, pp. 323–333, 2013.

[271] F. J. Massey Jr, "The kolmogorov-smirnov test for goodness of fit," *Journal of the American statistical Association*, vol. 46, no. 253, pp. 68–78, 1951.

[272] A. Hart, "Mann-whitney test is not just a test of medians: differences in spread can be important," *BMJ: British Medical Journal*, vol. 323, no. 7309, p. 391, 2001.

[273] T. A. Knijnenburg, L. F. Wessels, M. J. Reinders, and I. Shmulevich, "Fewer permutations, more accurate p-values," *Bioinformatics*, vol. 25, no. 12, pp. i161–i168, 2009.

[274] C. Chen and Y. Wei, "Computational issues for quantile regression," *Sankhyā: The Indian Journal of Statistics*, pp. 399–417, 2005.

[275] R. Koenker, "Confidence intervals for regression quantiles," in *Asymptotic statistics*, pp. 349–359, Springer, 1994.

[276] B. W. Silverman, *Density estimation for statistics and data analysis*, vol. 26. CRC press, 1986.

[277] I. S. Kohane, J. M. Drazen, and E. W. Campion, "A glimpse of the next 100 years in medicine," 2012.

[278] A. Killewald and J. Bearak, "Is the motherhood penalty larger for low-wage women? a comment on quantile regression," *American Sociological Review*, vol. 79, no. 2, pp. 350–357, 2014.

[279] D. Powell, "Quantile treatment effects in the presence of covariates," 2016.

[280] B. Melly and G. Santangelo, "The changes-in-changes model with covariates," *Universität Bern, Bern*, 2015.

[281] J. D. Angrist and J.-S. Pischke, *Mostly harmless econometrics: An empiricist's companion*. Princeton university press, 2008.

[282] A. A. Gushchin and D. A. Borzykh, "Integrated quantile functions: properties and applications," *arXiv preprint arXiv:1801.00977*, 2018.

[283] P. Chaudhuri, "On a geometric notion of quantiles for multivariate data," *Journal of the American Statistical Association*, vol. 91, no. 434, pp. 862–872, 1996.