ANCIENT DNA FROM PERMAFROST

ANCIENT DNA FROM PERMAFROST:

BERINGIAN SOIL PROVIDES INSIGHTS INTO PCR INHIBITION AND ANTIBIOTIC RESISTANCE

By

CHRISTINE E. KING, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree of

MASTER OF SCIENCE

McMaster University

FULL TITLE: Ancient DNA from permafrost: Beringian soil provides insights into PCR inhibition and antibiotic resistance

AUTHOR: Christine E. King, B.Sc. (McMaster University)

DEGREE: Master of Science (2011)

LOCATION: McMaster University, Hamilton, Ontario, Canada

DEPARTMENT: Biology

SUPERVISOR: Hendrik N. Poinar, PhD

NUMBER OF PAGES: viii, 85

ABSTRACT

The manuscripts included here describe two diverse projects that developed from a metagenomic pilot study of 30,000-year-old Beringian permafrost. To ensure that high yields of DNA were available for deep amplicon sequencing, our initial efforts focused on measuring PCR inhibition in low copy samples. In the first study, we demonstrate how these measurements can be applied to determine optimal combinations of PCR facilitators and to evaluate the tradeoff in DNA recovery that results from sample dilution. We also discuss the methods of inhibition measurement commonly used (amplification efficiency of the samples or quantification cycle shifts of internal positive controls) and show that neither method should be used alone. The results of our metagenomic study and the rigorous procedures used to authenticate the data are presented in the second report, which focuses on our discovery of antibiotic resistance genes in the same samples; here they provide contextual evidence that the site is free of modern contaminants and anthropogenic factors, offering an ideal setting to study the evolution of resistance prior to the antibiotic era. We report a highly diverse collection of ancient DNA sequences that cluster with those found in characterized tetracycline-, vancomycin-, and penicillin-resistant bacteria, and are distinct from homologues with other functions. To verify the functionality of these genes, we focused on the *vanHAX* operon, which functions exclusively as a resistant determinant; not only was the complete operon present, but with successful isolation and expression of *vanA*, we were also able to confirm the activity and specificity of the D-Ala-D-Lac ligase component. These results have serious implications for responsible antibiotic development and deployment, which must consider the potential existence of diverse resistance mechanisms in the environment.

ACKNOWLEDGEMENTS

First and foremost, I owe my deepest gratitude to my supervisor, Dr. Hendrik Poinar. I was drawn to his lab by his exciting research and inspired by his commitment to high quality science. To say 'this work would not have been possible without him' would certainly be an understatement; he has been a constant source of support and encouragement, without which I may not have attempted graduate work, let alone finished it. He is my mentor and my friend, and I am so proud to be his student.

I am also indebted to my colleagues at the McMaster Ancient DNA Center: Melanie Kuch, Carsten Schwarz, Régis Debruyne, Alison Devault, John Okello, Ben Novak, Kirsti Bos, Jodi Barta, Debi Poinar, and Jake Enk. No matter how I write this paragraph, a concise acknowledgement feels terribly insufficient; you've affected me in such different ways, I can only hope I've expressed my appreciation, respect, and admiration personally. I have learned so much from all of you, both in the lab and in life.

A very special thank you must also go to Vanessa D'Costa, Brian Golding, Gerry Wright, and Barb Reuter, who have been endlessly helpful (and patient). It has been an honour and a pleasure to work with you.

Finally, thank you to Mom, Dad, Dustin, Nadine, Milo, and Lucy, who always manage to lift my spirits. I love you all so much. And of course to my dear Grandpa – this is for you.

PREFACE

Two multi-authored papers are included in this thesis; both articles are presented herein with permission from the authors. Minor reformatting was necessary for consistency throughout the thesis. My contributions to each study are detailed below.

PART I: PCR INHIBITION

> King, C.E., R. Debruyne, M. Kuch, C. Schwarz, and H.N. Poinar, *A quantitative approach to detect and overcome PCR inhibition in ancient DNA extracts.* BioTechniques, 2009. **47**(5): p. 941-949

In the above study, I performed the inhibition experiments and analyses and wrote the manuscript. Additional sample types (non-permafrost) were prepared by my co-authors, who also assisted in the experimental design.

PART II: ANTIBIOTIC RESISTANCE

> D'Costa, V.M.*, C.E. King*, L. Kalan, M. Morar, W. Sung, C. Schwarz, D. Froese, G. Zazula, F. Calmels, R. Debruyne, G.B. Golding, H.N. Poinar, and G.D. Wright, *Antibiotic resistance is ancient.* Submitted for publication, 2011.

This report was written by G.D.W., H.N.P., V.M.D., and myself, with contributions from D.F. and G.Z. The initial coring and geologic analyses of the site were performed by D.F., G.Z., and F.C., followed by subsampling by C.S., V.M.D., and myself. All subsequent ancient DNA lab work was my responsibility, as was optimization of the qPCR assays (designed by V.M.D). I performed the 454 sequencing and cloning/sequencing of the *vanHAX* amplicons; V.M.D. cloned and sequenced the *bla* amplicons. Independent replication was carried out by R.D., with further study of the VanA protein by L.K. and M.M. The metagenomic data was processed by W.S., G.B.G., and myself, while analysis of the resistance gene sequences was done by G.B.G. and V.M.D. Individual sections of the SOM were written by the corresponding co-authors.

*V.M.D. and myself share first authorship.

NOTE: Separate reference lists follow each manuscript as well as the SOM in part II. A fourth list at the end of the thesis includes additional references cited in the introduction, conclusion, and in the overview material that precedes both articles. Each list is numbered independently and may include overlapping references.

CONTENTS

FIGURES vii

TABLES viii

INTRODUCTION: PERMAFROST METAGENOMICS 1

PART I: PCR INHIBITION 3

A quantitative approach to detect and overcome PCR inhibition in ancient DNA extracts (King *et al.*, 2009) 4
Introduction 5
Materials and methods 7
Results 10
Discussion 14
Acknowledgments 16
References 17

PART II: ANTIBIOTIC RESISTANCE 31

Antibiotic resistance is ancient (D'Costa *et al.*, 2011) 33
Acknowledgments 41
References 42

Antibiotic resistance is ancient: Supplementary online material (D'Costa *et al.*, 2011) 43
1. Cryostratigraphic Observations 44
2. Sampling Procedure 47
3. DNA Extraction 49
4. PCR Inhibition Assessment 49
5. Contamination Assessment 50
6. 454 Amplicons 54
7. *vanHAX* Cluster 77
8. Other Resistance Genes 86
9. Reports of Ancient Bacteria and Resistance in Permafrost 87
10. References 88

CONCLUSION 90

REFERENCES 91

FIGURES

PART I: PCR INHIBITION

1 Effects of PCR facilitators and dilution factors on optimization criteria 11
2 Effects of PCR facilitators and dilution factors on inhibition levels 12
3 IPC amplification plots 13

PART II: ANTIBIOTIC RESISTANCE

1 Stratigraphic profile and location of Bear Creek site 35
2 Genetic diversity of ancient antibiotic resistance elements 37
3 Ancient vancomycin resistance elements 39
S1 Geochemistry data 45
S2 Detailed stratigraphic profile of Bear Creek 45
S3 Subsampling procedure 48
S4 Contamination assessment 51
S5 De-multiplexing algorithm 55
S6 Top BLAST hits for 12S sequences 56
S7 Top BLAST hits for *rbc*L sequences 58
S8 Top BLAST hits for *trn*L sequences 59
S9 Top BLAST hits for 16S sequences 71
S10 Genetic diversity of permafrost VanX sequences 76
S11 Genetic diversity of permafrost VanHA sequences 79
S12 Genetic diversity of permafrost VanAX sequences 80
S13 Substrate specificity of D-Ala-D-X ligase $VanA_{A2}$ and $VanA_{A4}$ 83
S14 Genetic diversity of permafrost Bla sequences (subset 2) 87

TABLES

PART I: PCR INHIBITION

1 Sample information 7
2 Effects of PCR facilitator combinations at optimal dilution 10
3 Inhibition measurements by extraction method 12
S1 MIQE Checklist 20
S2 Inhibition test results 25

PART II: ANTIBIOTIC RESISTANCE

1 *vanHAX* permutation tests 40
S1 Average major element glass composition (%) of tephra 44
S2 Inhibition test results 50
S3 PCR assays 52
S4 PCR results – 454 amplicons 54
S5 Top BLAST hits for 12S sequences 57
S6 Top BLAST hits for *rbc*L sequences 60
S7 Top BLAST hits for *trn*L sequences 66
S8 Top BLAST hits for 16S sequences 72
S9 qPCR results – *vanHAX* cluster 78
S10 Kinetic characterization of $VanA_{A2}$ and $VanA_{A4}$ ligases 83
S11 Data processing and refinement statistics for $VanA_{A2}$/ATP binary complex 85

INTRODUCTION:

PERMAFROST METAGENOMICS

This thesis centers on two manuscripts that developed from my work with permafrost soil samples. The samples come from Eastern Beringia (modern day Yukon, Canada), a refugium that remained unglaciated during the Pleistocene ice ages and offered passage for the first human migration into North America. In present-day Beringia and other regions of Arctic tundra, primary productivity is low; consequently, this vegetation supports very few mammals. In contrast, the full-glacial environment of Beringia (colder and more arid) supported a tremendous diversity of flora and (mega)fauna. This apparent productivity paradox was summarized nicely by Charles Schweger, who asked "how does one keep a mammoth alive and well under the seemingly impossible conditions of ice-age Beringia?" [1]. Despite decades of research, a comprehensive description of this unique ecosystem – known as the mammoth steppe – has remained a matter of debate.

Much of the paleoecological information we have about Beringia is based on fossil remains, which represent only a small fraction of the organisms that lived there, due to their size or structure, and to the rarity of fossil preservation in general. Although ancient DNA is well preserved in such cold, dry conditions, molecular analyses have either focused on these rare remains, or have been limited (methodologically) to a few sequences from environmental samples (e.g. [2, 3]). Recent technological advances that enable high-throughput DNA sequencing and obviate the need for cloning [4] prompted us to further explore the potential of these environmental samples.

Several technical considerations were required to transition from a conventional metagenomic approach (e.g. deep sequencing of 16S genes from bacteria) to one suited for ancient DNA and not solely focused on bacterial sequences. Of primary concern was the choice between "shotgun" or total DNA sequencing, versus targeted amplicon sequencing; although more laborious, the latter option ultimately provides far more informative data in terms of species identification. This of course depends on careful selection of the target loci, which must be variable enough to enable deep taxonomic resolution, yet have conserved regions suitable for PCR primer binding across a broad range of taxa. On top of this, the amplicons must be small enough to permit amplification of degraded DNA.

As with all ancient DNA projects, careful consideration was also given to maximizing the DNA yields from our samples. This began with the choice of DNA extraction method, which must incorporate both thorough homogenization of the soil contents and efficient purification with limited DNA loss. Subsequently, much attention was given to monitoring PCR inhibition and quantifying the tradeoffs of sample dilution. These efforts are the subject of the first paper: *A quantitative approach to detect and overcome PCR inhibition in ancient DNA extracts.*

Early in the project I was asked to screen the samples for a variety of antibiotic resistance markers, which initiated an in-depth analysis of the vancomycin resistance operon, *vanHAX*. These results are the topic of the second paper: *Antibiotic resistance is ancient.* In this case, the results of the metagenomic study and the rigorous procedures employed to prevent and detect contamination, are crucial to the argument that the antibiotic resistance data is authentic.

PART I:

PCR INHIBITION

Most ancient DNA research has relied on our ability to amplify DNA using PCR; unfortunately, many common impurities can inhibit this reaction, and even prevent amplification entirely. In order to rule out false negative results and otherwise maximize the recovery of ancient DNA, it is essential that PCR inhibition be measured and circumvented as much as possible. The same can be said for quantitative PCR (qPCR) experiments, in which variations in amplification efficiency greatly affect the accuracy of the results [4].

In order to monitor inhibition, the MIQE guidelines [5] suggest measuring amplification efficiency using a dilution series of each sample in qPCR. Others have shown that the same information can be gleaned from individual amplification plots using various models of PCR kinetics [4, 6-18]. Of course, these approaches are of limited utility for samples with few copies of the target sequence, which may not be amplified in the presence of inhibitors and/or are below detection after dilution.

A more suitable approach for monitoring inhibition in such samples involves an internal positive control (IPC) (eg. [19-24]). In this case, the IPC is mixed with the sample and its performance compared to an IPC-only (uninhibited) standard reaction based on the shift in quantification cycle (C_q).

In the following paper, we use an IPC to evaluate PCR inhibition across a variety of samples and DNA extraction methods; with repeated inhibition testing, we demonstrate how DNA recovery can be maximized based on the optimal combination of PCR facilitators and amount of dilution for each sample.

These tests also revealed interesting patterns in the relationship between amplification efficiency and C_q. In most cases we observed high C_q values in IPC reactions with reduced amplification efficiency, as expected; however, many highly efficient reactions also showed high C_q values. This may indicate that the inhibitors do not affect the polymerase or other reaction components, but effectively reduce the number of amplifiable DNA molecules; alternatively, the inhibitors may lose their effect with successive cycles of PCR. We also observed the opposite pattern, in which minimal C_q shifts were produced by inefficient reactions, although this is simply a consequence of the threshold method of calculating C_q. Based on these observations, we advocate the use of both C_q and amplification efficiency measurements, as inhibition may not be detectable with either measurement alone.

A quantitative approach to detect and overcome PCR inhibition in ancient DNA extracts

Christine E. King[1,2], Régis Debruyne[1], Melanie Kuch[1], Carsten Schwarz[1], and Hendrik N. Poinar[1,2,3†]

[1]*McMaster Ancient DNA Centre, Department of Anthropology, McMaster University, Hamilton ON, Canada,* [2]*Department of Biology, McMaster University, Hamilton ON, Canada, and* [3]*Department of Pathology and Molecular Medicine, McMaster University, Hamilton ON, Canada*

BioTechniques 47(5):941-949 (November 2009) doi 10.2144/000113244

Received 13 May 2009; accepted 11 September 2009.

Keywords: PCR inhibitors; quantitative PCR; amplification efficiency; degraded DNA; forensic DNA analysis; PCR kinetics

Supplementary material for this article is available at www.BioTechniques.com/article/113244.

†To whom correspondence should be addressed: Email poinarh@mcmaster.ca

Inhibition is problematic in many applications of PCR, particularly those involving degraded or low amounts of template DNA, when simply diluting the extract is undesirable. Two basic approaches to monitoring inhibition in such samples using real-time or quantitative PCR (qPCR) have been proposed. The first method analyzes the quantification cycle (C_q) deviation of a spiked internal positive control. The second method considers variations in reaction efficiency based on the slopes of individual amplification plots. In combining these methods, we observed increased C_q values together with reduced amplification efficiencies in some samples, as expected; however, deviations from this pattern in other samples support the use of both measurements. Repeat inhibition testing enables optimization of PCR facilitator combinations and sample dilution such that DNA yields and/or quantitative accuracy can be maximized in subsequent PCR runs. Although some trends were apparent within sample types, differences in inhibition levels, optimal reactions conditions, and expected recovery of DNA under these conditions suggest that all samples be routinely tested with this approach.

INTRODUCTION

Since the advent of PCR, many inhibitory substances that interfere with the activity or availability of particular reaction components have been identified (1,2). The presence of inhibitors introduces a number of problems, ranging from reduced amplification capacity and reduced assay sensitivity to complete reaction failure. Thus, in situations where high (or simply consistent) amplicon concentrations or maximal sensitivity are desired, or when drawing conclusions from negative results, it is important that PCR inhibition be detected and circumvented as much as possible. Inhibition of qPCR presents additional concerns, as slight variations in amplification efficiency between samples can drastically affect the accuracy of template quantification (3).

When PCR inhibition is suspected, the simplest course of action is to dilute the template (and inhibitors), and make use of the sensitivity of PCR (1); however, for applications involving heavily degraded or otherwise low-copy templates, this solution is often undesirable—and indeed sometimes impossible—due to the further reduction of template amounts. In these situations, which are commonly encountered in the forensic and ancient DNA fields, routine detection and quantification of PCR inhibition is necessary.

In standard PCR experiments, negative results or unexpectedly low product yields may be indicative of inhibition, provided that the template is known to be present; alternatively, a known amount of non-endogenous DNA can be added to the sample and amplified as an internal positive control (IPC). These controls may be used in qPCR assays as well, enabling quantitative assessments of their performance. Based on modeling individual reaction kinetics and/or the calculation of amplification efficiency, qPCR also allows inhibited samples to be identified without additional IPC amplifications.

Internal positive controls

A number of studies have incorporated IPC assays for inhibition detection (4–9), as have some commercially available systems such as the Quantifiler Human DNA Quantification Kit (Applied Biosystems, Foster City, CA, USA). Typically, problematic samples are identified based on the shift in C_q (ΔC_q) observed relative to an uninhibited reaction. When using this type of control, one assumes that the effects of inhibitors on the IPC are predictive of those on other targets in subsequent PCR assays from the same sample. This assumption has been challenged, as some authors have noted differential susceptibility to PCR inhibition between assays (10,11).

Huggett et al. (11) found quantitatively small but significant differences between assays, with no correlation between the extent of inhibition and any particular characteristic of the template or primer sequences. Ståhlberg et al. (10) hypothesize that this may result from indirect inhibition, through competition for particular reaction components. If, for example, Mg^{2+} ions became limiting by the presence of EDTA in a sample, an assay with a greater Mg^{2+} requirement would be more susceptible to the same level of EDTA inhibition.

Thus, it is important that both IPC and target reactions perform well under (ideally) the same conditions. In particular, the concentrations of components known to facilitate PCR in the presence of inhibitors (e.g., BSA, *Taq* polymerase, etc.) must be the same. Furthermore, IPC and target assays with different enzymatic requirements may not be comparable; for instance, if inhibition of the exonuclease activity of *Taq* is also measured (e.g., by a *Taq*Man IPC assay), target assays performed without such probes may be less susceptible to inhibition than the IPC predicts.

Amplification efficiency

Based on the potential for differential susceptibility to inhibition, the Minimum Information for Publication of Quantitative Real-time PCR Experiments (MIQE) guidelines (12) instead advocate measuring amplification performance using a dilution series of each sample [e.g., Ståhlberg et al. (10)]. As noted previously, this may not be feasible for low-copy templates.

Building on the 2002 publications of Liu and Saint (13,14), various mathematical models that describe individual reaction kinetics have been developed and compared (3,15–25). Regardless of the particular model, it is clear that kinetic outliers can be identified when inhibitors influence amplification efficiency; however, certain inhibitory mechanisms may not be identified this way (20). Furthermore, highly inhibited samples cannot be differentiated from those having no template, or from a combination of low-copy template and moderate inhibition when there is simply no amplification observed.

To the best of our knowledge, no studies have incorporated the use of an IPC and amplification efficiency measurements in order to contrast their relative contribution in monitoring PCR inhibition. In this study, we present a comprehensive approach to detecting and quantifying PCR inhibition that incorporates both features, and recommend its use in routine sample processing (26), particularly when working with low-copy templates. Beyond simply identifying problematic samples, we demonstrate how quantifying inhibition effects can be used to determine an optimal combination of PCR facilitators (and dilution in some cases) for every DNA extract that maximizes quantitative accuracy and/or template recovery.

MATERIALS AND METHODS

Samples and DNA extraction

The tests described herein were performed on 47 DNA extracts from a diverse set of ancient samples (Table 1), including permafrost soil, mammoth bone and hair, and packrat paleofeces. DNA was extracted from the soil (~250 mg) according to Willerslev et al. (27), but using 25 mM TCEP [Tris (2-carboxyethyl) phosphine hydrochloride] in place of β-mercaptoethanol and with a 10-min vortexing step in Lysing Matrix E tubes (MP Biomedicals, Solon, OH, USA) for sample disruption. The extracts were purified either using columns as indicated or using a modification of the protocol of Boom et al. (28). In this case, the extracts were mixed with 4 mL binding buffer [5 M guanidinium thiocyanate (GuSCN), 50 mM Tris-Cl pH 8, 22.5 mM NaCl, 20 mM EDTA, 1.25% Triton-X 100] that had been previously incubated with 50 μL size-fractionated silicon dioxide (30 min with rotation at room temperature). Additional GuSCN was added to maintain a 5 M concentration upon addition of the extract and the pH was adjusted to 4.5–5.5 with glacial acetic acid to maximize binding efficiency (29). Following a minimum 1-h (room temperature) incubation, the silica-bound DNA was washed twice with 1 mL buffer (5 M GuSCN, 50 mM Tris-Cl pH 8, 22.5 mM NaCl) and once with 1 mL 80% ethanol (in 1× TE pH 7.5). The pellet was dried at 56°C for 5 min and the DNA eluted in 100 μL Buffer EB (Qiagen, Hilden, Germany) as per the procedure performed by Willerslev et al. (27). A subset of the soil samples were processed using the UltraClean Soil DNA Isolation Kit (MO BIO Laboratories, Carlsbad, CA, USA) according to the manufacturer's alternative protocol for maximum yields. DNA was extracted from the bone samples (~100 mg) according to Poinar et al. (30). As part of a separate study, the same protocol (without a demineralization step) was used for the feces samples. Although commonly used silica-based methods may offer high DNA purity, we have noted considerable DNA loss in exchange (unpublished results) and wished to evaluate whether an alternative protocol might yield enough DNA to outweigh the effects of increased PCR inhibition. The hair samples (54 mg and 2 mg) were processed as per Gilbert et al. (31).

Table 1. Sample information

Material	Species	Location	Extract	Age[a] (years BP)[b]
Bone	*M. primigenius*	Taimyr Peninsula, Siberia	B1	Unknown
			B2	55,800 ± 4500-2900
Feces	*Neotoma* sp.	Death Valley, CA, USA	F1	1942 ± 35
			F2	Unknown
Hair	*M. primigenius*	Taimyr Peninsula, Siberia	H1	20,380 ± 140
			H2	20,380 ± 140
Soil	...	Yukon, Canada	P1-10	ca. 25,300
			P11-20	ca. 80,000-90,000
			P21-32	ca. 740,000
			P33-34	ca. 80,000-90,000

[a]Ages of soil samples based on proximity to volcanic tephra. [b]BP, before present.

Inhibition assay

To assess the levels of PCR inhibition in the samples, we monitored the effect of each purified extract on the amplification of an IPC during qPCR. As amplification of any endogenous templates would distort the quantification results, we selected an assay targeting the human β-2-microglobulin gene (*B2M*), which should not be present in these extracts, and used an IPC template derived from the *B2M* cDNA sequence (GenBank accession no. NM_004048), such that amplification of any contaminating human genomic DNA could be distinguished. The primers used in this assay (5′–3′: TGACTTTGTCACAGCCCAAGATA and AATCCAAATGCGGCATCTTC) flank two intronic regions (~2 kb total), such that amplification of the cDNA-derived template yields an 85-bp product that is easily differentiated from amplification of longer *B2M* genomic DNA (32). As the assay is primarily used for quantification of human cDNA, the template (5′–3′: GAACCATG*TGACTTTGTCACAGCCCAAGATA*GTTAAGTGGGATC-GAGACATGTAAGCAGCATCATGGCGGTTT*GAAGATGCCGCATTTGGATT*GGATGA) was synthesized to include an internal base modification (underlined; primer binding sites are italicized) so that it may be distinguished from human cDNA contamination by sequencing as well, although this feature is not important for these experiments. We used an available synthetic ssDNA version of this template for convenience. As the same ssDNA template was used in all reactions, we perceive no disadvantages in its use compared to dsDNA template, nor must any corrections be made to the inhibition measurements.

Following gradient optimization (data not shown) of the annealing temperature and Mg^{2+} concentration, the assay demonstrated highly sensitive and reproducible amplification of serial dilutions of the IPC. Additional details are included in the MIQE checklist (Table S1).

In each 20 µL reaction, 12,500 copies of the IPC were amplified with 1× PCR Buffer II (10 mM Tris-Cl pH 8.3 and 50 mM KCl; Applied Biosystems), 2.5 mM $MgCl_2$, 300 nM primers, 400 µM (each) dNTP, 1 U of Ampli*Taq* Gold DNA Polymerase (Applied Biosystems), and 0.167× SYBR Green I solution (Invitrogen, Carlsbad, CA, USA). Following an initial denaturation at 95°C for 7 min, the reactions were subject to 50 cycles of 95°C, 59°C, and 72°C (30 s each), with data collection at the end of the annealing step. To verify that the cDNA-derived template was amplified, the melting temperature (Tm) of the product was monitored during a final cycle of 95°C for 1 min, 55°C for 30 s, and 95°C for 30 s.

Three reaction types were used in these tests: sample reactions, which included the IPC and one DNA extract at 10% of the final PCR volume; standard reactions, which included only the IPC; and no template control (NTC) reactions.

Inhibition measurement

Two measurements of inhibition were determined using the qPCR data for each sample reaction relative to the standard reactions. For the first, the Hill slope of each amplification plot was calculated by fitting a variable-slope sigmoidal dose-response curve to the raw fluorescence data using GraphPad Prism version 5.02 for Windows (GraphPad Software, San Diego, CA, USA). Under a sigmoidal model, this is the only parameter affecting amplification efficiency calculations in the early cycles of PCR (14,24) and these calculations of initial amplification efficiency are highly consistent with those derived from standard curves (19). Since an absolute measure of amplification efficiency was not required in this analysis, we compared the Hill slope values directly; the Hill slope of each sample reaction was expressed as a percentage of the average of the two corresponding standard reactions. For simplicity, this value (E) is referred to in the text as "amplification efficiency" or just "efficiency," with any deviations from 100% due to the effects of PCR inhibitors in the extract; however, we acknowledge that it is in fact a relative measurement of the shape of an amplification plot. For the second measurement, the C_q of each reaction was automatically calculated by the MxPro – Mx3000P QPCR System (Stratagene, La Jolla, CA, USA) using the amplification-based threshold-determination algorithm; C_q shifts (ΔC_q) were measured as the difference between the sample C_q and the average C_q of the two standard reactions.

To quantify the total effect of inhibition and any dilution of the extract on subsequent PCRs, we calculated the expected recovery (ER) value, where ER = (dilution factor $\times$ $2^{\Delta Cq})^{-1}$ $\times$ 100%. For example, from an extract showing ΔC_q = 1 when used directly in PCR (dilution factor = 1), we expect other assays to recover 50% of the true yield. Likewise, for an extract showing no inhibition (ΔC_q = 0) when tested at a 1/10 dilution, we can only expect to recover 10% of the templates available in the original extract.

Circumventing inhibition

For the soil samples, our goal was to optimize the PCR conditions (in terms of amplification facilitators) for each extract in order to minimize inhibition and reduce, if not eliminate, the need to dilute these extracts (i.e., maximize ER). Therefore, the inhibition tests were repeated with combinations of 0.75 mg/mL BSA, increased amounts of *Taq* (2.5 U per reaction), and using various dilutions of the extracts (in their corresponding elution buffer). Based on the performance of the most inhibited samples, the conditions were considered optimized when the ER was ≥20%. The bone, feces, and hair samples were tested with all combinations of BSA, additional *Taq*, and a range of dilutions. In order to preserve the extracts, sample reactions were performed only once with the exception of four inter-run duplicates used to assess experimental variation. One NTC and two standards reactions were included for every set of PCR conditions in every run.

RESULTS

All of the NTC reactions were negative, while all standard and positive sample reactions had Tm values of 81.8°C (SD_{Tm} = 0.5°C). These results indicate that no unspecific primer binding or contamination with human DNA had occurred. Based on the data from duplicate standard reactions, the mean intra-assay variation in C_q and Hill slope measurements was low, with SD_{Cq} = 0.12 cycles and CV_{HS} = 1.3%; between runs, CV_{HS} increased to 3.8%. Although inter-assay comparisons of C_q values for standard reactions are not possible, four sample reactions duplicated in different runs indicated that $SD_{\Delta Cq}$ was also low (0.08 cycles), with SD_E = 2.1%. No combination of PCR facilitators produced levels of variation significantly different from the mean values presented here. Based on this precision, we presume that the data from single sample reactions is reliable, although there is a risk that aberrant reactions may have gone unnoticed in the absence of technical replicates.

Initial inhibition tests (using undiluted extracts and no PCR facilitators) showed complete inhibition in all soil, feces, and hair extracts, with only bone extracts permitting any amplification of the IPC (Table S2). Both bone extracts (B1, B2) had good amplification efficiency (E = 94%), but with ΔC_q values of 2.4 or 6.1 cycles (Table 2); thus, even the least inhibited extract limited our expected recovery to <20% under these conditions.

As indicated in Table 2, the use of additional *Taq* improved the ΔC_q values (increasing the ER), but led to a decrease in amplification efficiency for the bone extracts. The cause of this decrease is unknown; however, adding BSA improved the ΔC_q values comparably and with optimal efficiency. Incorporating both facilitators brought the ΔC_q values to <1 (62 and 69% ER), such that no dilution would be beneficial in subsequent PCR runs with these extracts.

Table 2. Effects of PCR facilitator combinations at optimal dilution

Extract		Dilution Factor	PCR Facilitators			
			None	BSA	Taq	Both
Bone	B1	1	2.4, 94	1.1, 100	**0.5, 85**	**0.7, 100**
	B2	1	6.1, 94	1.1, 100	1.5, 88	**0.5, 100**
Feces	F1	100	-	1.4, 86	-	**1.0, 94**
	F2	50	-	2.0, 82	-	**1.3, 84**
Hair	H1	50	-	**0.5, 92**	4.6, 78	**0.4, 93**
	H2	50	-	0.7, 90	**0.4, 93**	**0.3, 97**

Inhibition test results [ΔC_q (cycles), E (%)], with results under the optimal PCR facilitator combinations bolded. (-), complete inhibition.

For the feces (F1, F2) and hair (H1, H2) extracts, both facilitators and dilution were required to reduce inhibition sufficiently, although BSA alone was responsible for the majority of the improvement for extracts F1, F2, and H1. Even at very high dilutions, additional *Taq* was unable to facilitate amplification in the feces extracts, but was particularly effective against inhibitors in H2. Unfortunately, the dilutions necessary to overcome inhibition in these extracts drop the maximum ER values below 2%, which may render the optimization process trivial for highly degraded and other low-copy templates.

In terms of reaction success versus complete inhibition among the soil extracts, *Taq* had no significant effect as a facilitator in contrast to BSA ($P = 3 \times 10^{-15}$). Not surprisingly, BSA had a less noticeable effect on reaction success when the extracts were diluted ($P = 3 \times 10^{-3}$ and $P = 5 \times 10^{-1}$ with 1/10 and 1/50 dilutions, respectively). Optimal ER values (from 20 to 77%) were achieved with BSA alone in ~31% of the extracts (Figure 1, solid black triangles), while the other 69% benefited from additional *Taq* as well (Figure 1, solid black circles). A subset of those more inhibited extracts showed optimal ER values using a 1/2 dilution in PCR (Figure 1, solid blue circles). Although the amplification efficiencies were improved by diluting the soil extracts further (Table S2), conditions providing higher ER values were more relevant for subsequent PCR runs.

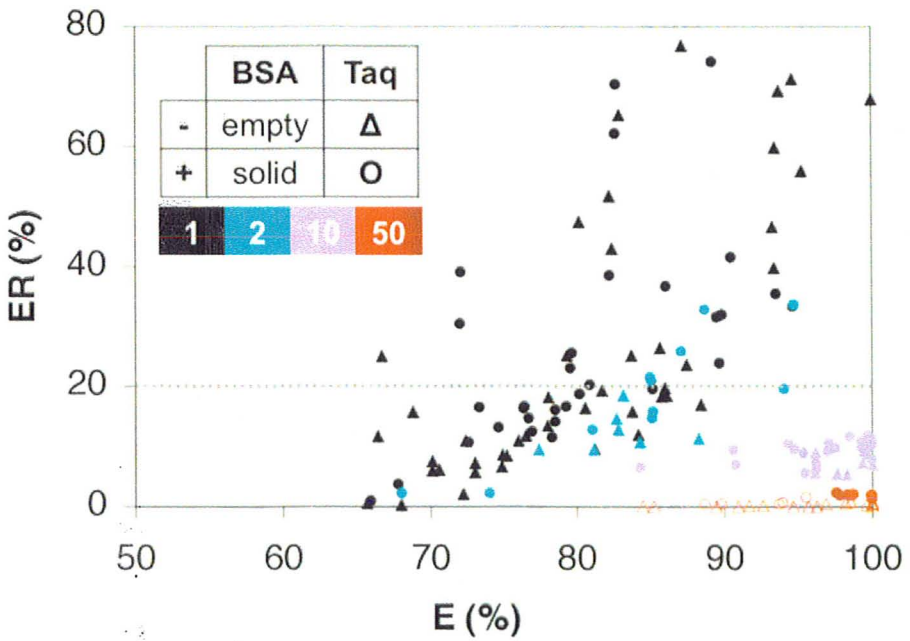


Figure 1. Effects of PCR facilitators and dilution factors on optimization criteria. Results of all successful IPC amplifications from the soil extracts are shown, with the minimum ER level of 20% noted. The use of BSA and additional *Taq* are indicated by the fill and shape (respectively) of each data point, with colors varying by dilution factor. ER, expected recovery; E, relative amplification efficiency.

Table 3. Inhibition measurements by extraction method

Soil Sample	Extraction Method		
	a	b	c
P5	2.6, 88	3.4, 81	7.6, 66
P10	1.1, 93	2.5, 86	5.6, 72
P25	0.6, 100	1.3, 93	NA
P30	0.7, 93	8.5, 68	NA

Inhibition test results [ΔC_q (cycles), E (%)] for undiluted extracts with BSA; (a) Willerslev et al. (27) extraction and purification protocol, (b) Willerslev et al. (27) extraction with alternative purification protocol, (c) UltraClean Soil DNA Isolation Kit (MO BIO Laboratories, Carlsbad, CA, USA). NA, not attempted.

A comparison of inhibition levels across soil DNA extraction methods is shown in Table 3. In all cases, the method of Willerslev et al. (27) demonstrated the lowest inhibition levels. Although we expect some variation in inhibition levels between these extracts due to the heterogeneity of soil, comparable tests of duplicate extracts (i.e., biological replicates) of sample P15 (Table S2) indicate that this variation is minimal, with $SD_{\Delta Cq}$ = 0.3 cycles and SD_E = 1.6%.

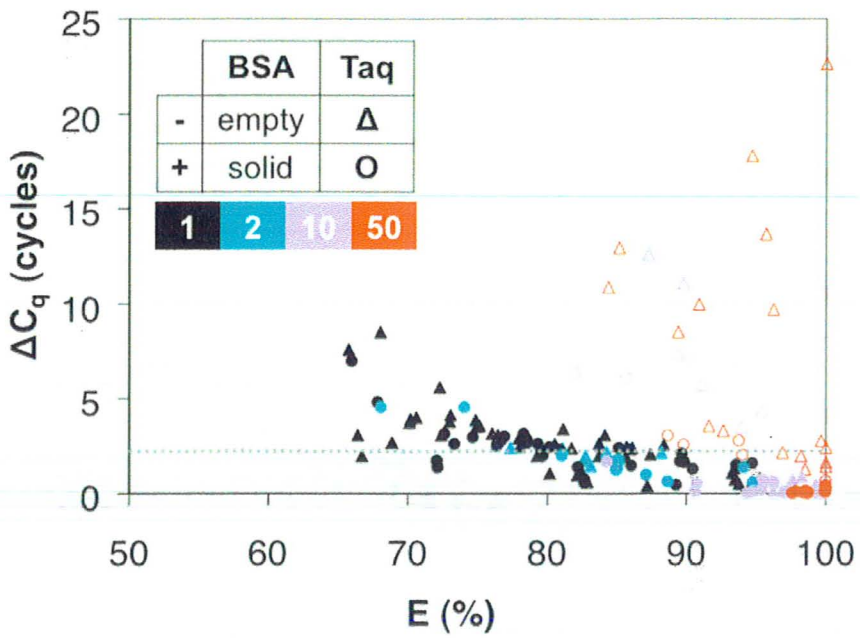


Figure 2. Effects of PCR facilitators and dilution factors on inhibition levels. Results of all successful IPC amplifications from the soil extracts are shown, with the maximum ΔC_q value of 2.3 (i.e. ER = 20% prior to any dilution) noted. The use of BSA and additional *Taq* are indicated by the fill and shape (respectively) of each data point, with colors varying by dilution factor. ΔC_q, cycle shift; E, relative amplification efficiency.

From the data in Figure 2, there is an apparent correlation between E and ΔC_q, which is expected given the threshold-based method of calculating C_q; as the amplification efficiency decreases, the placement of the fluorescence threshold becomes more sensitive, and the resulting ΔC_q values more variable. Thus, many of these poor efficiency reactions have deceptively low ΔC_q values, even though they are highly inhibited.

Interestingly, many reactions display a high ΔC_q value without a corresponding drop in reaction efficiency (Figure 2). This effect of inhibition is apparently overcome by the addition of BSA, as it is only visible in reactions lacking BSA, which required considerable dilution of the extracts to facilitate any amplification. Examples of these atypical efficiency- C_q relationships are shown in Figure 3.

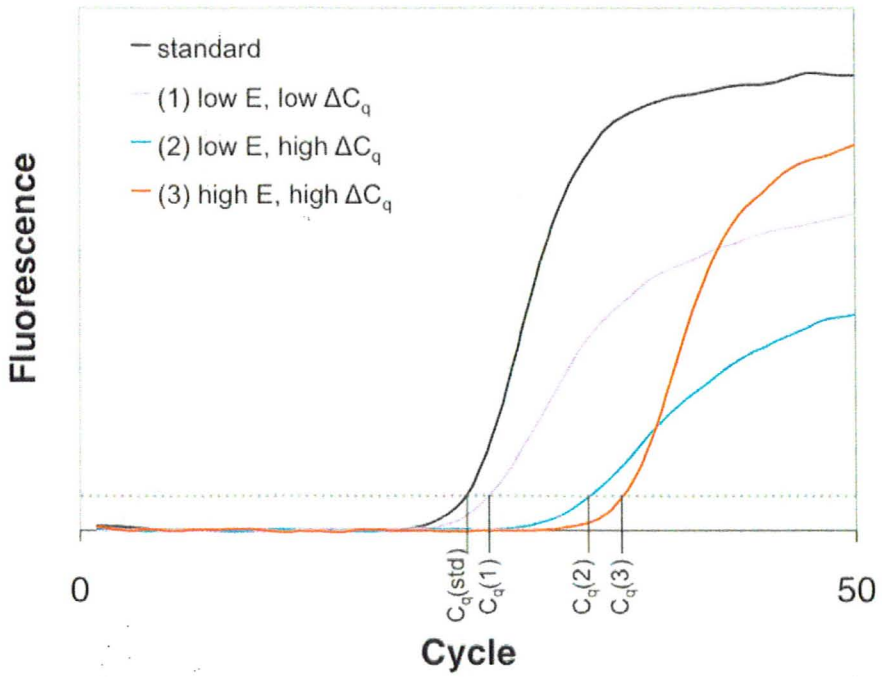


Figure 3. IPC amplification plots illustrating the observed efficiency-C_q shift relationships. Although high ΔC_q values were typically associated with low E values (blue), inhibition may not be apparent from either single measurement in other instances.

DISCUSSION

By incorporating both ΔC_q and efficiency measurements in our experiments, and by analyzing their results through pairwise comparisons, we have detected inhibition with various responses to PCR facilitators. We have typically observed C_q shifts in response to decreased amplification efficiency, as expected using threshold-based quantification. In some instances, however, a low ΔC_q and high ER can be associated with poor efficiency due to placement of the threshold. In these cases, and in any reaction where the efficiency differs from 100% (relative to a standard reaction), the ER value should only be considered an upper limit, but whose accuracy becomes greater as E approaches 100%.

In contrast, we observed many instances in which high ΔC_q values were associated with highly efficient reactions. This can be explained by the presence of thermolabile inhibitors that diminish with cycling; they may inhibit during early cycles of PCR, but the amplification efficiency will eventually recover and produce only a delayed C_q value. Alternatively, inhibitors that degrade or capture DNA (e.g., nucleases and DNA binding substances such as silica) may reduce the number of amplifiable IPC templates in an otherwise efficient PCR. This observation illustrates the importance of using an IPC for inhibition testing; without it (i.e., using efficiency calculations alone) these C_q shifts and large drops in ER would remain undetected.

Amplification of damaged templates may create the same pattern due to a reduction in amplification efficiency in the early cycles of PCR until newly synthesized/undamaged templates predominate (33). This is unlikely to affect our IPC, but may be an unavoidable source of error when quantifying damaged DNA relative to a pristine standard.

With repeat inhibition tests, we were able to optimize the PCR facilitator combinations and dilutions that result in the highest possible ER for each extract. For the bone and most soil samples, dilution could be avoided altogether; in contrast, inhibition in the feces and hair extracts could not be overcome with BSA and additional *Taq*, but instead required that 1/50 or even 1/100 dilutions be used in PCR. Although the resulting ER values are indeed optimal within the parameters of this study, the values are so low for these diluted extracts that additional PCR facilitators should be considered before subsequent PCR runs are attempted. If the values cannot be improved this way, additional purification or even alternative extraction procedures may be beneficial. As we have demonstrated, inhibition testing can also be useful in evaluating these protocols.

Although we recommend testing every extract for inhibition, it may not be necessary to test every combination of PCR facilitators. Some general trends exist within sample types using a particular extraction method: for instance, compared with BSA, additional *Taq* appeared to provide limited improvement for PCR results from our feces and soil extracts. This suggests that their major inhibitors do not act directly on the enzyme, provided that enough *Taq* was used to observe such an effect. In contrast, improvements in solely *Taq*-supplemented hair and bone reactions indicate that direct inhibitors of the enzyme are present. It is not surprising that BSA was able to overcome inhibition in all sample types, based on its combined enzyme-stabilizing and inhibitor-binding (34) abilities.

Recommendations

If the samples are to be used in subsequent quantitative comparisons, several considerations must be made to ensure accuracy. For threshold-based quantification, reaction conditions producing equivalent amplification efficiencies between samples and/or standards should be favored, even if ER is low (e.g., with dilution). If this is not feasible, differences in efficiency can be applied as correction factors (13). Relative measurements of efficiency are sufficient in the former situation, while the latter requires highly accurate, absolute values. Alternatively, the recently published Cy0 method (21) obviates the need to assess amplification efficiency separately from quantification, and may be particularly useful when efficiency is highly variable; furthermore, ΔCy0 values determined from IPC reactions should be analogous to ΔC_q, with their incorporation in ER calculations producing a convenient, single-parameter correction factor.

The accuracy of correction factors derived from IPC data also depends on the original assumption that the IPC and target assays are equally affected by inhibition. Intuitively, large differences in the amplicon sizes should be avoided; however, no particular sequence characteristics are clear predictors of inhibition susceptibility (11). Huggett et al. (11) suggest that assay performance be compared in the presence of a suspected inhibitor; however, this approach may be ineffective when the inhibitors are of unknown types and amounts, and additionally impractical when inhibitors vary between samples. If inhibition susceptibility is instead related to competition for reaction components (10), it should be sufficient to demonstrate that the assays behave similarly (i.e., within the accuracy requirements of the experiment) under a variety of conditions (e.g., concentration gradients of Mg^{2+}, primers, and facilitators).

Based on the data presented here, we suggest that inhibition testing be regularly incorporated in procedures involving low amounts of DNA, such as in the forensic and ancient DNA fields. While the results are generally helpful for improving PCR success, they should be considered essential in experiments involving qPCR comparisons and when interpreting negative results. These tests should incorporate an IPC, which allows for simultaneous measurement of amplification efficiency and ΔC_q. When employed routinely, this form of pre-PCR processing ensures both maximum access to template DNA and more accurate quantification of template amounts.

ACKNOWLEDGMENTS

This work was supported by funding from the Natural Sciences and Engineering Research Council of Canada (grant no. 299103-2004), the Canadian Research Chairs program (to H.P.), the Social Sciences and Humanities Research Council of Canada (grant no. 410-2004-0579), the Government of Ontario Early Researcher Award program (to H.P.), the Ontario Graduate Scholarship program (to C.E.K.), and McMaster University.

The authors thank Kirsti Bos, Alison Devault, Debi Poinar, John Okello, and Jaymi Zurek for helpful discussion; and Duane Froese, Grant Zazula, Bernard Buigues, and Felisa Smith for providing the samples.

The authors declare no competing interests.

REFERENCES

1. **Wilson, I.G.** 1997. Inhibition and facilitation of nucleic acid amplification. Appl. Environ. Microbiol. *63*:3741-3751.
2. **Rossen, L., P. Nørskov, K. Holmstrøm, and O.F. Rasmussen.** 1992. Inhibition of PCR by components of food samples, microbial diagnostic assays and DNA-extraction solutions. Int. J. Food Microbiol. *17*:37-45.
3. **Ramakers, C., J.M. Ruijter, R.H. Deprez, and A.F. Moorman.** 2003. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci. Lett. *339*:62-66.
4. **Cloud, J.L., W.C. Hymas, A. Turlak, A. Croft, U. Reischl, J.A. Daly, and K.C. Carroll.** 2003. Description of a multiplex *Bordetella pertussis* and *Bordetella parapertussis* LightCycler PCR assay with inhibition control. Diagn. Microbiol. Infect. Dis. *46*:189-195.
5. **Sutlović, D., M. Definis Gojanović, Š. Andelinović, D. Gugić, and D. Primorac.** 2005. *Taq* polymerase reverses inhibition of quantitative real time polymerase chain reaction by humic acid. Croat. Med. J. *46*:556-562.

6. **Deagle, B.E., J.P. Eveson, and S.N. Jarman.** 2006. Quantification of damage in DNA recovered from highly degraded samples – a case study on DNA in faeces. Front. Zool. *3*:11.
7. **Nolan, T., R.E. Hands, W. Ogunkolade, and S.A. Bustin.** 2006. SPUD: a quantitative PCR assay for the detection of inhibitors in nucleic acid preparations. Anal. Biochem. *351*:308-310.
8. **Swango, K.L., M.D. Timken, M.D. Chong, and M.R. Buoncristiani.** 2006. A quantitative PCR assay for the assessment of DNA degradation in forensic samples. Forensic Sci. Int. *158*:14-26.
9. **Niederstätter, H., S. Köchl, P. Grubwieser, M. Pavlic, M. Steinlechner, and W. Parson.** 2007. A modular real-time PCR concept for determining the quantity and quality of human nuclear and mitochondrial DNA. Forensic Sci. Int. Genet. *1*:29-34.
10. **Ståhlberg, A., P. Åman, B. Ridell, P. Mostad, and M. Kubista.** 2003. Quantitative real-time PCR method for detection of B-lymphocyte monoclonality by comparison of κ and λ immunoglobulin light chain expression. Clin. Chem. *49*:51-59.
11. **Huggett, J.F., T. Novak, J.A. Garson, C. Green, S.D. Morris-Jones, R.F. Miller, and A. Zumla.** 2008. Differential susceptibility of PCR reactions to inhibitors: an important and unrecognised phenomenon. BMC Res Notes *1*:70.
12. **Bustin, S.A., V. Benes, J.A. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, et al.** 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin. Chem. *55*:611-622.
13. **Liu, W. and D.A. Saint.** 2002. A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics. Anal. Biochem. *302*:52-59.
14. **Liu, W. and D.A. Saint.** 2002. Validation of a quantitative method for real time PCR kinetics. Biochem. Biophys. Res. Commun. *294*:347-353.
15. **Tichopad, A., A. Dzidic, and M.W. Pfaffl.** 2002. Improving quantitative real-time RT-PCR reproducibility by boosting primer-linked amplification efficiency. Biotechnol. Lett. *24*:2053-2056.
16. **Bar, T., A. Ståhlberg, A. Muszta, and M. Kubista.** 2003. Kinetic Outlier Detection (KOD) in real-time PCR. Nucleic Acids Res. *31*:e105.
17. **Peirson, S.N., J.N. Butler, and R.G. Foster.** 2003. Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis. Nucleic Acids Res. *31*:e73.
18. **Tichopad, A., M. Dilger, G. Schwarz, and M.W. Pfaffl.** 2003. Standardized determination of real-time PCR efficiency from a single reaction set-up. Nucleic Acids Res. *31*:e122.
19. **Rutledge, R.G.** 2004. Sigmoidal curve-fitting redefines quantitative real-time PCR with the prospective of developing automated high-throughput applications. Nucleic Acids Res. *32*:e178.

20. **Kontanis, E.J. and F.A. Reed.** 2006. Evaluation of real-time PCR amplification efficiencies to detect PCR inhibitors. J. Forensic Sci. *51*:795-804.
21. **Guescini, M., D. Sisti, M.B. Rocchi, L. Stocchi, and V. Stocchi.** 2008. A new real-time PCR method to overcome significant quantitative inaccuracy due to slight amplification inhibition. BMC Bioinformatics *9*:326.
22. **Rutledge, R.G. and D. Stewart.** 2008. Critical evaluation of methods used to determine amplification efficiency refutes the exponential character of real-time PCR. BMC Mol. Biol. *9*:96.
23. **Shain, E.B. and J.M. Clemens.** 2008. A new method for robust quantitative and qualitative analysis of real-time PCR. Nucleic Acids Res. *36*:e91.
24. **Swillens, S., B. Dessars, and H.E. Housni.** 2008. Revisiting the sigmoidal curve fitting applied to quantitative real-time PCR data. Anal. Biochem. *373*:370-376.
25. **Ruijter, J.M., C. Ramakers, W.M. Hoogaars, Y. Karlen, O. Bakker, M.J. van den Hoff, and A.F. Moorman.** 2009. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Res. *37*:e45.
26. **Rådström, P., R. Knutsson, P. Wolffs, M. Lövenklev, and C. Löfström.** 2004. Pre-PCR processing: strategies to generate PCR-compatible samples. Mol. Biotechnol. *26*:133-146.
27. **Willerslev, E., A.J. Hansen, J. Binladen, T.B. Brand, M.T. Gilbert, B. Shapiro, M. Bunce, C. Wiuf, et al.** 2003. Diverse plant and animal genetic records from Holocene and Pleistocene sediments. Science *300*:791-795.
28. **Boom, R., C.J. Sol, M.M. Salimans, C.L. Jansen, P.M. Wertheim-van Dillen, and J. van der Noordaa.** 1990. Rapid and simple method for purification of nucleic acids. J. Clin. Microbiol. *28*:495-503.
29. **Rohland, N. and M. Hofreiter.** 2007. Comparison and optimization of ancient DNA extraction. BioTechniques *42*:343-352.
30. **Poinar, H.N., C. Schwarz, J. Qi, B. Shapiro, R.D.E. MacPhee, B. Buigues, A. Tikhonov, D.H. Huson, et al.** 2006. Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. Science *311*:392-394.
31. **Gilbert, M.T., L.P. Tomsho, S. Rendulic, M. Packard, D.I. Drautz, A. Sher, A. Tikhonov, L. Dalen, et al.** 2007. Whole-genome shotgun sequencing of mitochondria from ancient hair shafts. Science *317*:1927-1930.
32. **Gilbert, M.T., T. Haselkorn, M. Bunce, J.J. Sanchez, S.B. Lucas, L.D. Jewell, E. Van Marck, and M. Worobey.** 2007. The isolation of nucleic acids from fixed, paraffin-embedded tissues-which methods are useful when? PLoS One *2*:e537.
33. **Sikorsky, J.A., D.A. Primerano, T.W. Fenger, and J. Denvir.** 2004. Effect of DNA damage on PCR amplification efficiency with the relative threshold cycle method. Biochem. Biophys. Res. Commun. *323*:823-830.
34. **Kreader, C.A.** 1996. Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. Appl. Environ. Microbiol. *62*:1102-1106.

Table S1. MIQE Checklist

Item	D/E	y/n	Comments
Experimental Design			
Definition of experimental and control groups	E	y	...
Number within each group	E	y	...
Assay carried out by core lab or investigator's lab?	D	y	Investigator's lab
Acknowledgement of authors' contributions	D	y	...
Sample			
Description	E	y	...
Volume/mass of sample processed	D	y	...
Microdissection or macrodissection	E	n	n/a
Processing procedure	E	n	n/a
If frozen - how and how quickly?	E	n	n/a
If fixed - with what, how quickly?	E	n	n/a
Sample storage conditions and duration (especially for FFPE samples)	E	y	All samples stored at room temperature since collection except permafrost soil cores (-40°C)
Nucleic Acid Extraction			
Procedure and/or instrumentation	E	y	...
Name of kit and details of any modifications	E	y	...
Source of additional reagents used	D	y	...
Details of DNase or RNase treatment	E	n	n/a
Contamination assessment (DNA or RNA)	E	n	n/a
Nucleic acid quantification	E	n	n/a
Instrument and method	E	n	n/a
Purity (A260/A280)	D	n	n/a
Yield	D	n	n/a
RNA integrity method/instrument	E	n	n/a
RIN/RQI or C_q of 3' and 5' transcripts	E	n	n/a
Electrophoresis traces	D	n	n/a
Inhibition testing (C_q dilutions, spike or other)	E	y	...
Reverse Transcription			
Complete reaction conditions	E	n	n/a
Amount of RNA and reaction volume	E	n	n/a
Priming oligonucleotide (if using GSP) and concentration	E	n	n/a
Reverse transcriptase and concentration	E	n	n/a
Temperature and time	E	n	n/a
Manufacturer of reagents and catalogue numbers	D	n	n/a
C_qs with and without RT	D	n	n/a
Storage conditions of cDNA	D	n	n/a
qPCR Target Information			
If multiplex, efficiency and LOD of each assay.	E	n	n/a
Sequence accession number	E	y	See RTPrimerDB (ID:1529)
Location of amplicon	D	y	See RTPrimerDB (ID:1529)
Amplicon length	E	y	...
In silico specificity screen (BLAST, etc)	E	y	Primers align with human and other primate B2M sequences
Pseudogenes, retropseudogenes or other homologs?	D	y	None found
Sequence alignment	D	n	...
Secondary structure analysis of amplicon	D	y	mfold analysis of IPC, 25°C, 2.5 mM Mg^{2+}, 50 mM Na^{+}: ΔG = -14.55 kcal/mol
Location of each primer by exon or intron (if applicable)	E	y	see RTPrimerDB (ID:1529)
What splice variants are targeted?	E	n	n/a
qPCR Oligonucleotides			
Primer sequences	E	y	...
RTPrimerDB Identification Number	D	y	ID:1529; no probe, primers used in SYBR Green assay
Probe sequences	D	n	n/a
Location and identity of any modifications	E	y	...
Manufacturer of oligonucleotides	D	y	Integrated DNA Technologies (IDT), Coralville, IA, USA
Purification method	D	y	Desalting

Table S1 (continued). MIQE Checklist

Item	D/E	y/n	Comments
	qPCR Protocol		
Complete reaction conditions	E	y	...
Reaction volume and amount of cDNA/DNA	E	y	...
Primer, (probe), Mg^{2+} and dNTP concentrations	E	y	...
Polymerase identity and concentration	E	y	...
Buffer/kit identity and manufacturer	E	y	...
Exact chemical constitution of the buffer	D	y	...
Additives (SYBR Green I, DMSO, etc.)	E	y	...
Manufacturer of plates/tubes and catalog number	D	y	VWR (West Chester, PA, USA), catalog #93001-118
Complete thermocycling parameters	E	y	...
Reaction setup (manual/robotic)	D	y	Manual
Manufacturer of qPCR instrument	E	y	...
	qPCR Validation		
Evidence of optimisation (from gradients)	D	y	Optimized for specificity based on [Mg^{2+}] and annealing temperature gradients
Specificity (gel, sequence, melt, or digest)	E	y	Melting/dissociation curves
For SYBR Green I, C_q of the NTC	E	y	...
Standard curves with slope and y-intercept	E	y	Experiments using a 6-point, 5-fold dilution series of the IPC (12500 to 4 copies/reaction) in duplicate, over 4 runs: mean slope = -3.229, mean Y-intercept = 40.50
PCR efficiency calculated from slope	E	y	Mean efficiency = 104%
Confidence interval for PCR efficiency or standard error	D	y	(Interassay) CV of the mean efficiency = 2.2%
r^2 of standard curve	E	y	Mean r^2 = 0.995
Linear dynamic range	E	y	...
C_q variation at lower limit	E	y	Average SD of C_q for duplicate reactions with 4 copies = 0.33 cycles
Confidence intervals throughout range	D	n	...
Evidence for limit of detection	E	y	4 copies/reaction detected in 100% of attempts (n=8)
If multiplex, efficiency and LOD of each assay.	E	n	n/a
	Data Analysis		
qPCR analysis program (source, version)	E	y	...
C_q method determination	E	y	...
Outlier identification and disposition	E	n	n/a
Results of NTCs	E	y	...
Justification of number and choice of reference genes	E	n	n/a
Description of normalisation method	E	n	n/a
Number and concordance of biological replicates	D	y	...
Number and stage (RT or qPCR) of technical replicates	E	y	...
Repeatability (intra-assay variation)	E	y	...
Reproducibility (inter-assay variation, %CV)	D	y	...
Power analysis	D	n	...
Statistical methods for result significance	E	y	Chi-squared test
Software (source, version)	E	y	...
C_q or raw data submission using RDML	D	n	...

Checklist from: http://www.rdml.org/guidelines.php showing (E) essential and (D) desirable information to be included in research reports using qPCR. Any essential items not included (n) were not applicable (n/a) to the study. The real-time PCR primer and probe database (RTPrimerDB): http://medgen.ugent.be/ rtprimerdb/index.php; mfold: http://mfold.bioinfo.rpi.edu/cgi-bin/dna-form1.cgi.

Table S2. Inhibition test results for all extracts and PCR facilitator/dilution factor combination attempted

Extract	BSA (mg/ml)	Taq (units)	Dilution Factor	ΔC_q (cycles)	ER (%)	E (%)
B1	0	1	1	2.4	18.6	94
			2	0.7	31.9	98
			10	0	10.8	99
		2.5	1	0.5	**71**	85
			2	0.1	47.1	96
			10	0	10.6	95
	0.75	1	1	1.1	46.7	100
			2	0.4	36.9	100
			10	0	10.9	100
		2.5	1	0.7	**62.2**	100
			2	0.2	42.5	100
			10	0.1	9.6	100
B2	0	1	1	6.1	1.5	94
			2	2	12.6	92
			10	0.3	8	97
		2.5	1	1.5	35.5	88
			2	0.2	44.6	92
			10	0	11.9	100
	0.75	1	1	1.1	45.4	100
			2	0.1	45.7	100
			10	0	12.5	100
		2.5	1	0.5	**69**	100
			2	0	53.8	97
			10	0	11.1	99
F1	0	1	1	-	-	-
			2	-	-	-
			10	-	-	-
			50	-	-	-
			100	-	-	-
		2.5	1	-	-	-
			2	-	-	-
			10	-	-	-
			50	-	-	-
			100	-	-	-

Extract	BSA (mg/ml)	Taq (units)	Dilution Factor	ΔC_q (cycles)	ER (%)	E (%)
F1 (continued)	0.75	1	1	-	-	-
			2	-	-	-
			10	-	-	-
			50	3.2	0.2	77
			100	1.4	**0.4**	86
		2.5	1	-	-	-
			2	-	-	-
			10	-	-	-
			50	2	**0.5**	83
			100	1	**0.5**	94
F2	0	1	1	-	-	-
			2	-	-	-
			10	-	-	-
			50	-	-	-
		2.5	1	-	-	-
			2	-	-	-
			10	-	-	-
			50	-	-	-
	0.75	1	1	-	-	-
			2	-	-	-
			10	11.4	0	61
			50	2	0.5	82
		2.5	1	-	-	-
			2	-	-	-
			10	8.6	0	55
			50	1.3	**0.8**	84
H1	0	1	1	-	-	-
			10	-	-	-
			50	-	-	-
			100	13.3	0	78
		2.5	1	-	-	-
			10	-	-	-
			50	4.6	0.1	78
			100	1.5	0.4	89

Table S2 (continued). Inhibition test results for all extracts and PCR facilitator/dilution factor combination attempted

Extract	BSA (mg/ml)	Taq (units)	Dilution Factor	ΔC_q (cycles)	ER (%)	E (%)
H1 (continued)	0.75	1	1	-	-	-
			10	4.4	0.5	67
			50	0.5	**1.4**	92
			100	0.2	0.9	97
		2.5	1	-	-	-
			10	4.3	0.5	66
			50	0.4	**1.6**	93
			100	0.3	0.8	98
H2	0	1	1	-	-	-
			10	20.7	0	64
			50	-	-	-
			100	3.1	0.1	83
		2.5	1	-	-	-
			10	5.6	0.2	61
			50	0.4	**1.5**	93
			100	0.3	0.8	94
	0.75	1	1	-	-	-
			10	8.9	0	55
			50	0.7	1.2	90
			100	0.5	0.7	94
		2.5	1	-	-	-
			10	2.1	2.3	78
			50	0.3	**1.6**	
			100	0.8	0.6	91
P1-b	0	1	1	-	-	-
			10	5.7	0.2	91
			50	1.6	0.7	100
	0.75	1	1	1.9	**26.4**	86
		2.5	10	0	10.4	100
P2-b	0.75	1	1	2.6	16.4	81
		2.5	1	1.7	**31.4**	90
P3-b	0.75	1	1	2.3	19.8	86
		2.5	1	1.7	**31.9**	90
P4-b	0.75	1	1	3.1	11.9	84
		2.5	1	2.1	**23.8**	90
P5-a	0.75	1	1	2.6	17	88
		2.5	1	1.6	**33.2**	95

Extract	BSA (mg/ml)	Taq (units)	Dilution Factor	ΔC_q (cycles)	ER (%)	E (%)
P5-b	0	1	1	-	-	-
			10	7.3	0.1	89
	0.75	1	1	3.4	9.6	81
			2	2.2	11.2	88
			1	2.4	**19.5**	85
		2.5	2	1.4	**19.5**	94
			10	0	9.7	100
P5-c	0.75	1	1	7.6	0.5	66
		2.5	1	4.8	**3.6**	68
P6-b	0	1	1	-	-	-
			10	12.6	0	87
			50	2.4	0.4	100
	0.75	1	1	1.2	**42.9**	82
		2.5	10	0	11.6	99
P7-b	0.75	1	1	2	**25.2**	84
P8-b	0.75	1	1	2.1	**23.7**	87
P9-b	0.75	1	1	2.4	18.6	86
		2.5	1	1.5	**35.4**	94
P10-a	0.75	1	1	1.1	**46.7**	93
P10-b	0	1	1	-	-	-
			10	11.1	0	90
			50	2.8	0.3	100
	0.75	1	1	2.5	18.3	86
		2.5	1	1.3	**41.5**	90
			10	0.1	9.1	97
P10-c	0.75	1	1	5.6	2	72
		2.5	1	3.1	**11.4**	78
P11-b	0	1	1	-	-	-
			1.1	-	-	-
			10	-	-	-
			10.9	-	-	-
			50	17.8	0	95
			54.5	10.9	0	84
		2.5	1.1	-	-	-
			10.9	-	-	-
			54.5	2.6	0.3	90

Table S2 (continued). Inhibition test results for all extracts and PCR facilitator/dilution factor combination attempted

Extract	BSA (mg/ml)	Taq (units)	Dilution Factor	ΔC_q (cycles)	ER (%)	E (%)
P11-b (continued)	0.75	1	1	4	6.2	71
			1.1	3.1	11	72
			2	2	12.8	83
			10.9	0.8	5.4	98
			54.5	0.5	1.3	100
		2.5	1	2.9	13.1	75
			1.1	2.5	16.3	76
			2	1.2	**21.4**	85
			10	0	10.3	94
			10.9	0.6	6	96
			50	0	2	98
			54.5	0.4	1.4	100
P12-b	0.75	1	1	2.5	18.2	78
		2.5	1	2	**25.5**	80
P13-b	0.75	1	1	3.2	11	76
			2	1.4	18.5	83
		2.5	1	2.4	18.7	80
			2	1	**25.8**	87
P14-b	0.75	1	1	4.1	5.7	73
			2	2.4	9.5	81
		2.5	1	3	**12.4**	77
			2	1.8	**14.6**	85
			50	2	0.5	98
P15-b-1	0	1	1	-	-	-
			10	-	-	-
			50	13.7	0	96
	0.75	1	1	3.6	8.4	75
			2	1.8	14.6	83
		2.5	1	2.6	16	79
			2	1.3	**20.8**	85
			10	0.1	9.6	97
			50	0.1	1.8	99
P15-b-2	0.75	1	1	2.9	13.5	78
		2.5	1	2.6	16.6	79
			2	0.6	**32.6**	89

Extract	BSA (mg/ml)	Taq (units)	Dilution Factor	ΔC_q (cycles)	ER (%)	E (%)
P16-b	0	1	1	-	-	-
			10	-	-	-
			50	9.7	0	96
	0.75	1	1	1	**51.8**	82
		2.5	10	0	11.1	100
			50	0.1	1.9	98
P17-b	0.75	1	1	3.8	7.2	73
		2.5	1	2.3	**20.2**	81
P18-b	0.75	1	1	3.1	11.8	76
		2.5	1	2.1	**23**	80
P19-b	0.75	1	1	2.4	19.3	82
		2.5	1	1.5	**36.6**	86
P20-b	0	1	1	-	-	-
			10	-	-	-
			50	-	-	-
		2.5	1	-	-	-
			10	-	-	-
		5	1	-	-	-
	0.75	1	1	3.9	6.7	75
			1	3.5	8.7	75
			2	2.2	10.8	84
			10	0.9	5.2	98
			50	0.6	1.3	100
		2.5	1	2.8	**14.1**	79
			1	2.8	**14.7**	77
			2	1.7	**15.7**	85
			10	0.1	9.3	91
			10	0.1	9.5	95
			50	0	1.9	99
		5	1	2.9	13.1	80
P21-a	0	1	1	-	-	-
			10	-	-	-
			50	3.6	0.2	92
	0.75	1	1	0.4	**76.8**	87
		2.5	10	0	11.7	100
			50	0	2.2	98

Table S2 (continued). Inhibition test results for all extracts and PCR facilitator/dilution factor combination attempted

Extract	BSA (mg/ml)	Taq (units)	Dilution Factor	ΔC_q (cycles)	ER (%)	E (%)
P22-a	0.75	1	1	2	**25.2**	79
P23-a	0	1	1.1	-	-	-
			10.9	-	-	-
			54.5	22.6	0	100
		2.5	1.1	-	-	-
			10.9	-	-	-
			54.5	3	0.2	89
	0.75	1	1	3.7	7.5	70
			1.1	3.9	6	70
			2	2.4	9.4	77
			10.9	0.7	5.6	96
			54.5	0.2	1.6	100
		2.5	1	2.6	**16.6**	76
			1.1	3.1	10.5	73
			2	2	12.7	81
			10.9	0.8	5.5	96
			54.5	0.1	1.7	100
P24-a	0.75	1	1	2.7	15.8	84
		2.5	1	1.4	**38.4**	82
P25-a	0	1	1	-	-	-
			10	4.4	0.5	95
			50	1.4	0.8	100
	0.75	1	1	0.6	**67.8**	100
		2.5	10	0	10.7	99
P25-b	0.75	1	1	1.3	39.8	93
		2.5	1	0.4	**74**	89
			2	0.6	33.6	95
P26-a	0	1	1	-	-	-
			10	-	-	-
			50	8.5	0	89
	0.75	1	1	3.1	11.7	66
			1	1.4	**39**	72
		2.5	10	0	10.5	97
			50	0.1	1.9	100
P27-a	0.75	1	1	0.5	**71.2**	95
P28-a	0.75	1	1	0.5	**69.3**	94
P29-a	0.75	1	1	0.8	**55.9**	95

Extract	BSA (mg/ml)	Taq (units)	Dilution Factor	ΔC_q (cycles)	ER (%)	E (%)
P30-a	0	1	1	-	-	-
			10	3.5	0.9	94
			50	1.3	0.8	99
	0.75	1	1	0.7	**59.9**	93
		2.5	10	0	9.7	99
P30-b	0.75	1	1	8.5	0.3	68
			1	7	0.8	66
		2.5	2	4.5	2.1	74
			2	4.5	2.1	68
			5	1.7	**6.4**	84
			10	0.5	**6.9**	91
	1	2.5	2	4.3	2.5	73
			5	1.5	**6.9**	87
			10	0.5	**7.1**	91
P31-a	0	1	1	-	-	-
			10	-	-	-
			50	9.9	0	91
		2.5	1	-	-	-
			10	-	-	-
			50	2	0.5	94
	0.75	1	1	2	25.1	67
			10	0.1	9.1	96
			50	0.1	1.9	100
		2.5	1	1.7	**30.4**	72
			10	0.2	8.8	95
			50	0.2	1.7	100
P32-a	0	1	1	-	-	-
			10	-	-	-
			50	12.9	0	85
		2.5	1	-	-	-
			10	-	-	-
			50	2.8	0.3	94
	0.75	1	1	2.7	**15.7**	69
			10	0.5	7.2	99
			50	0.2	1.8	100

Table S2 (continued). Inhibition test results for all extracts and PCR facilitator/dilution factor combination attempted

Extract	BSA (mg/ml)	Taq (units)	Dilution Factor	ΔC_q (cycles)	ER (%)	E (%)
P32-a (continued)	0.75 (continued)	2.5	1	2.6	**16.4**	73
			10	0.5	7.1	96
			50	0.3	1.6	100
P33-b	0	1	1	-	-	-
			50	2.1	0.5	97
		2.5	1	-	-	-
			10	6.1	0.1	86
			50	0.7	1.3	100
	0.75	1	1	0.6	**65.3**	83
			10	0.1	9.3	99
			50	0.2	1.8	100
		2.5	1	0.5	**70.2**	83
			10	0.3	8.3	100
			50	0	2	99
P34-b	0	1	1	-	-	-
			10	-	-	-
			50	3.3	0.2	93
		2.5	1	-	-	-
			10	6.3	0.1	82
			50	0.5	1.4	96
	0.75	1	1	1.1	47.5	80
			10	0.5	6.9	100
			50	0.2	1.8	100
		2.5	1	0.7	**62**	83
			10	0.4	7.7	100
			50	0.1	1.8	100

Optimal ER values (within 15% of the maximum ER for each extract) are bolded. Extracts are labeled according to sample type [(B) bone, (F) feces, (H) hair, (P) permafrost soil] and extraction method (a, b, or c - see main text). BSA, bovine serum albumin; ΔC_q, cycle shift; ER, expected recovery; E, relative amplification efficiency; (-), complete inhibition of IPC amplification.

PART II:

ANTIBIOTIC RESISTANCE

Throughout the history of clinical antibiotic use, resistant bacteria have consistently emerged and confounded the utility of these drugs [25, 26]. The proliferation of resistance among pathogenic microbes by horizontal gene transfer (HGT) and natural selection has been the focus of much research, highlighting the effects of antibiotic misuse in the clinic and in agriculture. In contrast, investigation of the origins of resistance mechanisms and their entry into this cycle has been relatively minimal, being otherwise perceived as a modern phenomenon that coincides with the antibiotic era.

Recently, the existence of a natural reservoir of resistance determinants has been explored, revealing environmental bacteria resistant to a variety of antibiotics [27-29]. In some cases these genes provide resistance mechanisms for the antibiotic producers themselves, while others confer resistance inadvertently, having unrelated functions in the environment [29]. This may indicate that pathogens acquire resistance from this environmental reservoir, but may also be interpreted in the other direction, or even as the result of antibiotic contamination at the study site and recent evolution therein.

With these possibilities in mind, others have attempted to study "pristine" environments, free of anthropogenic influences [30]; however, given the potential mobility (e.g. by water, wind, animals) of antibiotics and/or resistant bacteria, even remote locales may be far from pristine [29]. We addressed this concern in the following article, entitled *Antibiotic resistance is ancient*, using the Beringian permafrost samples.

Both the cryostratigraphy of the site and the results of our metagenomic analysis suggest that the permafrost has not thawed since its deposition approximately 30,000 years ago; the latter including flora (e.g. *Poa*, *Festuca*, *Artemisia*, *Salix*) and fauna (e.g. *Bison*, *Equus*, *Ovis*, *Microtus*, *Ellobius*) consistent with the Late Pleistocene macrofossil record in the region [31-35] and excluding common Holocene species (e.g. elk, moose, spruce). In the absence of liquid water and subsequent leaching of modern contaminants, these samples offer a truly pristine environment for the study of natural resistance reservoirs.

Several reports claim to have cultured resistant strains of ancient bacteria from Siberian permafrost [36-40], but the viability of ancient cells remains controversial; furthermore, these particular results are debatable due to the absence of adequate experimental controls to rule out contamination [41]. In contrast, we employed a direct sequencing approach (i.e. culture-independent PCR) with rigorous procedures to minimize and detect contamination. The necessity of these steps cannot be overstated, as there are no simple means to distinguish contaminating material from endogenous material in an environmental sample.

With these procedures in place, we successfully amplified fragments of *tetM* (a ribosome-protecting protein), *vanX* (a D-Ala-D-Ala dipeptide hydrolase), and *bla* (a β-lactamase). The diversity of the ancient sequences was unexpectedly high, even at the protein level. In most cases, these sequences cluster with those found in characterized tetracycline-, vancomycin- and penicillin-resistant strains, respectively, and are distinct from homologues with other known functions.

In addition to authenticating our presence/absence data, a second objective was to determine if these ancient genes were also functional as resistant determinants and were not simply homologous sequence fragments. This is essential to our conclusion that resistance is ancient, as one could argue that this functionality was the result of further evolutionary changes during the antibiotic era. This was done both indirectly and directly. First, we confirmed the contiguity of the genes in the *vanHAX* operon (only together conferring vancomycin resistance, and having no other known function in this form) by amplifying across the gene boundaries; second, we sequenced the complete *vanA* gene, and tested its activity and specificity as a D-Ala-D-Lac ligase.

Taken together, these results offer the first conclusive evidence that functional resistance determinants were present in the environment prior to modern antibiotic use. The dangers of our antibiotic use remain and should not be disregarded; rather, these results have serious implications for the development of new antibiotics. Confirming the effectiveness of a drug against pathogens and even against symbiotic microbes [42] is insufficient; every effort must be made to screen environmental bacteria in order to contend with the extraordinarily diverse antibiotic resistome.

Antibiotic resistance is ancient

Thirty thousand year-old permafrost contains functional antibiotic resistance genes.

Vanessa M. D'Costa[1,2]*, Christine E. King[3,4]*, Lindsay Kalan[1,2], Maryia Morar[1,2], Wilson Sung[4], Carsten Schwarz[3], Duane Froese[5], Grant Zazula[6], Fabrice Calmels[5], Regis Debruyne[7], G. Brian Golding[4], Hendrik N. Poinar[1,3,4†], Gerard D. Wright[1,2†]

[1]Michael G. DeGroote Institute for Infectious Disease Research, McMaster University, Hamilton, Ontario, Canada, L8N 3Z5, [2]Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario, Canada, L8N 3Z5, [3]McMaster Ancient DNA Centre, Department of Anthropology, McMaster University, Hamilton, Ontario, Canada, L8S 4L9, [4]Department of Biology, McMaster University, Hamilton, Ontario, Canada, L8S 4K1, [5]Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, Alberta, Canada, T6G 2E3, [6]Yukon Palaeontology Program, Department of Tourism and Culture, Yukon Government, P.O. Box 2703, Whitehorse, Yukon, Canada, Y1A 6V6, and [7]Muséum National d'Histoire Naturelle, UMR 7206 Eco-anthropologie, 57 rue Cuvier, CP139, 75231 Paris Cedex 05

* These authors contributed equally to this work

This version of the manuscript submitted for publication, March 2011

†To whom correspondence should be addressed: Email wrightge@mcmaster.ca (G.D.W.); poinarh@mcmaster.ca (H.N.P.)

Antibiotic resistance has often been perceived as a modern phenomenon, primarily resulting from selective pressures imposed by human implementation of antibiotics. However, one of the paradoxes of antibiotic use is the rapid emergence of resistance following clinical deployment. We performed targeted metagenomic analyses of rigorously authenticated ancient DNA from 30,000 year-old Beringian permafrost sediments and identified a highly diverse collection of genes encoding resistance to β-lactam, tetracycline, and glycopeptide antibiotics. Structure and function studies on the complete vancomycin resistance element VanA confirmed its similarity to modern variants. These results conclusively show that antibiotic resistance is a natural phenomenon that predates the modern selective pressure of clinical antibiotic use.

The discovery of the first antibiotics over 70 years ago initiated a period of innovation and implementation in human and animal health and agriculture. These remarkable discoveries, however, were tempered in all cases by the emergence of resistant microbes (*1, 2*). This history is often interpreted to mean that antibiotic resistance is a modern phenomenon; giving strength to this view is the fact that some collections of microbes that predate the antibiotic era are highly susceptible to antibiotics (*3*).

Recent studies of modern environmental and human commensal microbial genomes demonstrate a much larger concentration of antibiotic resistance genes than previously recognized (*4-6*). Additionally, metagenomic studies have revealed diverse homologs of known resistance genes broadly distributed across environmental locales (*7*). This widespread dissemination of antibiotic resistance elements is inconsistent with a hypothesis of contemporary emergence and instead suggests a richer natural history of resistance (*2*). Indeed, estimates of the origin of natural product antibiotics range from 2B to 40M years ago (*8, 9*), suggesting that resistance should be similarly old. Previous publications claim to have cultured resistant bacteria from Siberian permafrost (e.g. ref (*10*)), but these results remain contentious (see supporting online material (SOM), section 9).

In order to determine whether contemporary resistance elements are modern or have originated prior to our use of antibiotics, we analyzed DNA sequences recovered from Late Pleistocene permafrost sediment. The samples were collected east of Dawson City, Yukon, at the Bear Creek (BC) site (Fig. 1); prominent forms of ground ice (ice wedges and surface icings) are preserved in the exposure, immediately overlain by a distinctive volcanic ash layer, the Dawson tephra (*11, 12*) (Figs. S1-S2, Table S1). The tephra has been dated at several sites in the area to ca. 25,300 radiocarbon (^{14}C) years BP, or about 30,000 calendar years (*11, 13*). The cryostratigraphic context is similar to other sites in the area preserving relict permafrost and indicates that the permafrost has not thawed since the time of deposition (SOM, section 1). In the absence of fluid leaching, the site represents an ideal source of uncontaminated and securely dated ancient DNA.

Two frozen sediment cores (BC1 and BC4), 10 cm apart, were obtained 50 cm below the tephra. In accordance with appropriate protocol (*14*), we monitored contamination introduced during coring by spraying the drilling equipment and the outer surface of the cores with high concentrations of *E. coli* harbouring the *gfp* (green fluorescent protein) gene from *Aequorea victoria* (SOM, section 2a).

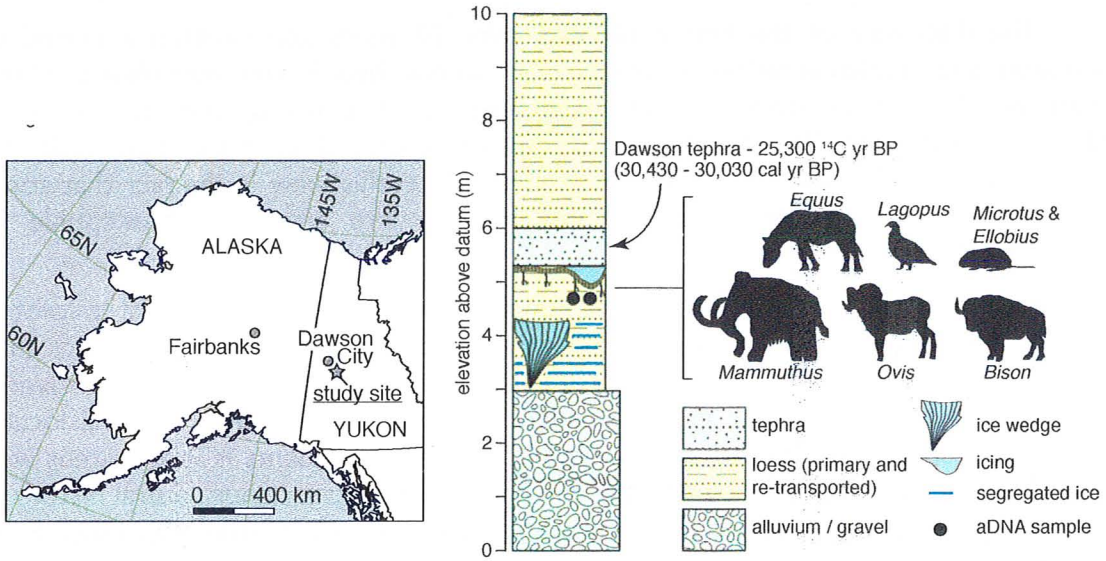


Fig. 1. Stratigraphic profile and location of Bear Creek site. Elevation in meters above base of exposure. Permafrost samples from below Dawson tephra were dated to ca. 30,000 years BP. Preservation of the ice below and above the sample indicates the sediments have not thawed since deposition. Silhouettes represent mammals and birds identified from ancient DNA that are typical of the regional late Pleistocene environment.

After fracturing the samples (Fig. S3), total DNA was extracted from a series of five subsamples taken along the radius of each core (SOM, sections 2b-4). Quantitative PCR (qPCR) analysis confirmed extremely high yields of *gfp* on both core exteriors, with 0.1% or less of this amount at the centre (SOM, section 5, Fig. S4). This supports negligible leaching and cross-contamination during sub-sampling.

A crucial step lending support for the authenticity of the ancient DNA was to confirm the presence of DNA derived from flora and fauna characteristic of a late Pleistocene age, and the absence of common modern or Holocene floral and faunal sources. To explore the vertebrate and plant diversity, we amplified fragments of the mitochondrial 12S rRNA and chloroplast *trn*L and *rbc*L genes (Table S3). Products were sequenced using the 454 GS-FLX platform and identified by BLAST analysis of GenBank sequences (SOM, section 6).

The vertebrate sequences included abundant Late Pleistocene megafauna such as *Bison*, *Equus*, and *Ovis*, as well as rodents (*Microtus*, *Ellobius*) and the rock ptarmigan, *Lagopus mutus* (Fig. S6, Table S5). *Mammuthus* was detectable at low copy numbers using a mammoth-specific qPCR assay, which is consistent with the low ratio of these fossils relative to bison and horse in the region (*12, 15*). The *rbc*L and *trn*L sequences revealed many plant groups that are also well documented in Beringia, including the grasses *Poa* and *Festuca*, sage (*Artemisia*), and willow (*Salix*) (*16*) (Figs. S7-S8, Tables S6-S7). Importantly, no sequences of common Holocene vertebrates (e.g. elk, moose) or plants (e.g. spruce) were identified despite sequence conservation across the primer binding sites; these results are consistent with other reports (*17, 18*) that have argued against DNA leaching in permafrost environments.

We focused our investigation of bacterial 16S rRNA sequences on the Actinobacteria, known for their ability to synthesize diverse secondary metabolites, and for harboring antibiotic resistance genes (*4*). Deep sequencing of 16S amplicons (SOM, section 6) revealed genera commonly found in soil and permafrost microbial communities (*19*) including *Aeromicrobium*, *Arthrobacter*, *Frankia*, and other Actinobacteria that are well known for producing cytotoxic secondary metabolites, including antibiotics (Fig. S9, Table S8). Analysis of the sequences derived from control reactions (Table S4) suggested that modern contaminants do not contribute to the aDNA sequences; not only were the copy numbers 1,000-30,000x lower than from the permafrost extracts, but with the exception of unclassified bacteria, there was also very little overlap in the genera identified (Fig. S9, Table S8). Querying the core sequences against the contaminant data set using BLAST further confirmed their disparity: only 1% of the reads had 95-100% identity to a contaminant sequence, with a single sequence showing 100% identity.

We next developed a series of assays to detect genes encoding resistance to several major classes of antibiotics, and that represent diverse strategies of drug evasion (e.g. target modification, target protection, enzymatic drug inactivation) (SOM, sections 6 and 8). Determinants included the ribosome protecting protein TetM that confers resistance to tetracyclines, the D-Ala-D-Ala dipeptide hydrolase VanX that is a component of the vancomycin resistance cluster, the aminoglycoside modifying acetyltransferase AAC(3), a penicillin-inactivating β-lactamase related to the TEM group (Bla), and the ribosome methyltransferase Erm that protects against macrolides, lincosamides, and type B streptogramins. Amplification of *vanX*, *tetM* and *bla* fragments was successful and triplicate PCR products from multiple extracts were sequenced.

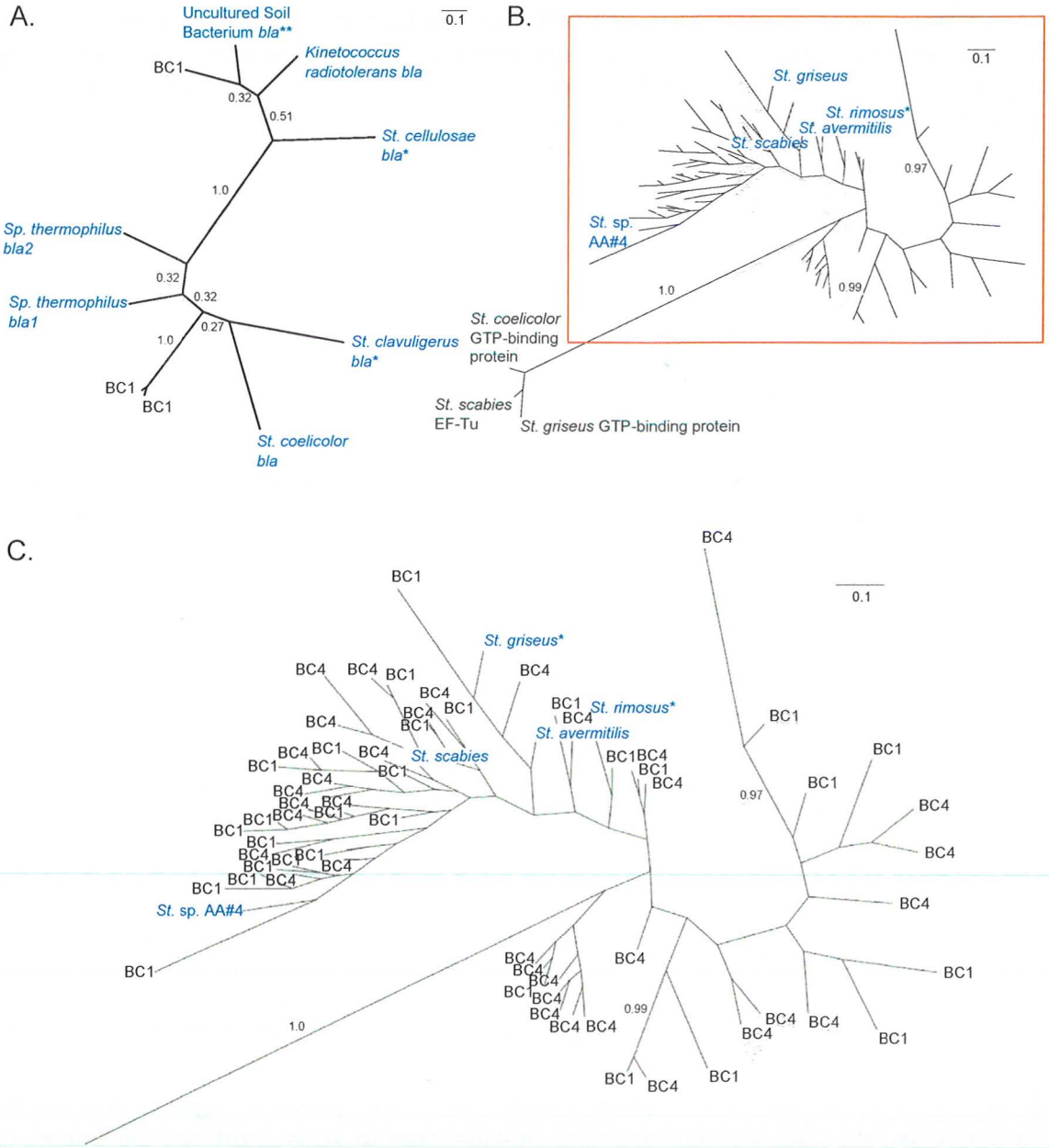


Fig. 2. Genetic diversity of ancient antibiotic resistance elements. Unrooted phylogenies of translated *bla* (**A**), and *tetM* (**B-C**) sequences; (C) details the highlighted section in (B). Blue denotes predicted resistance enzymes, and green those associated with other functions; permafrost-derived sequences are labeled by the originating core name. Sequences where resistance activity has been biochemically verified are noted with a single asterisk (*) (SOM, section 6f). The scale bar represents 0.1 substitutions per site. For (B) and (C), posterior support values are shown.

We identified a number of *tetM*-related genes in the permafrost, the majority of which were most closely related to the actinomycete subset of ribosomal protection proteins, including the biochemically characterized self-resistance element OtrA from the oxytetracycline producer *Streptomyces rimosus* (*20*) (Figs. 2B-C). The β-lactamase sequences demonstrated amino acid identities between 53 and 84% with known determinants and clustered with one of two groups of enzymes: characterized β-lactamases from streptomycetes and uncharacterized β-lactamase-like hydrolytic proteins (Figs. 2A, S14). Most intriguing was the identification of *vanX* gene fragments, which spanned the entire phylogenetic space of characterized vancomycin resistance determinants found in the clinic and the environment. These branch away from the cellular dipeptidases that are the likely progenitors the *vanX* family (Fig. S10).

Vancomycin resistance took the clinical community by surprise when it emerged in pathogenic enterococci in the late 1980s (*21*). In both clinical pathogens (*21*) and contemporary soil environments (*4*), resistance results from the acquisition of a three-gene operon *vanH-vanA-vanX* (*vanHAX*). These enzymes collectively reconstruct bacterial peptidoglycan to terminate in D-alanine-D-lactate in place of the canonical D-alanine-D-alanine, which is required for vancomycin binding and subsequent antibiotic action. While most forms of resistance are attributed to a single gene, this complex mechanism is exclusively associated with resistance and thus its presence provides unambiguous confirmation of its role as a resistance determinant.

With few exceptions, the *vanHAX* operon is invariant in genetic organization; it therefore offers a matchless template to confirm its presence using PCR assays that span the *vanHA* and *vanAX* boundaries. Two short qPCR assays were designed to confirm this contiguity (SOM, sections 7a-b, Fig. 3A). Positive results, including particularly high yields of the smallest amplicon, A6X (Table S9), encouraged us to attempt amplification across both boundaries (i.e. the complete *vanA* gene) in a single 1.2 kb amplicon. We also targeted fragments anchored on either boundary and extending as far as possible into *vanA*. None of the sequences from these products, or those generated by an independent lab (SOM, section 7c), were present in GenBank. Importantly, no contaminants were detected in more than 300 control reactions.

Phylogenic analyses showed that many of the ancient *vanHAX* sequences cluster with characterized glycopeptide-resistant strains of Actinobacteria containing *vanHAX* cassettes (e.g. streptomycetes, glycopeptide-producing *Amycolatopsis* species, and the nitrogen-fixing *Frankia* sp. EAN1 spec) (Figs. 3B, S11-12). Another group falls between the actinobacterial sequences and the Firmicutes-derived cluster, which includes environmental *Paenibacillus* isolates and the pathogenic *Enterococci*, and may reflect an intermediate group.

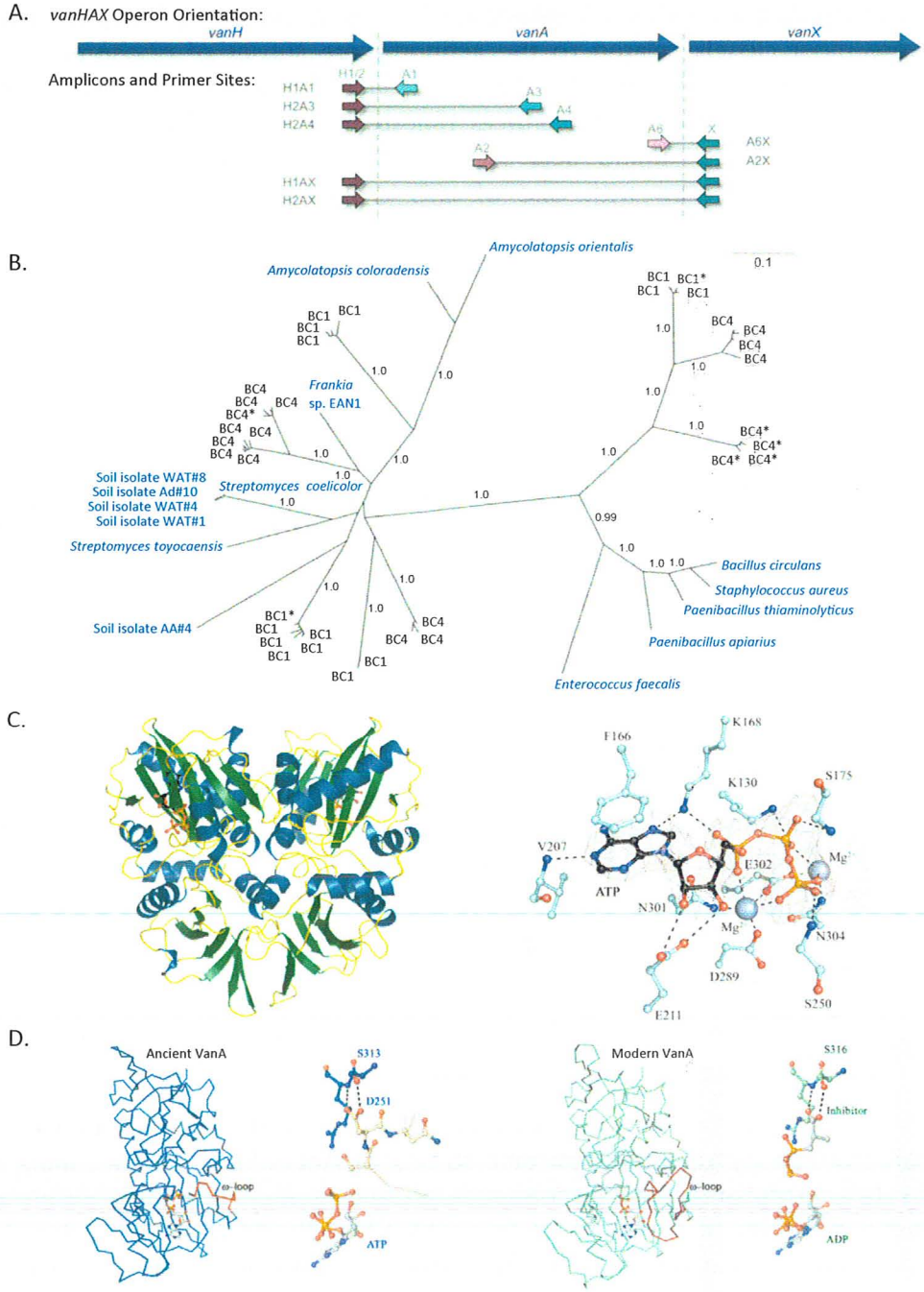


Fig. 3. Ancient vancomycin resistance elements. (**A**) *vanHAX* amplicons used in this study, with primer names noted above each arrow. (**B**) Unrooted phylogeny of translated *vanA* sequences; blue denotes glycopeptide-resistant strains with *vanHAX* clusters confirmed to confer resistance; sequences containing stop codons but homology throughout are noted with a single asterisk (*) (SOM, section 7b). (**C**) $VanA_{A2}$ structure; ribbon diagram of the $VanA_{A2}$ dimer colored by secondary structure (left), and ball-and-stick representation of ATP binding (right); electron density shown is an Fo-Fc map contoured at 3σ. (**D**) Comparison of modern and ancient VanA monomer structures; the ω-loop is colored red and detailed in the ball-and-stick figures; ligands are shown in grey. Dashed lines represent hydrogen bonds.

Permutation tests were performed using the PRSS algorithm (22) (1000 permutations each) to confirm that the sequences were statistically similar to vancomycin resistance genes (*vanHAX*) present in modern *Streptomyces*. As shown in Table 1, all *vanHA*-spanning clones have significant similarity to *vanH* and *vanA*, while all *vanAX*-spanning clones have significant similarity to *vanA* and *vanX*, respectively.

Table 1. *vanHAX* permutation tests

Amplicon	#	Length (bp)	Probability of similarity by chance alone to *S. coelicolor genes**		
			vanH	*vanA*	*vanX*
H1A1	164	203-213	3.59e-03	4.39e-17	0.24
H1A1**	12	209-216	2.83e-03	8.16e-16	0.28
H2A3	24	573-605	9.83e-03	1.27e-54	0.22
H2A4	79	666-681	4.33e-03	6.15e-53	0.18
A6X	159	170-179	0.11	6.87e-08	5.64e-09
A6X**	11	176-179	0.04	2.96e-08	3.63e-09
A2X	96	735-796	0.11	1.80e-59	1.35e-06
HAX***	40	1173-1204	5.95e-03	9.32e-92	6.47e-07

*Gene cluster confers VanB phenotype; **Clones from independent replication in France; ***includes both H1AX and H2AX

To ascertain whether the complete *vanA* sequences are indeed functional and do not represent PCR artifacts or pseudogenes, we synthesized four open reading frames from the 40 H1AX/H2AX sequences (SOM, section 7d). Two of these generated soluble proteins suitable for purification to homogeneity. Enzymatic characterization indicated that these ligases were indeed D-alanine-D-lactate-specific (Fig. S13), and steady state analysis revealed steady state kinetic parameters consistent with contemporary enzymes derived from both the clinic and the environment (Table S10). These results clearly demonstrate that the *vanHAX* genes identified in the ancient samples encode enzymes capable of bona fide antibiotic resistance.

We further confirmed the link between 30,000-year old VanA and contemporary enzymes by determining the 3D-structure of $VanA_{A2}$ by X-ray crystallography (SOM, section 7e, Table S11). The quaternary and tertiary structures of $VanA_{A2}$, crystallized in the ATP-bound form, show the overall D-Ala-D-X ligase fold of modern enzymes including VanA from vancomycin-resistant *Enterococcus faecium* (*23*) (Fig. 3C). Superposition of ancient and modern VanA (Fig. 3D) reveals conservation of quaternary and tertiary structure with minor differences in Mg^{2+} and ATP γ-phosphate coordination. The ω-loop comprises the biggest structural change; thirteen N-terminal residues (233-246) are absent in the diffraction pattern of $VanA_{A2}$, including His241 (His 244 in modern VanA), responsible for the lactate selectivity. The last seven ω-loop residues (247-253) have clear electron density, undergoing a dramatic 13 Å shift. These structural differences however are not reflected in enzyme function.

This work firmly establishes that antibiotic resistance genes pre-date our use of antibiotics and offers the first direct evidence that antibiotic resistance is an ancient, naturally occurring phenomenon widespread in the environment. This is consistent with the rapid emergence of resistance in the clinic and predicts that new antibiotics will select for pre-existing resistance determinants that have been circulating within the microbial pangenome for millennia. This reality must be guiding principle in our stewardship of existing and new antibiotics.

ACKNOWLEDGMENTS

This work was supported by Canada Research Chairs to H.N.P. and G.D.W, Canadian Institutes of Health Research operating grant to G.D.W. (MOP-79488) and scholarship to V.D., as well as grants from the Natural Sciences and Engineering Research Council of Canada to H.N.P. and scholarship to C.E.K.

The authors declare no competing interests.

REFERENCES

1. D. M. Livermore, *J. Antimicrob. Chemother.* **64**, i29 (2009).
2. G. D. Wright, *Nature Rev. Microbiol.* **5**, 175 (2007).
3. V. M. Hughes, N. Datta, *Nature* **302**, 725 (1983).
4. V. M. D'Costa, K. M. McGrann, D. W. Hughes, G. D. Wright, *Science* **311**, 374 (2006).
5. G. Dantas, M. O. A. Sommer, R. D. Oluwasegun, G. M. Church, *Science* **320**, 100 (2008).
6. M. O. A. Sommer, G. Dantas, G. M. Church, *Science* **325**, 1128 (2009).
7. H. K. Allen, L. A. Moe, J. Rodbumrer, A. Gaarder, J. Handelsman, *ISME J.* **3**, 243 (2009).
8. R. H. Baltz, *SIM News* **55**, 186 (2005).

9. B. G. Hall, M. Barlow, *Drug Resist. Update.* **7**, 111 (2004).
10. S. Z. Mindlin, V. S. Soina, M. A. Petrova, Z. M. Gorlenko, *Russ. J. Genet.* **44**, 27 (2008).
11. D. G. Froese, G. D. Zazula, A. V. Reyes, *Quat. Sci. Rev.* **25**, 1542 (2006).
12. D. G. Froese *et al.*, *GSA Today* **19**, 5 (2009).
13. F. Brock, D. G. Froese, R. G. Roberts, *Quat. Geochronol.* **5**, 625 (2010).
14. E. Willerslev, A. J. Hansen, H. N. Poinar, *Trends Ecol. Evol.* **19**, 141 (2004).
15. C. R. Harington, F. V. Clulow, *Can. J. Earth Sci.* **10**, 697 (1973).
16. G. D. Zazula *et al.*, *Nature* **423**, 603 (2003).
17. J. Haile *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 22352 (2009).
18. J. Haile *et al.*, *Mol. Biol. Evol.* **24**, 982 (2007).
19. D. Gilichinsky *et al.*, in *Psychrophiles: From Biodiversity to Biotechnology,* R. Margesin, F. Schinner, J.-C. Marx, C. Gerday, Eds. (Springer-Verlag, Berlin, 2008), pp. 83-102.
20. D. Doyle, K. J. McDowall, M. J. Butler, I. S. Hunter, *Mol. Microbiol.* **5**, 2923 (1991).
21. P. Courvalin, *Clin. Infect. Dis.* **42**, S25 (2006).
22. W. R. Pearson, D. J. Lipman, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2444 (1988).
23. D. I. Roper, T. Huyton, A. Vagin, G. Dodson, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8921 (2000).

SUPPLEMENTARY ONLINE MATERIAL

Antibiotic resistance is ancient

Vanessa M. D'Costa[1,2]*, Christine E. King[3,4]*, Lindsay Kalan[1,2], Maryia Morar[1,2], Wilson Sung[4], Carsten Schwarz[3], Duane Froese[5], Grant Zazula[6], Fabrice Calmels[5], Regis Debruyne[7], G. Brian Golding[4], Hendrik N. Poinar[1,3,4†], Gerard D. Wright[1,2†]

[1]Michael G. DeGroote Institute for Infectious Disease Research, McMaster University, Hamilton, Ontario, Canada, L8N 3Z5, [2]Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario, Canada, L8N 3Z5, [3]McMaster Ancient DNA Centre, Department of Anthropology, McMaster University, Hamilton, Ontario, Canada, L8S 4L9, [4]Department of Biology, McMaster University, Hamilton, Ontario, Canada, L8S 4K1, [5]Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, Alberta, Canada, T6G 2E3, [6]Yukon Palaeontology Program, Department of Tourism and Culture, Yukon Government, P.O. Box 2703, Whitehorse, Yukon, Canada, Y1A 6V6, and [7]Muséum National d'Histoire Naturelle, UMR 7206 Eco-anthropologie, 57 rue Cuvier, CP139, 75231 Paris Cedex 05

* These authors contributed equally to this work

†To whom correspondence should be addressed: Email wrightge@mcmaster.ca (G.D.W.); poinarh@mcmaster.ca (H.N.P.)

1. CRYOSTRATIGRAPHIC OBSERVATIONS

The Bear Creek site (63.98° N, 138.22° W) is located 11 km east of Dawson City, Yukon, in the Klondike placer mining district (Fig. 1, main text). Mining activities at the site exposed the (vertical) sampling surface, which was described and sampled in June and July of 2007. The sampling site is within a prominent unit of ice-rich loess (wind-blown silt), hosting relict ground ice that began accumulating in the region about 34,000 years ago (29,000 radiocarbon (^{14}C) years BP) (*1-4*).

The exposure contains a prominent volcanic ash layer, the Dawson tephra, which was identified by major element geochemistry by electron microprobe at the University of Alberta (Table S1, Figs. S1, S2). The Dawson tephra is the most common tephra bed in the area and is easily distinguished from other tephras by its major element geochemistry (Fig. S1). The tephra is dated to 25,300 ^{14}C years BP, or about 30,000 calendar years ago (*1, 5, 6*) and provides chronology for the cores in this study.

Table S1. Average major element glass composition (%) of tephra collected at Bear Creek, and Dawson tephra reference glass from the Quartz Creek site (*2*)

Component	Bear Creek (UA 1878)	Dawson (UA 1517)
SiO_2	74.34 (0.42)	74.17 (0.40)
TiO_2	0.26 (0.02)	0.26 (0.02)
Al_2O_3	13.76 (0.25)	13.77 (0.24)
FeOt	2.05 (0.11)	2.06 (0.11)
MgO	0.23 (0.02)	0.23 (0.02)
CaO	1.29 (0.08)	1.29 (0.08)
MnO	0.06 (0.04)	0.06 (0.04)
Na_2O	4.39 (0.24)	4.53 (0.21)
K_2O	3.41 (0.10)	3.42 (0.08)
Cl	0.22 (0.02)	0.22 (0.03)
H_2Od	3.45 (1.86)	3.20 (2.14)
n	33	20

Glass analyses completed using wavelength-dispersive analyses on a Cameca SX 100 microprobe at the University of Alberta (15 keV accelerating voltage, 10 um beam diameter, and 6 nA current); 10 um beam diameter, 6 nA current). FeOt is total iron oxide as FeO; H_2Od, estimated water by difference; *n*, number of analyses; standard deviations of the normalized values are given in parentheses.

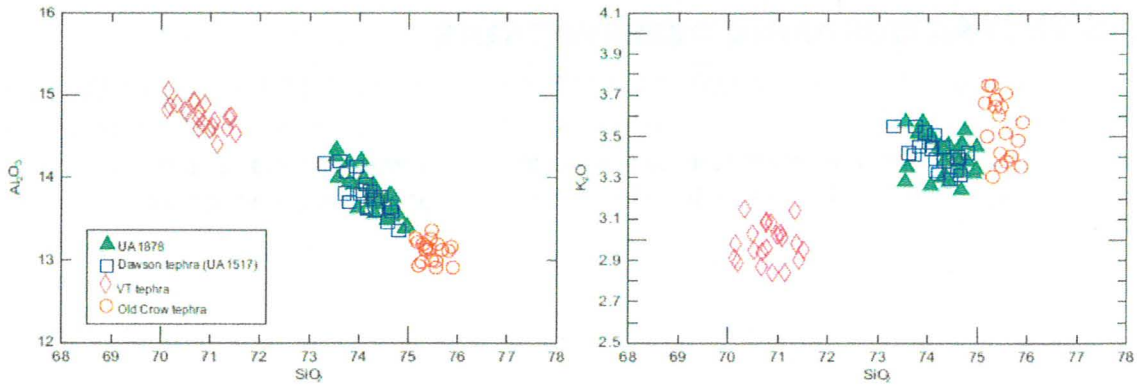


Fig. S1. Geochemistry data. Bivariate plots of major element glass geochemistry of the Dawson tephra from the Bear Creek site (UA1878) and a reference sample of Dawson tephra from Quartz Creek dated to 25,300 ^{14}C yrs BP (2, 5). Major element chemistry easily distinguishes the Dawson tephra from other common tephras in the region (here shown with SiO_2-Al_2O_3 and SiO_2-K_2O plots).

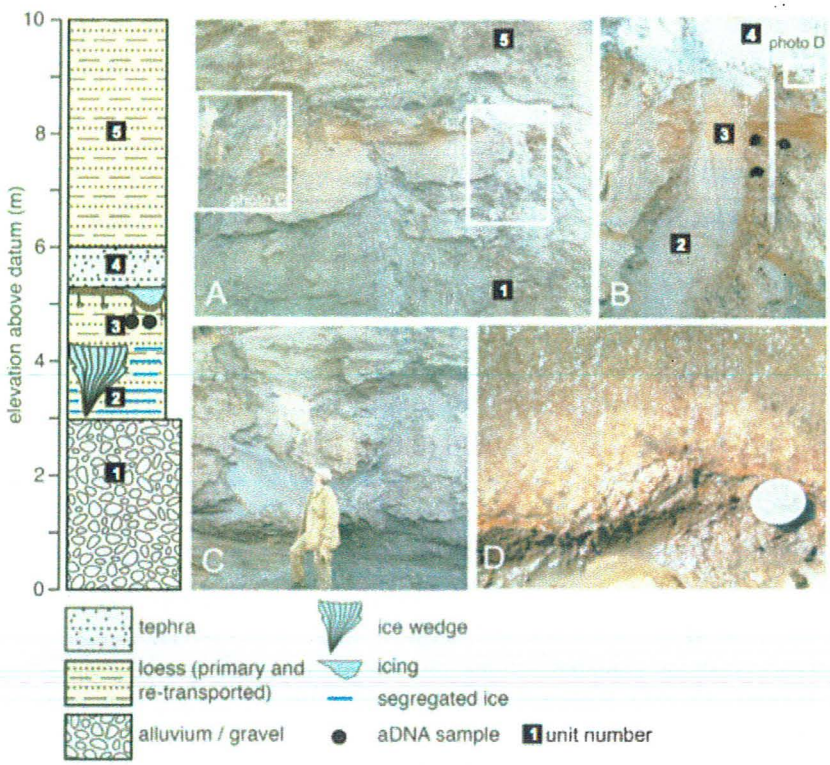


Fig. S2. Detailed stratigraphic profile of Bear Creek. (**A**) Study site with inset boxes of photo B and C showing sample locations. (**B**) Sampling site where two cores were drilled from paleo-active layer in Unit 3; upper two samples in photo represent the two analyzed core samples (ice axe 80 cm long). (**C**) Ice wedge overlain by Dawson tephra (white sediment in centre of photo), about 5 m lateral to the sampling site (see inset box in A for location). (**D**) Close up of pool ice preserved on surface of Unit 3 (paleo-active layer) including plant macrofossils entombed in the ice.

The sampled sequence consists of 3 m of alluvial sediment overlain by 7 m of largely ice-rich eolian silt (or loess), containing prominent ground ice bodies. We divide the ice-rich silt into three units that we describe in turn below from the base of the exposure to above the sampling level.

Unit 1 is the lower unit of silt from 3 to 4.2 m above the base of the exposure, above the alluvial gravel, and contains an abundance of ice, including vertically-foliated ice bodies (ice wedges). The ice wedges have stepped lateral boundaries, or shoulders, indicating that the wedges grew when permafrost aggraded upward (syngenetically) with loess deposition (Fig. S2). The top of these ice wedges is truncated by a thaw contact that extends to the base of the overlying Unit 2, about 0.9 m below the Dawson tephra. The ice-rich silts of Unit 1 contain regular horizontal ice lenses ranging from 1-20 mm thick, and rare vertical ice cracks, attributable to cryodessication as the permafrost table aggraded (*7*). The presence of the regular, thin horizontal ice lenses suggests that the permafrost table aggraded slowly enough to allow the development of segregated ice lenses; this process is common in aggrading permafrost (*8-10*) and common to most of the ice-rich silts present in the Klondike area.

Unit 2, from 4.2 to ~5.2 m, is a prominent reddish-brown, ice-poor silt that extends across the exposure, and is sharply overlain over most of that extent by a thick occurrence (up to 80 cm) of the Dawson tephra (Fig. S1). The bottom of this unit truncates the surface of the underlying ice wedges, while the top of this unit is either directly overlain by the Dawson tephra, or a clear ice body is present which infills a slight depression, up to 30 cm, on the surface of this unit. This ice body is massive pure ice, and locally preserves an in situ vegetation surface within the ice (Figs. S1B and S1D). The ice body shows no evidence of lensing or disruption of the surrounding sediments, suggesting that it formed as a pool of ice that accumulated on the surface prior to deposition of the Dawson tephra. The silts of Unit 2 are 'dry' permafrost, lacking completely in visible ice lenses, though some moisture, up to ca. 10% is present, suggesting a faster freezing gradient not permitting segregation to take place (*11, 12*). This condition occurs annually when permafrost freezes back the active layer during the fall season (*13, 14*). The Dawson tephra locally preserves a sparse vegetation surface, and the surface pool ice found on low spots on the surface of Unit 2.

The contact between ice and ash is sharp, consistent with rapid deposition and absence of subsequent thaw of the tephra or underlying silts, where the DNA samples were collected. It is likely that the deposition of the tephra was sufficiently rapid such that the active layer was shifted upward, above the level of the icing and paleoactive layer; this would prevent subsequent thaw of the pool ice and inhibit segregation of even minor pore ice within the paleoactive layer. This condition is similar to other sites in the Klondike where surface icings and buried snowbanks are preserved below the Dawson tephra by the rapid sedimentation on the surface (*1*). Observations made on the paleoactive layer unit, suggest that this loess material was frozen back during the fall season preceding the tephra deposition. Collectively, these unthawed surficial ice bodies and the paleoactive layer observations, indicate that the permafrost has not thawed since the time of deposition, about 30,000 years ago, which strongly argues against any water migration in these sediments post-deposition.

2. SAMPLING PROCEDURE

2.a) Coring

Excavations were made into frozen sediments, generally less than 30 cm from the surface. The cores were stored individually in plastic bags and shipped frozen to the McMaster Ancient DNA Centre, where they were stored at -40°C. All subsequent procedures (prior to PCR cycling) were performed in dedicated clean rooms, physically separated from laboratories containing thermocyclers and PCR products.

To monitor and control for surface contamination (*15-18*), we sprayed the cores with a lab strain of *Escherichia coli* harboring a jellyfish gene, *gfp*. We prepared an *Escherichia coli* culture harboring the ampicillin-resistant pET23a-*gfp* construct (obtained from Dr. R. Truant). The construct was transformed into *E. coli* BL21 (DE3) and inoculated into a 25 mL culture of Luria-Bertani (LB) broth supplemented with 100 μg/mL of ampicillin. After overnight growth at 37°C and 250 rpm, it was subcultured at 1% (v/v) into a 1 L culture of LB supplemented with ampicillin. After growth at 37°C for approximately 8 hours, the culture was stored at 4°C during transportation to the sampling site.

2.b) Subsampling

Subsamples are often taken after cutting away the outer layer of each core; however, this cutting motion may spread contaminants over the newly exposed surface. To prevent this, two shallow, parallel cuts were made approximately 5 cm apart along the circumference of the cores using a Dremel tool; the cores were then easily fractured at these points using a hammer and chisel, creating cross-sectional disks with untouched surfaces (Fig. S3).

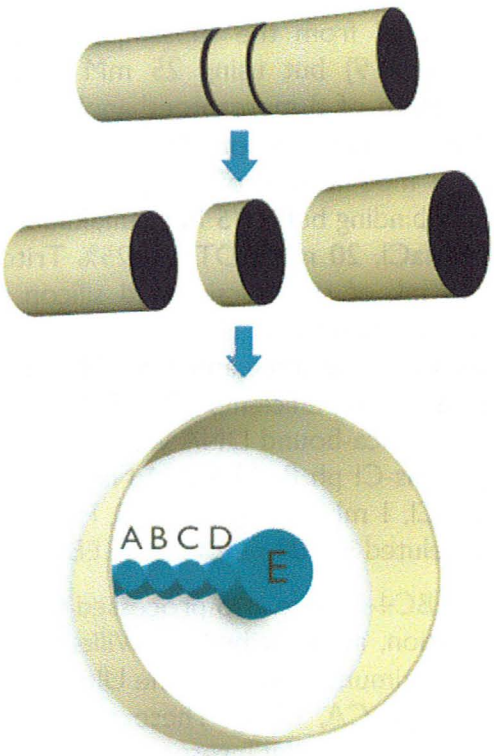


Fig. S3. Subsampling procedure. A disk approximately 5 cm thick was isolated by fracturing the larger core and subsamples A-E taken from the newly exposed surface.

Using a stationary drill press, subsamples were taken along the radius of each disk. Working from the center outward, this included a sample 1 cm in diameter (denoted as E), and four 0.5 cm samples denoted D, C, B, and A. We expected high concentrations of the *gfp* control in sample A, and decreasing concentrations toward the center of each disk, depending on the degree of thawing and leaching since coring.

To minimize melting, each sample was processed independently and tools were cooled in liquid nitrogen whenever possible. All tools were decontaminated with commercial bleach (~5%) and either baked at 375°F for 30 min or treated with UV light (100 $\mu J/cm^2$) for 20 min, depending on their material. To monitor potential airborne contamination introduced during subsampling, tubes containing 1 mL of 0.1x TE buffer (10 mM Tris-Cl pH 8, 0.1 mM EDTA) were placed on all work surfaces and left open throughout the procedure. Additional air controls were set up a few days before and several weeks later to monitor any potential contaminants normally present in the room.

3. DNA EXTRACTION

Total DNA was extracted from the subsamples (~250 mg) and air controls according to Willerslev et al., (*19*) but using 25 mM TCEP (Tris (2-carboxyethyl) phosphine hydrochloride) in place of β-mercaptoethanol and with a 10 min vortex in Lysing Matrix E tubes (MP Biomedicals, Solon, OH, USA) for sample disruption. The extracts were purified using a modification of the protocol of Boom et al. (*20*): first, they were mixed with 4 mL of binding buffer (5 M guanidinium thiocyanate (GuSCN), 50 mM Tris-Cl pH 8, 22.5 mM NaCl, 20 mM EDTA, 1.25% Triton-X 100) that had been previously incubated with 50 mL of size-fractionated silicon dioxide for 30 min with rotation at room temperature. Additional GuSCN was added to maintain a 5 M concentration upon addition of the extract and the pH was adjusted to 4.5-5.5 with glacial acetic acid to maximize binding efficiency (*21*). Following incubation of at least 1 hour at room temperature, the silica-bound DNA was washed twice with 1 mL of wash buffer (5 M GuSCN, 50 mM Tris-Cl ph 8, 22.5 mM NaCl) and once with 1 ml of 80% ethanol in 1x TE (10 mM Tris-Cl, 1 mM EDTA), pH 7.5. Finally, the pellets were dried at 56°C for 5 min and the DNA eluted in 100 µL of Buffer EB (Qiagen, Hilden, Germany).

Subsamples BC1-E and BC4-E were also processed following the above protocol, but using columns for purification, as described by Willerslev et al. (*19*) and following the "alternative protocol for maximum yields" for the UltraClean Soil DNA Isolation Kit (MO BIO Laboratories, Carlsbad, CA, USA). These additional extracts are identified elsewhere with the notations "per" and "kit", respectively. Negative extraction controls were included with all three methods.

4. PCR INHIBITION ASSESSMENT

All extracts, including air controls and extraction blanks were tested for PCR inhibition as described by King et al. (*22*). Based on these experiments, a specific dilution and amount of Taq DNA polymerase to be used in PCR were determined for each extract, in order to maximize the potential recovery of DNA and the accuracy of our qPCR assays. The best conditions and their inhibition test performance are indicated in Table S2.

5. CONTAMINATION ASSESSMENT

Quantitative PCR (qPCR) assays for human mitochondrial DNA (mtDNA) and the *gfp* construct were designed to evaluate contamination levels across the radii of the cores. Primers and qPCR conditions are shown in Table S3.

Table S2. Inhibition test results

Extract	Dilution Factor	Taq (U)	ER (%)	E (%)
Samples				
BC1-A	1	1	26	86
BC1-B	1	2.5	31	90
BC1-C	1	2.5	32	90
BC1-D	1	2.5	24	90
BC1-E	2	2.5	20	94
BC1-E kit	1	2.5	4	68
BC1-E per	1	2.5	33	95
BC4-A	1	1	43	82
BC4-B	1	1	25	84
BC4-C	1	1	24	87
BC4-D	1	2.5	35	94
BC4-E	1	2.5	42	90
BC4-E kit	1	2.5	11	78
BC4-E per	1	1	47	93
Air Controls				
ACB	1	1	47	96
ACH	1	1	50	98
Extraction Blanks				
BC-B1	1	1	50	94
BC-B2	1	1	51	97
kit blk	1	1	55	106
per blk	1	1	69	98
AC-B1	1	1	53	99
AC-B2	1	1	54	96

ER, expected recovery; E, relative amplification efficiency (*18*).

For the *gfp* construct, primers were designed to amplify a 225 bp fragment based on the sequence from *A. victoria* (Genbank accession number X83959). To prepare qPCR standards, the *gfp* fragment was isolated from pET23a-*gfp* by digesting with BglII, followed by agarose gel electrophoresis and purification using the QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany). The purified products (in 1x TE) were then quantified by fluorometry and serial dilutions (from 1e5 to 1e-1 copies/μL) used for absolute quantification in qPCR. Based on repeated amplification of the low copy standards, we determined the sensitivity of the assay to be approximately 1 copy/μL; the 1 copy/μL standard was consistently amplified, while the 0.1 copy/μL standard amplified in roughly 10% of all reactions, as expected.

Human mtDNA standards were prepared from a pool of PCR products (primers L16055 to H16410 from Handt et al. (*23*), which are external to the qPCR locus). The pool was purified using the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany), quantified, and used as described above; sensitivity was between 1-10 copies/µl.

To preserve the extracts, each was screened only once using these assays. Standards and no template control (NTC) reactions were included with both amounts of Taq (1U and 2.5U).

The approximate concentrations of *gfp* and human mtDNA are shown in Fig. S4. A clear pattern was observed for *gfp*, with the highest concentration on the exterior of the cores (Fig. S3, subsample A), as expected, and 0.1% or less of that maximum reaching subsample B. This may be the result of leaching, as the exterior of the core begins to melt, or simply a consequence of the proximity of subsamples A and B (the drill bit used to sample A touches the edge of sample B; there is no space between them). In either case, these results indicate that the fracturing procedure was successful in ensuring that the core centers are free of contaminants. Extract BC1-E did contain traces of *gfp*, however; since other extracts of this sample, all control extracts, and the NTCs were clean, it is unclear at which point this stochastic contamination occurred, but not surprising given the inordinate amount of this particular template present on the surface of the core. Low-level human DNA contamination (<3 copies/µl) was detected in some extracts, but also in the NTC reactions for this assay, suggesting contamination in lab reagents rather than from the sampling procedures.

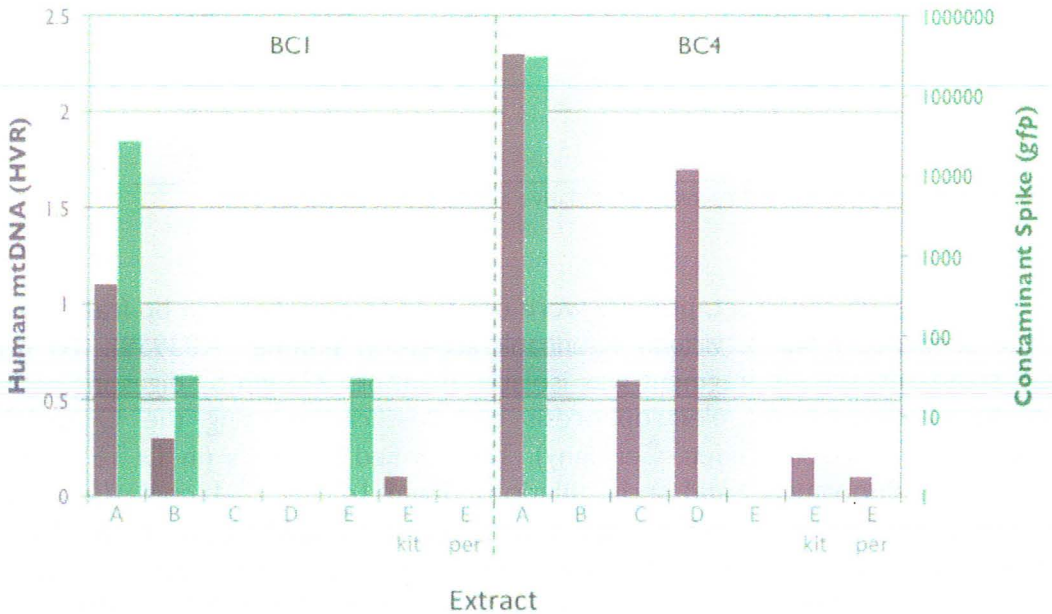


Fig. S4. Contamination assessment. Approximate copy numbers of human mtDNA and *gfp* contaminants in each extract, measured in copies/µL.

Table S3. PCR assays

Assay	Target	Primer Name	Primer Sequence (5'-3')	Size (bp)	Ta (°C)	PCR Notes[1]	Sensitivity[2] (copies/mL)	Source
Contamination Assessment								
HVR	Human (mtDNA)	L16209 H16303	CCCCATGCTTACAAGCAAGT TGGCTTTATGTACTATGTAC	131	57	...	1-10	*(19)*
gfp	E. coli plasmid pET23a-GFP	gfp F gfp R	TGATGCAACATACGGAAAACTTACCCTTA CACGTGTCTTGTAGTTCCCGTCATCT	225	65	...	1	...
454 Amplicon Sequencing								
12S	Vertebrates (mtDNA)	12Sa' 12So	CTGGGATTAGATACCCCACTAT GTCGATTATAGGACAGGTTCCTCTA	150	57	No SG	N/A	*(57)*
16S	Actinobacteria	16S F 16S R	GGCGAACGGGTGAGTAACACG CCATTGTGCAATATTCCCCACTG	275	58	10% DMSO; 60 sec denaturation	N/A	...
*rbc*L	Plants (cpDNA)	rbcL h1a F rbcL h2a R	GGCAGCATTCCGAGTAACTCCTC CGTCCTTTGTAACGATCAAG	138	55	No SG	N/A	(57)
*trn*L	Plants (cpDNA)	trnL 49425 F trnL 49466 R	GGGCAATCCTGAGCCAA CCATTGAGTCTCTGCACCTATC	49-182	55	No SG	N/A	*(49)*
vanX	Glycopeptide resistance	vanX F vanX R	CCAAGTACGCCACSTGGGACAAC CTTCGTCCGGCCGTCCTCC	220	60	10% DMSO; 60 sec denaturation; no SG	N/A	...
tetM/otrA	Tetracycline resistance	tetM/otrA F tetM/otrA R	CCCCGGGMCACYCCGACTTC CGSCMCGGTCGATCTTGTTSA	175	65	10% DMSO; 60 sec denaturation; no SG	N/A	...

Table S3 (continued). PCR assays

Assay	Target	Primer Name	Primer Sequence (5'-3')	Size (bp)	Ta (°C)	PCR Notes[1]	Sensitivity[2] (copies/mL)	Source
vanHAX Cluster								
H1A1	*vanHA* boundary	H1 F A1 R	CCGCAYAYSGCCTATTACACG GGCCGTCGCAGAGYHKCCAGG	255	66	...	1-10	...
H2A3	*vanHA* boundary	H2 F A3 R	CCGCAYAYSGCCTATTACAC ACACCTTGSTGACGCCGAASGA	645	59	60 sec extension	10-100	...
H2A4	*vanHA* boundary	H2 F A4 R (12-oligo pool)	CCGCAYAYSGCCTATTACAC TCGATCAGCACCTTCGTGTCGTA TCGATCAGGATCTTCGCGTCGTA TCGATCAGGACCTTCGAGTCGTA TCGATCAGGACCTTCGAGACGTA TCGATCAAGACCCTGTCGTCGTA ACGGCTTCTTCGATCAGCACCTT ACGGCCTCTTCGATCAGCACCTT ACAGCCTCTTCGATCAGTACCTT ATCGCCTGCTCGATCACAACCTT ACCGCCTCTTCGATCACAACCTT ACCGCCTCTTCGATCACAATCTT ACCCCCTCTTCGATCACAACCTT	710	60	5% DMSO; 60 sec extension	1-10	...
A6X	*vanAX* boundary	A6 F X R	CAACGAGGTCAACACCCTGCCC CCGGTTTGCCGGTGAAGTTGTC	225	66	...	1-10	...
A2X	*vanAX* boundary	A2 F X R	CATCCCSTACGTSGGCTG CCGGTTTGCCGGTGAAGTTGTC	785	62	5% DMSO; 60 sec extension	1-10	...
H1AX	*vanHA* and *vanAX* boundaries	H1 F X R	CCGCAYAYSGCCTATTACACG CCGGTTTGCCGGTGAAGTTGTC	1225	60	5% DMSO; 90 sec extension	10-100	...
H2AX	*vanHA* and *vanAX* boundaries	H2 F X R	CCGCAYAYSGCCTATTACAC CCGGTTTGCCGGTGAAGTTGTC	1225	59	5% DMSO; 90 sec extension	1-10	...
Other Resistance Assays								
aac(3)	Aminoglycoside resistance	aac(3) F aac(3) R	CTGGAACGACGCYCCGCCSTAC GGSGCGCCSAGSAGCAGYA	300	64	5% DMSO; 2 mM $MgCl_2$; 60 sec denaturation	1-10	...
bla	b-lactam resistance	bla F bla R	CCTCGCCCGGCGCGTCCASTA AGMAGGTTGGCCGCSGTGTTGTC	155	65	60 sec extension	25-50	...
erm	Macrolide resistance	erm F erm R	GCTCTCGCAGAACTTCCTCGC AGMCKGGGGTCGATCTCGTAG	175	60	7.5% DMSO; 1.5 mM $MgCl_2$; 60 sec extension	10-100	...

1) PCRs included 1x PCR Buffer II (Applied Biosystems), 2.5 mM $MgCl_2$, 0.75 mg/mL BSA, 0.3 μM primers, 0.4 mM (each) dNTP, 0.167x SYBR Green I (Invitrogen) in DMSO, (SG) 1 or 2.5 U of AmpliTaq Gold DNA Polymerase (Applied Biosystems), based on the inhibition test results, and 2 μL of extract in a final volume of 20 μL; reactions were cycled 45-55 times (30 sec at 95C, 30 sec annealing, 30 sec at 72C) with a 7 min initial denaturation and 15 min final extension; 2) 95% limit of detection, as per the MIQE guidelines (*58*). Ta, annealing temperature.

6. 454 AMPLICONS

6.a) Amplification and Sequencing

Several PCR assays (Table S3) were used to explore the diversity of vertebrates, plants, and bacteria present in the cores. To explore the origins of antibiotic resistance, we also included assays targeting small fragments of *vanX* (vancomycin resistance) and *tetM* (tetracycline resistance) from *Streptomyces* and closely related genera. These targets represent biochemically-characterized resistance determinants with sequence homology to genes in clinical pathogens (*24-26*).

All PCR products were replicated three times from extracts BC1-E and BC4-D or -E, with the exception of the 16S products (assuming bacterial DNA yields would be high enough from a single PCR that our scale of sequencing would not reach saturation) and the *tetM* assay, for which PCR success was limited (Table S4).

Table S4. PCR results – 454 amplicons

Extract	12S	*rbc*L	*trn*L	16S	*tetM*	*vanX*
Samples						
BC1-E	**3/3**	**3/3**	**3/3**	**1/1**	**1/5**	**3/3**
BC4-D	...	...	...	**1/1**	0/2	...
BC4-E	**3/3**	**3/3**	**3/3**	...	**1/3**	**3/3**
Air Controls						
ACB	0/1	0/1	0/1	0/1	0/1	0/1
ACH	0/1	0/1	0/1	0/1	0/1	0/1
Extraction Blanks						
BC-B1	0/2	0/2	0/3	**1/1**	0/3	0/2
BC-B2	0/2	0/2	0/3	**1/1**	0/3	0/2
AC-B1	**1/1**	0/1	0/1	0/1	0/1	0/1
AC-B2	0/1	0/1	0/1	0/1	0/1	0/1
PCR NTCs						
...	0/7	0/7	0/9	**6/6**	0/12	0/6

Fraction of positive PCRs, with successful amplifications in bold and contamination in red. NTCs, no-template controls.

For each core, the three PCRs were pooled at approximately equal concentrations and re-amplified with 454 fusion primers using the same conditions (Table S3) but with a 2°C increase in annealing temperature. All fusion primers included a 10-nucleotide multiplex identifier (MID) sequence to distinguish the two cores. The 12 libraries were purified using AcroPrep Filter Plates (30K, Omega membrane; Pall, Port Washington, NY, USA) and quantified by fluorometry. The libraries were pooled such that each 16S library would be represented by approximately 2e4 reads, and all others by approximately 2e3 reads.

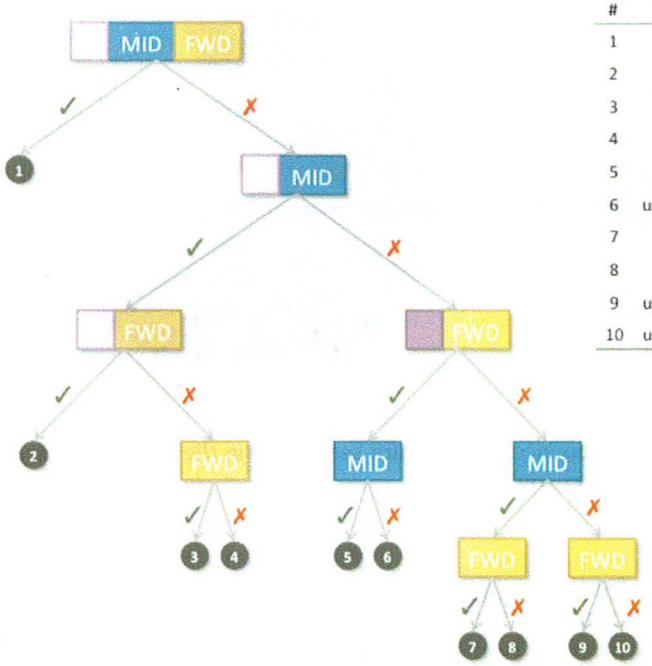
Two additional libraries were prepared in order to identify the source of contaminants. 12S contaminants were detected in one extraction blank from the air control extraction set, while 16S contaminants were detected in both extraction blanks and PCR blanks (Table S4). These were processed as described above and pooled to obtain approximately 100 and 2e3 sequences, respectively. Emulsion PCR and sequencing were carried out according to the GS-FLX protocols (454 Life Sciences, Branford, CT, USA) (27).

6.b) De-multiplexing and Identification

A Perl script was written to trim the PCR primers and de-multiplex the sequences based on their unique combinations of MIDs and forward PCR primers. The algorithm is best explained by a decision tree, as shown in Fig. S5. In some cases we noted extra sequence upstream of the MID or forward primer or downstream from the reverse primer; thus, the algorithm first attempts to find perfect-matching MIDs and primers while allowing for extra sequence (denoted by the blank or shaded boxes in Fig. S5). If that fails, an inexact match is attempted while not allowing for extra sequence. An inexact match means that there are *n* changes compared to the expected MID or primer; *n* for primers may be specified by the user but MIDs were always allowed 0 to 2 changes based on their designed specifications. A change may represent a substitution or a one-nucleotide insertion or deletion.

The sequences were then sorted into bins based on the unique sample and amplicon combinations. Reads that could not be sorted were placed in the 'unknown' category, reads without any sequence between the primers were places in the 'discarded' category, and sequences of 30 nucleotides or less were placed in the 'short' category. We allowed up to 5 changes in the primers and further separated the reads into those with few combined changes (3 or less) and many combined changes (4 or 5) in the primers.

Finally, the sequences were identified using BLAST (word size 15, E value <0.001) and the GenBank database; the top hit for each read is reported in Tables S5-S8.



#	MID	FWD
1	exact	exact
2	exact	exact
3	exact	inexact
4	exact	unidentified
5	inexact	exact
6	unidentified	exact
7	inexact	inexact
8	inexact	unidentified
9	unidentified	inexact
10	unidentified	unidentified

Fig. S5. De-multiplexing algorithm. Each node in the tree represents the search for an MID or forward primer, or both simultaneously. An empty box represents sequence before a perfect-matching MID or forward primer. A shaded box represents sequence that might harbor an MID with one or more differences. An MID or forward primer without any empty or shaded box contains at least one difference.

6.c) Vertebrate 12S Sequences

Based on the top BLAST hit for each sequence, we identified seven vertebrate families, with Bovidae and Equidae appearing most frequently (Table S5 and Fig. S6). With the exception of a small portion of human DNA, the presence of these groups is expected based on fossil remains from Beringia (*28*). We did not detect any mammoth (*Mammuthus*) DNA using this approach; however, further inspection revealed several mismatches between the 12S primers and mammoth reference sequences. An 84 bp qPCR assay with mammoth-specific *cytB* primers (*29*) did reveal trace amounts (<2 copies/μL) of mammoth DNA in these extracts, which mirrors the low ratio of mammoth to bison (*Bison*) and horse (*Equus*) fossils in the region (*2, 30*). Sequences from the contaminant product include only human and chicken DNA, which are documented contaminants of some laboratory reagents (*31*).

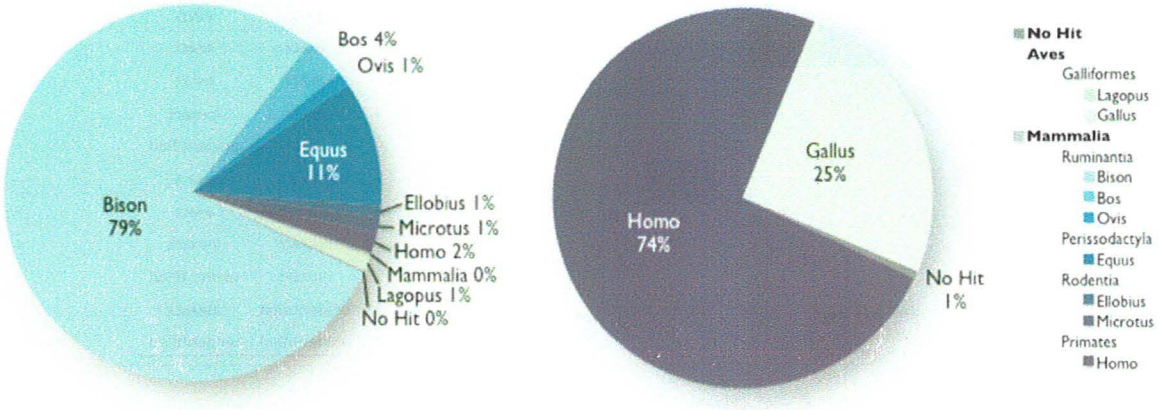


Fig. S6. Top BLAST hits for 12S sequences. Vertebrate sequences in the cores (left) are distinct from those in the contaminant PCR product (right).

Table S5. Top BLAST hits for 12S sequences

Phylum	Class	(Sub)Order	Family	Subfamily	Genus	Species	# Core	# Cont[1]
Chordata	Aves	Galliformes	Phasianidae	Phasianinae	Gallus	gallus	...	40
				Tetraoninae	Lagopus	mutus	70	...
	Mammalia	Chiroptera	Pteropodidae	Macroglossinae	Syconycteris	australis	1	...
		Perissodactyla	Equidae	N/A	Equus	caballus	571	...
		Primates	Hominidae	Homininae	Homo	sapiens	92	118
		Rodentia	Cricetidae	Arvicolinae	Ellobius	talpinus	51	
					Microtus	arvalis	70	...
						kikuchii	1	...
						lusitanicus	3	...
						savil	1	...
		Ruminantia	Bovidae	Bovinae	Bison	bison	4166	...
					Bos	frontalis	10	...
						grunniens	148	...
						primigenius	4	...
						taurus	25	...
				Caprinae	Capra	hircus	10	...
					Ovis	aries	66	...
				Cephalophinae	Cephalophus	zebra	2	...
			Moschidae	N/A	Moschus	moschiferus	1	...
						No Hit	6	1
						Total	5296	159

1) contaminant sequences from extraction blank AC-B1 (Table S4)

6.d) Plant *rbc*L and *trn*L Sequences

Figs. S7 and S8 show the distribution of plant sequences based on the *rbc*L and *trn*L assays, respectively, and complete listings of the top BLAST hits are provided in Tables S6 and S7. Although hundreds of genera were putatively identified, DNA damage and PCR/sequencing errors likely inflate the diversity. Additionally, identification beyond the family level is generally not supported with these loci.

Nevertheless, these assays reveal many plant families that are well-documented in Beringia, based on macrofossil and pollen data; the grasses *Poa* (bluegrass) and *Festuca* (fescue), sage (*Artemisia*), and willow (*Salix*) are particularly common in paleoecological records (*28, 31*). Although the two assays clearly differ in their specificity, both show a predominance of Asteraceae and Poaceae (blue and purple sections in Figs. S7 and S8), which are consistent with regional paleoecological data (*28, 31*).

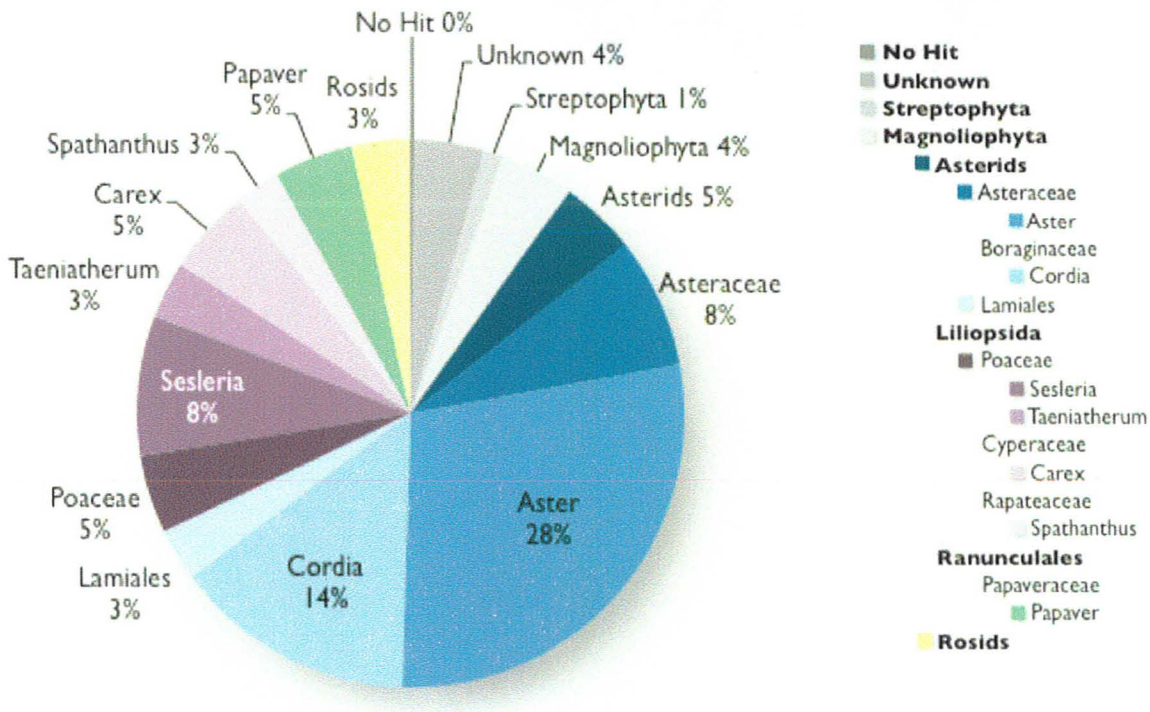


Fig. S7. Top BLAST hits for *rbc*L sequences

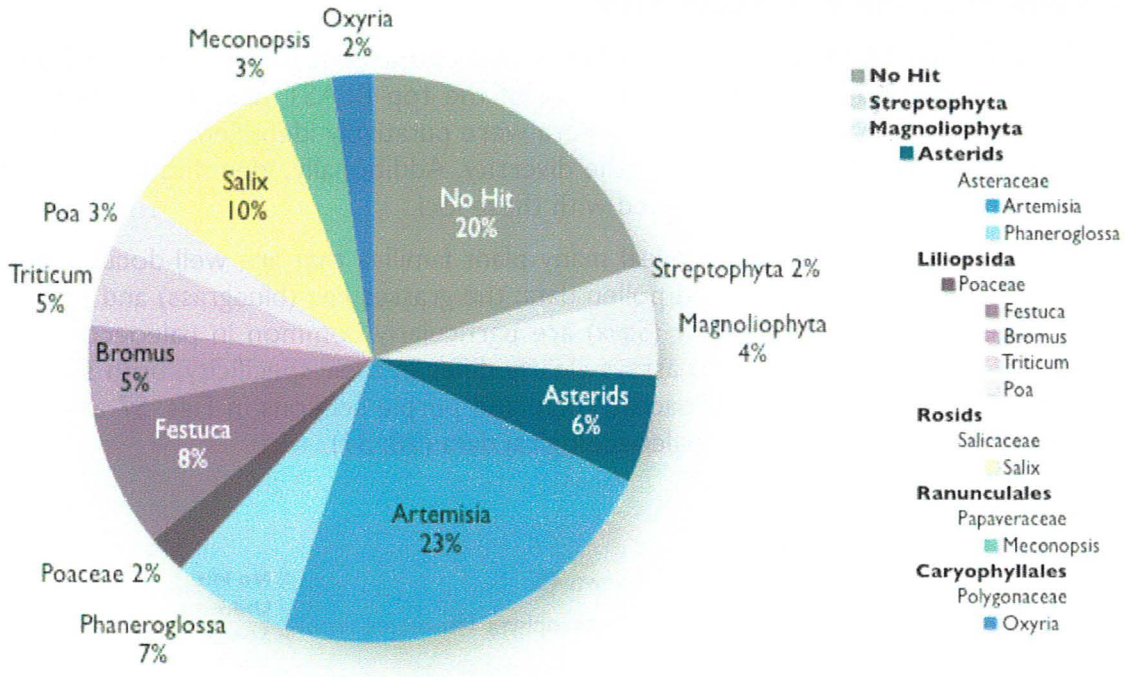


Fig. S8. Top BLAST hits for *trn*L sequences

Table S6. Top BLAST hits for *rbc*L sequences

Phylum	No Rank	(Sub)Class	Order	Family	Subfamily	Genus	#
Streptophyta	Coniferophyta	Coniferopsida	Coniferales	Pinaceae	N/A	Pinus	10
	Magnoliophyta	Asterids	Apiales	Apiaceae	Apioideae	Angelica	24
						Berula	4
						Bupleurum	6
						Saposhnikovia	1
				Araliaceae	N/A	Hedera	4
						Hydrocotyle	5
						Panax	3
						Pseudopanax	1
						Tetrapanax	1
				Pittosporaceae	N/A	Pittosporum	1
			Aquifoliales	Aquifoliaceae	N/A	Ilex	1
				Phyllonomaceae	N/A	Phyllonoma	1
			Asterales	Asteraceae	Asteroideae	Artemisia	2
						Aster	1282
						Calendula	1
						Cosmos	1
						Duhaldea	10
						Emilia	7
						Euryops	1
						Gaillardia	7
						Helenium	1
						Helianthus	1
						Pteronia	5
						Solidago	2
						Symphyotrichum	10
						Tagetes	16
						Ursinia	2
					Carduoideae	Carthamus	18
						Centaurea	8
						Dicoma	8
					Cichorioideae	Centratherum	2
						Crepis	2
						Elephantopus	2
						Gazania	92
						Lactuca	12
						Leontodon	17
						Prenanthes	2
					Mutisioideae	Gerbera	98
						Mutisia	6
						Perezia	4
						Trichocline	9
				Campanulaceae	N/A	Codonopsis	1
						Cyphia	1
						Lobelia	1
				Menyanthaceae	N/A	Nymphoides	1
				Stylidiaceae	N/A	Phyllachne	1

Table S6 (continued). Top BLAST hits for *rbc*L sequences

Phylum	No Rank	(Sub)Class	Order	Family	Subfamily	Genus	#
Streptophyta (continued)	Magnoliophyta (continued)	Asterids (continued)	Bruniales	Columelliaceae	N/A	Columellia	2
			Cornales	Cornaceae	N/A	Alangium	1
				Hydrostachyaceae	N/A	Hydrostachys	3
			Dipsacales	Adoxaceae	N/A	Sambucus	2
			Ericales	Polemoniaceae	N/A	Cantua	4
						Polemonium	2
				Primulaceae	N/A	Androsace	8
						Coris	1
						Lysimachia	1
				Styracaceae	N/A	Styrax	1
			Garryales	Garryaceae	N/A	Aucuba	8
			Gentianales	Apocynaceae	Asclepiadoideae	Leptadenia	1
						Schizostephanus	2
				Loganiaceae	N/A	Strychnos	15
				Rubiaceae	Cinchonoideae	Coutarea	1
						Hymenodictyon	4
						Suberanthus	2
					Ixoroideae	Oxyanthus	1
			Lamiales	Acanthaceae	Thunbergioideae	Thunbergia	1
				Bignoniaceae	N/A	Campsis	3
						Macfadyena	1
						Tecomaria	2
				Lamiaceae	Lamioideae	Lamium	14
					Viticoideae	Vitex	1
				Lentibulariaceae	N/A	Pinguicula	13
						Utricularia	1
				Oleaceae	N/A	Ligustrum	4
						Nyctanthes	4
				Orobanchaceae	N/A	Agalinis	3
						Castilleja	1
						Hyobanche	1
				Plantaginaceae	N/A	Aragoa	1
						Callitriche	3
						Linaria	41
						Plantago	18
				Scrophulariaceae	N/A	Lyperia	3
						Sutera	5
				Verbenaceae	N/A	Petrea	18
						Stachytarpheta	1

Table S6 (continued). Top BLAST hits for *rbc*L sequences

Phylum	No Rank	(Sub)Class	Order	Family	Subfamily	Genus	#
Streptophyta (continued)	Magnoliophyta (continued)	Asterids (continued)	Solanales	Convolvulaceae	N/A	Porana	1
				Hydrophyllaceae	N/A	Phacelia	1
			N/A	Boraginaceae	N/A	Anchusa	1
						Bourreria	1
						Cordia	644
						Ehretia	10
						Myosotis	72
				Icacinaceae	N/A	Chlamydocarya	1
		Liliopsida	Alismatales	Araceae	Aroideae	Arisaema	6
						Pistia	1
					Lemnoideae	Spirodela	1
				Tofieldiaceae	N/A	Tofieldia	1
			Asparagales	Asparagaceae	Scilloideae	Ornithogalum	1
				Hypoxidaceae	N/A	Curculigo	1
				Iridaceae	N/A	Trimezia	1
				Xanthorrhoeaceae	Hemerocallidoideae	Eccremis	2
			Dioscoreales	Dioscoreaceae	N/A	Tacca	1
			Pandanales	Stemonaceae	N/A	Stemona	1
			Poales	Cyperaceae	N/A	Carex	231
						Cladium	4
						Cyperus	6
						Neesenbeckia	1
						Schoenoplectus	12
						Scleria	2
				Poaceae	Bambusoideae	Phyllostachys	1
					Danthonioideae	Chaetobromus	2
					Ehrhartoideae	Oryza	2
						Phaenosperma	1
					Pooideae	Arrhenatherum	36
						Bromus	2
						Elymus	2
						Festuca	5
						Gaudinia	88
						Parapholis	1
						Phalaris	1
						Poa	64
						Polypogon	2
						Sesleria	372
						Taeniatherum	152
						Vulpia	8
				Rapateaceae	N/A	Cephalostemon	2
						Rapatea	1
						Spathanthus	117
				Typhaceae	N/A	Typha	1

Table S6 (continued). Top BLAST hits for *rbc*L sequences

Phylum	No Rank	(Sub)Class	Order	Family	Subfamily	Genus	#
Streptophyta (continued)	Magnoliophyta (continued)	Rosids	Brassicales	Brassicaceae	N/A	Arabidopsis	16
						Raphanus	33
						Vella	8
			Crossosomatales	Staphyleaceae	N/A	Turpinia	2
			Fabales	Fabaceae	Caesalpinioideae	Bauhinia	1
					Papilionoideae	Astragalus	8
						Cicer	3
						Echinosophora	1
						Lessertia	13
						Lupinus	12
				Polygalaceae	N/A	Polygala	1
			Fagales	Fagaceae	N/A	Castanea	2
						Lithocarpus	2
						Quercus	7
			Malpighiales	Achariaceae	N/A	Kiggelaria	2
				Euphorbiaceae	Acalyphoideae	Ricinus	5
					Crotonoideae	Manihot	1
				Salicaceae	N/A	Bembicia	4
					N/A	Salix	4
			Malvales	Thymelaeaceae	N/A	Passerina	1
			Myrtales	Melastomataceae	N/A	Melastoma	7
				Myrtaceae	N/A	Metrosideros	2
			Rosales	Moraceae	N/A	Antiaris	2
				Rosaceae	Rosoideae	Potentilla	11
						Sanguisorba	1
						Waldsteinia	1
				Ulmaceae	N/A	Ampelocera	1
			Zygophyllales	Zygophyllaceae	Morkillioideae	Morkillia	2
		N/A	Caryophyllales	Aizoaceae	N/A	Carpanthea	1
				Caryophyllaceae	N/A	Cerastium	5
						Drymaria	2
						Silene	3
						Stellaria	11
				Nepenthaceae	N/A	Nepenthes	2
				Phytolaccaceae	N/A	Phytolacca	1
				Polygonaceae	N/A	Bistorta	2
						Rheum	12
						Rumex	3
				Simmondsiaceae	N/A	Simmondsia	1

Table S6 (continued). Top BLAST hits for *rbc*L sequences

Phylum	No Rank	(Sub)Class	Order	Family	Subfamily	Genus	#
Streptophyta (continued)	Magnoliophyta (continued)	N/A (continued)	Dilleniales	Dilleniaceae	N/A	Dillenia	1
						Hibbertia	6
						Pinzona	4
						Schumacheria	5
			Laurales	Lauraceae	N/A	Beilschmiedia	1
			Ranunculales	Papaveraceae	N/A	Corydalis :	1
						Papaver	213
				Ranunculaceae	N/A	Aconitum	1
						Anemone	20
						Asteropyrum	1
						Caltha	1
						Clematis	5
						Coptis	3
						Delphinium	1
						Ranunculus	7
						Thalictrum	9
			Santalales	Olacaceae	N/A	Olacaceae	2
				Opiliaceae	N/A	Lepionurus	16
				Santalaceae	N/A	Pyrularia	1
			Saxifragales	Saxifragaceae	N/A	Heuchera	3
						Schefflera	4
			N/A	Sabiaceae	N/A	Meliosma	9
	Environmental embryophyte						42
						Unknown	187
						No Hit	13
						Total	4509

Table S7. Top BLAST hits for *trn*L sequences.

Phylum	No Rank	(Sub)Class	Order	Family	Subfamily	Genus	#
Streptophyta	Coniferophyta	Coniferopsida	Coniferales	Cupressaceae	N/A	Juniperus	16
				Pinaceae	N/A	Pinus	151
	Magnoliophyta	Asterids	Apiales	Apiaceae	Apioideae	Aletes	1
						Bupleurum	20
						Cnidium	62
						Lomatium	1
						Phlojodicarpus	2
						Pleurospermum	4
						Podistera	74
						Tauschia	16
			Aquifoliales	Cardiopteridaceae	N/A	Cardiopteris	1
			Asterales	Asteraceae	Asteroideae	Achillea	23
						Arnica	3
						Artemisia	2784
						Blumea	31
						Carphephorus	2
						Cephalosorus	1
						Coreocarpus	1
						Emilia	2
						Erigeron	71
						Helichrysum	15
						Jacobaea	1
						Layia	1
						Myriocephalus	1
						Pallenis	1
						Perityle	26
						Phaneroglossa	828
						Psacalium	1
						Pterygopappus	2
						Relhania	1
						Santolina	4
						Solidago	1
						Synotis	2
						Tagetes	5
						Tanacetum	1
						Tephroseris	2
						Tripleurospermum	2
						Wamalchitamia	1
					Carduoideae	Saussurea	4

Table S7 (continued). Top BLAST hits for *trn*L sequences.

Phylum	No Rank	(Sub)Class	Order	Family	Subfamily	Genus	#
Streptophyta (continued)	Magnoliophyta (continued)	Asterids (continued)	Asterales (continued)	Asteraceae (continued)	Cichorioideae	Arctotheca	1
						Arctotis	38
						Bothriocline	11
						Hirpicium	1
						Lactuca	17
						Leontodon	1
						Lessingianthus	1
						Lychnophora	1
						Tragopogon	5
						Vernonia	1
					Mutisioideae	Nassauvia	3
						Pachylaena	1
					Wunderlichioideae	Wunderlichia	1
				Campanulaceae	N/A	Campanula	2
						Phyteuma	1
			Cornales	Cornaceae	N/A	Nyssa	1
				Loasaceae	N/A	Huidobria	30
						Kissenia	1
						Klaprothia	4
						Mentzelia	1
						Presliophytum	1
			Dipsacales	Adoxaceae	N/A	Sinadoxa	1
				Caprifoliaceae	N/A	Valeriana	3
			Ericales	Diapensiaceae	N/A	Diapensia	1
				Lecythidaceae	N/A	Eschweilera	1
				Polemoniaceae	N/A	Phlox	18
						Polemonium	1
				Primulaceae	N/A	Androsace	1
				Styracaceae	N/A	Halesia	1
			Lamiales	Orobanchaceae	N/A	Pedicularis	15
				Plantaginaceae	N/A	Penstemon	1
						Plantago	38
			Solanales	Convolvulaceae	N/A	Jacquemontia	1
				Solanaceae	Petunioideae	Petunia	1
			N/A	Boraginaceae	N/A	Cordia	6
						Echiochilon	1
						Eritrichium	9
						Lithospermum	3
						Mertensia	107
						Myosotis	36
						Onosma	1
				Icacinaceae	N/A	Poraqueiba	2

Table S7 (continued). Top BLAST hits for *trn*L sequences.

Phylum	No Rank	(Sub)Class	Order	Family	Subfamily	Genus	#
Streptophyta (continued)	Magnoliophyta (continued)	Liliopsida	Asparagales	Asparagaceae	Scilloideae	Neopatersonia	1
			Poales	Poaceae	Ehrhartoideae	Prosphytochloa	1
					Pooideae	Aegilops	2
						Alopecurus	2
						Avena	1
						Avenula	3
						Briza	9
						Bromus	603
						Calamagrostis	72
						Cinna	1
						Dactylis	1
						Dupontia	2
						Echinopogon	2
						Festuca	968
						Hordeum	18
						Koeleria	2
						Phippsia	6
						Poa	369
						Puccinellia	65
						Trisetum	93
						Triticum	567
		Rosids	Brassicales	Brassicaceae	N/A	Alyssopsis	2
						Alyssum	18
						Arabis	1
						Braya	34
						Chaunanthus	1
						Crucihimalaya	1
						Descurainia	27
						Draba	78
						Parrya	1
						Physaria	8
						Smelowskia	15
			Fabales	Fabaceae	Papilionoideae	Cytisophyllum	1
						Laburnum	1
						Lupinus	9
						Oxytropis	3
			Fagales	Fagaceae	N/A	Fagus	2
			Malpighiales	Euphorbiaceae	Euphorbioideae	Homalanthus	1
						Pseudosenefeldera	1
				Humiriaceae	N/A	Humiria	2
				Salicaceae	N/A	Salix	1178
						Scolopia	1

Table S7 (continued). Top BLAST hits for *trn*L sequences.

Phylum	No Rank	(Sub)Class	Order	Family	Subfamily	Genus	#
Streptophyta (continued)	Magnoliophyta (continued)	Rosids (continued)	Rosales	Rosaceae	Dryadoideae	Dryas	30
					Rosoideae	Fragaria	1
						Horkeliella	2
						Potentilla	23
						Sanguisorba	3
					Spiraeoideae	Torminalis	2
						Vauquelinia	1
		N/A	Caryophyllales	Caryophyllaceae	N/A	Stellaria	29
				Polygonaceae	N/A	Bistorta	95
						Calligonum	4
						Eriogonum	4
						Fallopia	10
						Oxyria	291
						Rheum	4
						Rumex	2
			Gunnerales	Myrothamnaceae	N/A	Myrothamnus	1
			Ranunculales	Papaveraceae	N/A	Meconopsis	414
						Papaver	77
				Ranunculaceae	N/A	Anemone	46
						Ranunculus	6
						Thalictrum	5
			Saxifragales	Crassulaceae	N/A	Rhodiola	2
				Saxifragaceae	N/A	Saxifraga	10
	Environmental embryophyte						33
						No Hit	2402
						Total	12206

6.e) Bacterial 16S Sequences

The distribution of 16S sequences from the cores and from the contaminant products (Table S4, red) are shown in Fig. S9 and complete listings of the top BLAST hits are provided in Table S8. The genus *Arthrobacter* was among the most commonly encountered, consistent with previous findings (*32*). Additionally, the samples harboured an abundance of *Aeromicrobium* as well as numerous strains of Micromonosporaceae and *Streptomyces*, all commonly isolated from terrestrial environments (*33-37*).

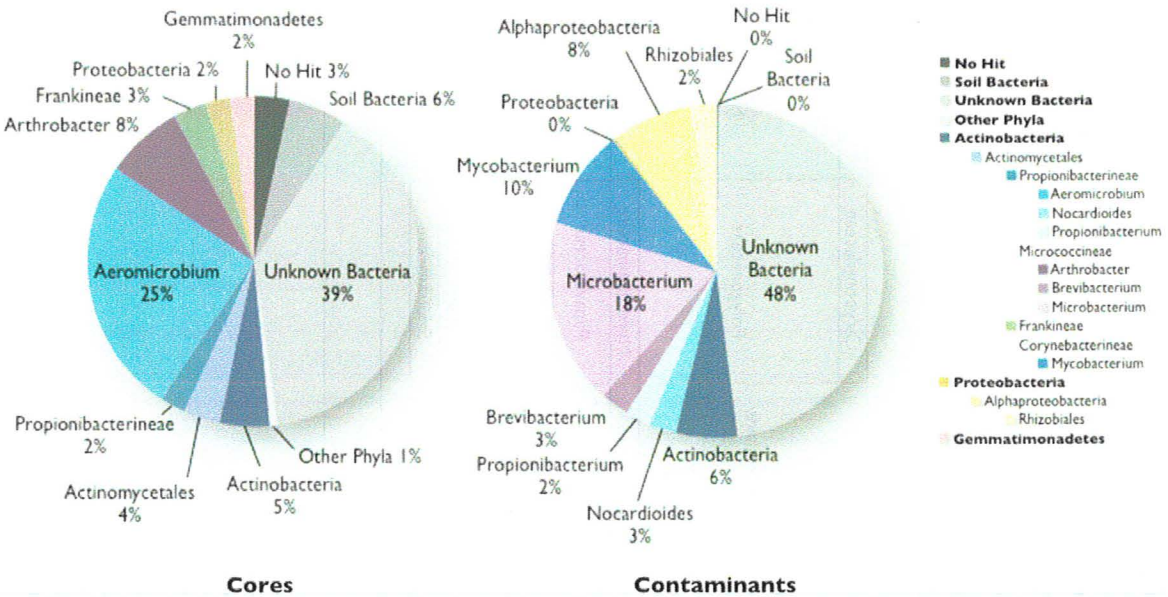


Fig. S9. Top BLAST hits for 16S sequences. Bacterial sequences in the cores (left) are distinct from those in the contaminant PCR products (right).

Table S8. Top BLAST hits for 16S sequences

Phylum	Class	Order	Suborder	Family	Genus	# Core	# Cont[1]
Acidobacteria	Acidobacteria	Acidobacteriales	N/A	Acidobacteriaceae	Acidobacterium	1	...
					Other/Unknown	7	...
		Other/Unknown	...	...	...	24	...
Actinobacteria	Actinobacteria	Acidimicrobiales	Acidimicrobineae	Acidimicrobiaceae	Other/Unknown	4	...
			Other/Unknown	...	...	24	...
		Actinomycetales	Corynebacterineae	Dietziaceae	Dietzia	3	...
				Mycobacteriaceae	Mycobacterium	148	129
				Nocardiaceae	Nocardia	1	...
					Rhodococcus	3	...
			Frankineae	Geodermatophilaceae	Blastococcus	29	...
					Modestobacter	57	...
					Other/Unknown	3	...
				Nakamurellaceae	Humicoccus	1	...
				Sporichthyaceae	Sporichthya	10	...
					Other/Unknown	66	...
				Other/Unknown	...	513	...
			Kineosporiineae	Kineosporiaceae	Other/Unknown	3	...
			Micrococcineae	Bogoriellaceae	Georgenia	3	...
				Brevibacteriaceae	Brevibacterium	...	37
				Cellulomonadaceae	Cellulomonas	6	...
					Other/Unknown	9	...
				Intrasporangiaceae	Arsenicicoccus	2	...
					Janibacter	...	1
					Knoellia	3	...
					Phycicoccus	2	...
					Tetrasphaera	1	...
					Other/Unknown	2	...
				Microbacteriaceae	Agreia	1	...
					Agrococcus	6	...
					Agromyces	26	...
					Clavibacter	1	...
					Cryobacterium	1	...
					Frigoribacterium	35	...
					Leifsonia	6	...
					Microbacterium	20	240
					Subtercola	1	...
					Other/Unknown	1	...
				Micrococcaceae	Arthrobacter	1593	...
					Kocuria	1	...
					Micrococcus	4	...
					Nesterenkonia	1	...
				Promicromonosporaceae	Cellulosimicrobium	13	...
					Promicromonospora	8	...
				Other/Unknown	...	86	...

Table S8 (continued). Top BLAST hits for 16S sequences

Phylum	Class	Order	Suborder	Family	Genus	# Core	# Cont[1]
Actinobacteria (continued)	Actinobacteria (continued)	Actinomycetales (continued)	Micromonosporineae	Micromonosporaceae	Actinoplanes	10	...
					Catellatospora	12	...
					Couchioplanes	2	...
					Dactylosporangium	6	...
					Micromonospora	60	...
					Polymorphospora	7	...
					Solwaraspora	1	...
					Verrucosispora	54	...
					Virgisporangium	13	...
					Other/Unknown	17	...
			Propionibacterineae	Nocardioidaceae	Aeromicrobium	5259	...
					Friedmanniella	3	...
					Kribbella	22	...
					Nocardioides	49	35
					Other/Unknown	12	...
				Propionibacteriaceae	Microlunatus	29	...
					Propionibacterium	...	32
					Other/Unknown	...	1
				Other/Unknown		402	...
			Pseudonocardineae	Actinosynnemataceae	Other/Unknown	8	...
				Pseudonocardiaceae	Amycolatopsis	4	...
					Crossiella	1	...
					Pseudonocardia	37	...
					Other/Unknown	9	...
			Streptomycineae	Streptomycetaceae	Streptomyces	54	...
					Other/Unknown	1	...
				Other/Unknown		3	...
			Streptosporangineae	Nocardiopsaceae	Other/Unknown	1	...
				Streptosporangiaceae	Streptosporangium	2	...
				Thermomonosporaceae	Actinomadura	2	...
					Other/Unknown	5	...
			Other/Unknown			46	2
		Bifidobacteriales	N/A	Bifidobacteriaceae	Bifidobacterium	...	1
		Coriobacteriales	Coriobacterineae	Coriobacteriaceae	Other/Unknown	1	...
		Rubrobacterales	Rubrobacterineae	Rubrobacteraceae	Other/Unknown	1	...
				Other/Unknown		7	...
		Other/Unknown				971	74
Chloroflexi	Other/Unknown					2	...
Deinococcus-Thermus	Deinococci	Deinococcales	N/A	Deinococcaceae	Other/Unknown	1	...
Firmicutes	Clostridia	Clostridiales	N/A	Clostridiaceae	Clostridium	26	...
					Other/Unknown	43	...
				Eubacteriaceae	Eubacterium	21	...
	Other/Unknown					40	...

Table S8 (continued). Top BLAST hits for 16S sequences

Phylum	Class	Order	Suborder	Family	Genus	# Core	# Cont[1]
Gemmatimonadetes	Gemmatimonadetes	Other/Unknown				487	...
Proteobacteria	Alphaproteobacteria	Caulobacterales	N/A	Caulobacteraceae	Brevundimonas	...	1
					Caulobacter	...	22
		Parvularculales	N/A	Parvularculaceae	Parvularcula	3	...
		Rhizobiales	N/A	Aurantimonadaceae	Fulvimarina	2	...
				Bradyrhizobiaceae	Afipia	1	...
					Bosea	4	15
					Bradyrhizobium	25	2
					Nitrobacter	1	...
					Other/Unknown	2	...
				Brucellaceae	Other/Unknown	2	...
				Hyphomicrobiaceae	Devosia	...	8
					Hyphomicrobiaceae	21	...
					Hyphomicrobium	3	...
					Pedomicrobium	120	...
					Rhodomicrobium	12	...
					Rhodoplanes	153	...
				Methylobacteriaceae	Meganema	3	...
					Methylobacterium	2	7
					Microvirga	10	...
				Methylocystaceae	Methylocystis	2	...
					Other/Unknown	2	...
				Phyllobacteriaceae	Mesorhizobium	1	...
				Rhizobiaceae	Rhizobium	...	1
		Rhodospirillales	Other/Unknown			5	...
		Sphingomonadales	N/A	Erythrobacteraceae	Porphyrobacter	5	...
				Sphingomonadaceae	Novosphingobium	...	7
					Sphingomonas	14	7
					Other/Unknown	5	1
		Other/Unknown				7	66
	Betaproteobacteria	Burkholderiales	N/A	Alcaligenaceae	Alcaligenes	1	...
				Comamonadaceae	Comamonas	...	3
		Rhodocyclales	N/A	Rhodocyclaceae	Other/Unknown	2	...
	Deltaproteobacteria	Desulfuromonadales	N/A	Geobacteraceae	Geobacter	8	...
		Myxococcales	Sorangiineae	Polyangiaceae	Other/Unknown	2	...
		Other/Unknown				25	...
	Gammaproteobacteria	Enterobacteriales	N/A	Enterobacteriaceae	Erwinia	1	...
		Legionellales	N/A	Legionellaceae	Legionella	3	...
		Other/Unknown				23	1
	Other/Unknown					27	...
					Soil bacteria	1183	2
					Unknown	8068	634
					No Hit	728	2
					Total	20936	1331

1) Contaminant sequences from extraction blanks BC-B1, BC-B2, and PCR NTCs highlighted in Table S4.

6.f) *tetM* Sequence Analysis

BLAST results for the putative *tetM* sequences indicated that most were GTP-binding proteins and others elongation factors, which have similar ribosome binding sites (*38*); these were excluded from further analysis. The remaining 794 unique *tetM* sequences were translated and a subset of these subject to phylogenetic analysis using MrBayes (*39*) (with a mixed amino acid model, gamma-distributed rates, 4 chains, 2 million generations, sampling every 1000, 1000-sample burn-in, and manually checked for convergence) (Figs. 2C-B, main text).

6.g) *vanX* Sequence Analysis

The 2911 unique *vanX* sequences were translated and a subset of these subject to phylogenetic analysis using MrBayes, as described above; the resulting tree is shown in Fig. S10.

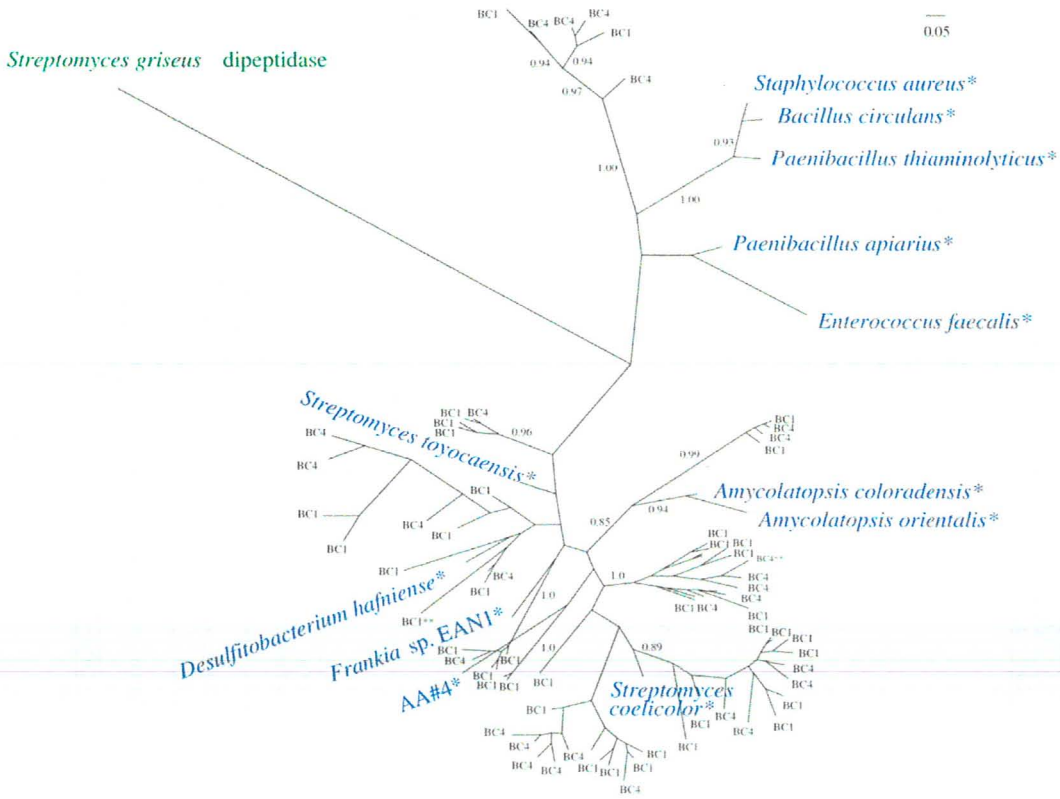


Fig. S10. Genetic diversity of permafrost VanX sequences. Unrooted phylogeny of translated *vanX* sequences. Blue denotes enzymes that are predicted to belong to the VanX family and green denotes those putatively associated with other functions. Permafrost-derived sequences are labeled by the originating core name. Strains that are associated with glycopeptide resistance are labeled with a single asterisk (*) (*20, 33, 51-56*) and those containing putative stop codons but homology throughout with a double asterisk (**). The scale bar represents 0.05 substitutions per site. Posterior support values over 0.7 are shown.

7. *vanHAX* CLUSTER

7.a) Amplification and Sequencing

The *vanHAX* assays (Fig. 3A, main text) were optimized ($MgCl_2$, DMSO, annealing temperature, etc.) using qPCR and their sensitivities evaluated as described above; the details of each assay are listed in Table S3. Successful amplifications were replicated three times from multiple extracts stemming from two subsamples (preferably E and D); Table S9 shows the amplification success and qPCR yields, where available (i.e. excluding reactions producing any secondary products).

In preparation for cloning, amplifications producing secondary products were cored and reamplified from the agarose plugs; all other products were diluted 1/1000 with 0.1x TE and reamplified for 20 cycles following the original PCR conditions. All products were cloned using the TOPO-TA cloning kit with pCR 2.1-TOPO vector and One Shot TOP10 Chemically Competent *E. coli* (Invitrogen, Carlsbad, CA, USA); insert sizes were confirmed by PCR screening with the M13 primers. PCR products of the correct size were purified using AcroPrep Filter Plates (30K, Omega membrane; Pall, Port Washington, NY, USA) and quantified following gel electrophoresis by visual comparison to a standardized DNA ladder. Approximately 1 ng per 100 bp of template DNA was used for sequencing with 0.3 μL of BigDye v3.1 (Applied Biosystems, Foster City, CA, USA) in 7 μL reactions. After cycling, the reactions were purified and sequenced by Sanger methods at the MobixLab (McMaster University). At least three clones were sequenced from each PCR replicate listed in Table S9. Most products were sequenced in both directions, allowing all full-length consensus sequences to be built only from high quality bases (QV≥20).

Table S9. qPCR results – *vanHAX* cluster

Extract	HA			AX		HAX	
	H1A1	H2A3	H2A4	A6X	A2X	H1AX	H2AX
Samples							
BC1-D	**3/3**	**4/10**	**1/13**	**3/3**	**5/9**	0/1	0/1
	31	*n.d.*	<1	1873	*n.d.*	...	...
BC1-E	**3/3**	...	**3/4**	**3/3**	**3/3**	0/1	0/1
	334	...	28	11232	*n.d.*	...	...
BC1-E kit	**3/3**	**3/3**	**1/7**	**3/3**	**3/3**	**1/3**	**2/3**
	11	18	7	387	*1*	*n.d*	23
BC1-E per	**3/3**	...	**3/3**	**3/3**	**3/3**	0/7	**1/3**
	155	...	1	14567	*n.d.*	...	*n.d.*
BC4-C	...	...	**4/12**	...	**4/5**	0/1	0/1
	...	...	2	...	*2*	...	...
BC4-D	**3/3**	...	**5/8**	**3/3**	**3/3**	0/1	0/1
	48	...	7	3218	*2*	...	...
BC4-E[1]	**3/3**	0/1	0/4	**3/3**	0/4	0/1	0/1
	14	...	...	482	...	...	...
BC4-E kit	**3/3**	...	**3/3**	**3/3**	**3/3**	**2/3**	**3/3**
	11	...	14	2293	43	*n.d*	77
BC4-E per	**3/3**	...	**3/9**	**3/3**	**4/5**	0/4	0/1
	6	...	1	4603	*4*	...	...
Air Controls							
ACB	0/1	0/1	0/1	0/1	0/1	0/1	0/1
ACH	0/1	0/1	0/1	0/1	0/1	0/1	0/1
Extraction Blanks							
BC-B1	0/1	0/1	0/1	0/1	0/1	0/1	0/1
BC-B2	0/1	0/1	0/1	0/1	0/1	0/1	0/1
kit blk	0/1	0/1	0/1	0/1	0/1	0/1	0/1
per blk	0/1	0/1	0/1	0/1	0/1	0/1	0/1
AC-B1	0/1	0/1	0/1	0/1	0/1	0/1	0/1
AC-B2	0/1	0/1	0/1	0/1	0/1	0/1	0/1
PCR NTCs							
...	0/34	0/41	0/53	0/34	0/49	0/30	0/26

Fraction of positive PCRs, with successful amplifications in bold; shaded cells contain the average qPCR values, excluding any reactions with secondary products (italics); 1) dilutions of BC4-E used for most PCRs, due to limited extract volume; n.d., no data (i.e. all reactions had secondary products); NTCs, no-template controls.

7.b) *vanHAX* Sequence Analysis

For each of the 164 clones from the smallest *vanHA* amplicon (H1A1), the two open reading frames were translated and concatenated; a subset of these (42 sequences) were subject to phylogenetic analysis using MrBayes, as described above. The 159 clones from the smallest *vanAX* amplicon (A6X) were processed in the same manner. The resulting trees are shown in Figs. S11 and S12.

The complete *vanA* open reading frames from 40 clones of the 1.2 kb amplicons (H1AX and H2AX) were translated and selected translations subject to phylogenetic analysis using MrBayes, as described above. The resulting tree is shown in Fig. 3B (main text).

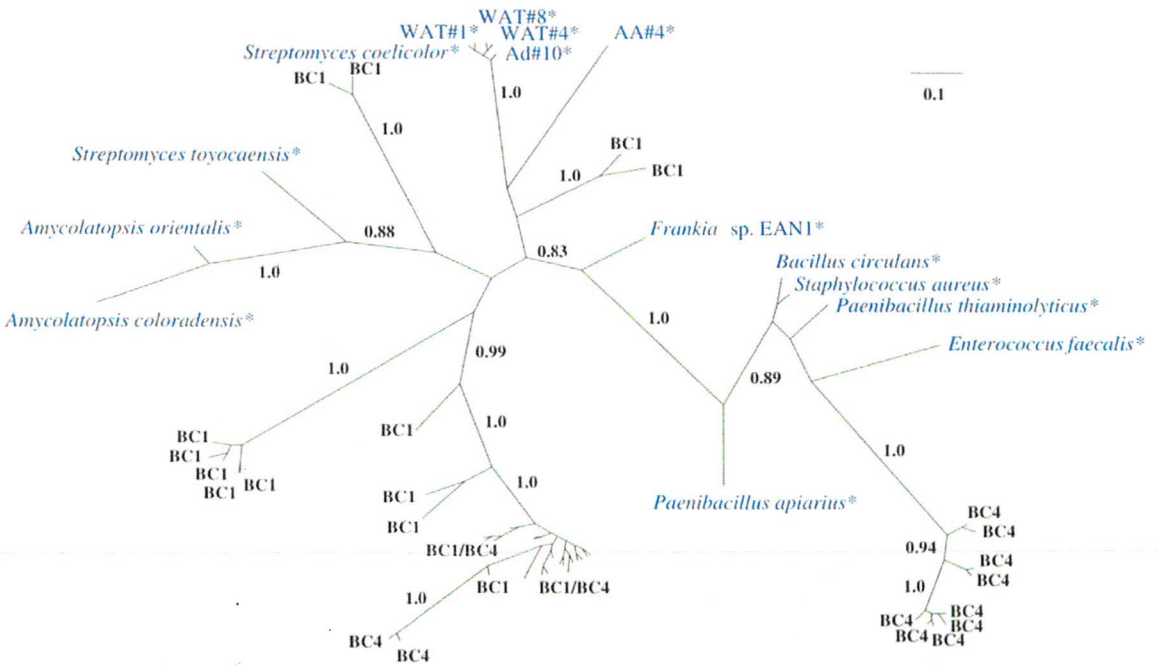


Fig. S11. Genetic diversity of permafrost VanHA sequences. Unrooted phylogeny of translated, concatenated sequences from the H1A1 screen. Permafrost-derived sequences are labeled by the originating core name. Blue denotes strains characterized as glycopeptide-resistant; strains that contain a *vanHAX* cluster are labeled with a single asterisk (*) (*20, 33, 51-56*). The scale bar represents 0.1 substitutions per site. Posterior support values over 0.7 are shown.

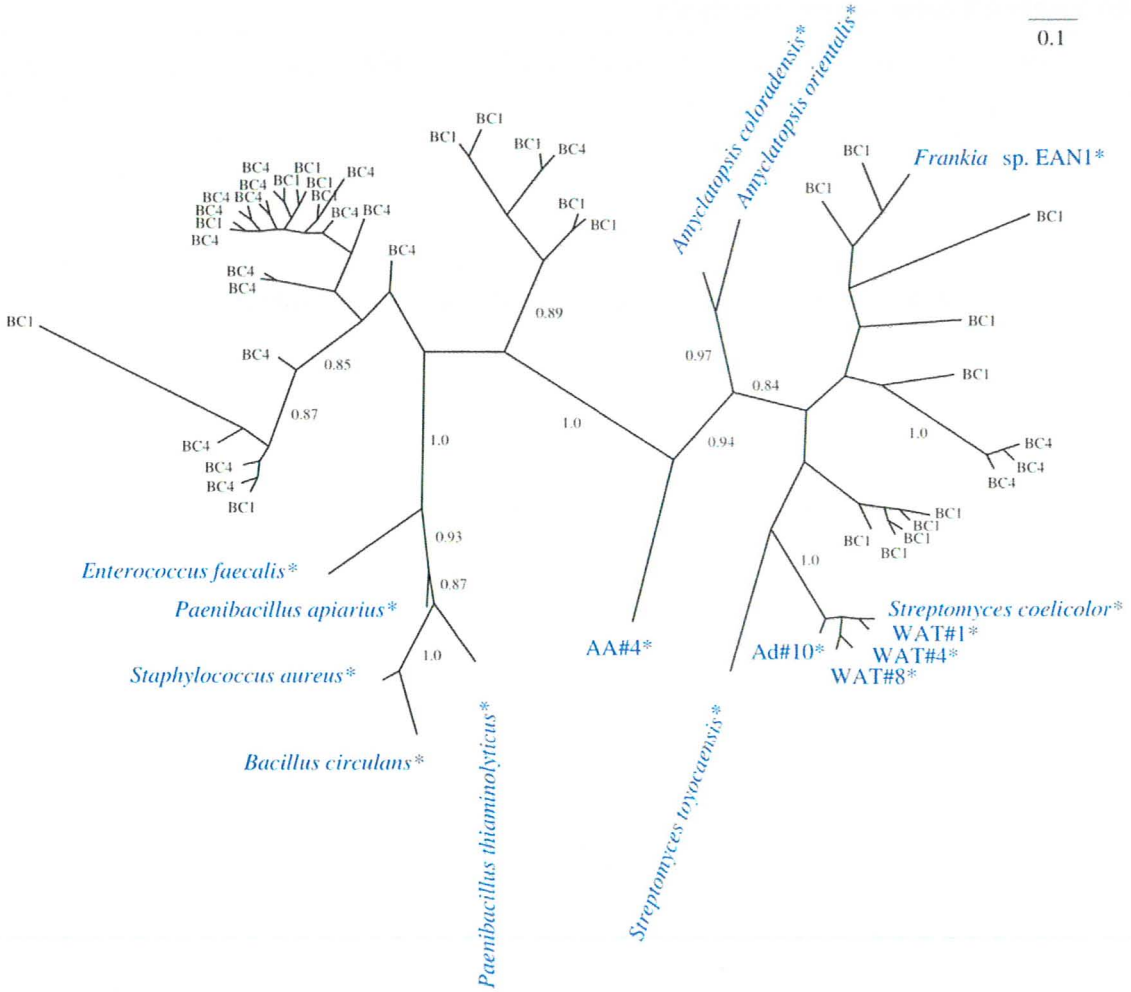


Fig. S12. Genetic diversity of permafrost VanAX sequences. Unrooted phylogeny of translated, concatenated sequences from the A6X screen. Permafrost-derived sequences are labeled by the originating core name. Blue denotes strains characterized as glycopeptide-resistant; strains that contain a *vanHAX* cluster are labeled with a single asterisk (*) (*20, 33, 51-56*). The scale bar represents 0.1 substitutions per site. Posterior support values over 0.7 are shown.

7.c) Replication

Replicates of the "kit" extractions for subsamples BC1-E and BC4-E were performed independently at the "Ancient DNA Box" (Service de Systématique Moléculaire) of the Muséum National d'Histoire Naturelle (Paris, France) as per the manufacturer's "alternative protocol for maximum yields" (MO BIO Laboratories, Carlsbad, CA, USA). These extracts were used for replicating amplifications of the *vanHA* (H1A1) and *vanAX* (A6X) regions using the same PCR conditions (Table S3) except for the Taq concentration (4-5 U instead of 2.5 U). Straight PCR products were blunted and cloned using the CloneJet PCR cloning kit (Fermentas, Estonia) and colony PCRs were performed using the pJET1.2 primers to check for presence of the insert. Six clones were sequenced for each of the four reactions (H1A1 or A6X with either extract), using the same procedure described above, but with the pJET1.2 primers. Twenty-three sequences with QV>20 were obtained and subject to permutation tests as described in the main text.

7.d) Biochemistry

Synthesis, expression, and purification of ancient D-ala-D-lac ligases

Four complete *vanA* (D-Ala-D-Lac ligase) sequences (the consensus sequences from four distinct sequence groups apparent in an initial subset of 20 clones) were selected from the H1AX/H2AX clones for further biochemical analysis. Each was synthesized with codon optimization for expression in *E. coli* and incorporated 5' *Nde* I and 3' *Hind* III restriction enzyme sites (GenScript, Piscataway, NJ, USA); the resulting genes were sub-cloned from pUC57 into the expression vector pET28a with an N-terminal His6 tag for downstream purification. Constructs were confirmed by sequencing and subsequently propagated in *E. coli* BL21 (DE3) (Novagen, Darmstadt, Germany) for expression. Two constructs produced enough soluble protein for downstream analysis, designated $vanA_{A2}$ and $vanA_{A4}$. High-level protein expression was achieved by growing cells in 1 L of Luria-Bertani broth to an optical density of 0.6 at 600 nm. Protein expression under the T7 promoter was induced by addition of isopropyl β-D-1-thiogalactopyranoside to a final concentration of 1 mM, followed by incubation at 16°C for 16-18 hours. Post-harvesting, cells were washed in 0.85% NaCl (w/v) and the pellet stored at -20°C or prepared for protein purification.

Nickel-NTA immobilized metal affinity chromatography (Qiagen, Hilden, Germany) was utilized to purify $VanA_{A2}$ and $VanA_{A4}$. Briefly, cell pellets were re-suspended in lysis buffer (50 mM HEPES; 500 mM NaCl; 20 mM Imidazole; 10 mM $MgCl_2$; 1 mM phenylmethanesulfonyl fluoride; 1 mM DNase; pH 7.5) and lysed by three passes on a continuous cell disruptor (Constant Systems Ltd., Daventry, UK) at 30 kPSI. Lysate was clarified by centrifugation at 27,000 x g for 30 min; the supernatant was collected and applied to a Nickel-NTA immobilized metal affinity column (Qiagen, Hilden, Germany). After elution with imidazole, fractions containing purified enzyme were pooled and dialyzed into 50 mM HEPES, 150 mM NaCl, pH 7.5 at 4°C. $VanA_{Ao}$, from the vancomycin producer *Amycolatopsis orientalis* C329.2, was purified as above from an *E. coli* BL21 (DE3) strain harboring pET28b*vanA*$_{Ao}$ previously described by Marshall et al. (*26*). *E. coli* W3110 harboring pTB2 for expression of the D-Ala-D-Ala ligase DdlB was previously reported (*40*).

Ddl assay

Substrate specificity of $VanA_{A2}$ and $VanA_{A4}$ was qualitatively determined using the pyruvate kinase/lactate dehydrogenase coupled assay to monitor ADP formation (*41*) (Table S10). Amino- and hydroxy acid substrate specificity was visually determined by thin-layer-chromatography (TLC) and using radiolabeled substrates (Fig. S13). Due to cost of [U-^{14}C]-D-Alanine, [U-^{14}C]-L-Alanine was isomerized to a racemic mixture of [^{14}C]-L/D-Alanine with one unit of *Bacillus stearothermophilus* alanine racemase (Sigma-Aldrich, St. Louis, MO, USA). Ligase reactions contained 50 mM HEPES pH 7.5; 10 mM $MgCl_2$; 40 mM KCl; 6 mM ATP; 1 μM enzyme; 0.1 μCi [U-^{14}C]-L/D-Ala; 1 mM D-Ala and 2 mM D-X substrate. Reactions were quenched with 50% methanol and applied onto a PEI-cellulose TLC plate (Sigma-Aldrich, St. Louis, MO, USA). The plates were developed in 12:3:5 butanol:acetic acid:water, dried and exposed to a phosphor-storage imaging screen. The screens were imaged using a Typhoon™ variable mode imager.

Michaelis-Menten kinetics were determined by using the software program GraFit version 4.0.21 (*42*) and initial rates were determined using the non-linear least squares method and equation 1.

$$v = (k_{cat}/E_t)[S]/(K_M + [S]) \qquad [1]$$

Table S10. Kinetic characterization of $VanA_{A2}$ and $VanA_{A4}$ ligases using the pyruvate kinase/lactate dehydrogenase assay

Enzyme	Product	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$)
$VanA_{A2}$	D-Ala-D-Ala	32 ± 3.5	0.82 ± 0.036	2.6×10^{1}
	D-Ala-D-Lac	0.22 ± 0.061	0.24 ± 0.023	1.1×10^{3}
	ATP	0.085 ±0.0092	0.18 ± 0.0054	2.2×10^{3}
$VanA_{A4}$	D-Ala-D-Ala	35 ± 3.2	0.29 ± 0.012	8.3×10^{1}
	D-Ala-D-Lac	0.55 ± 0.042	0.13 ± 0.0041	2.4×10^{2}
	ATP	0.043 ± 0.0034	0.1 ± 0.0035	2.5×10^{3}

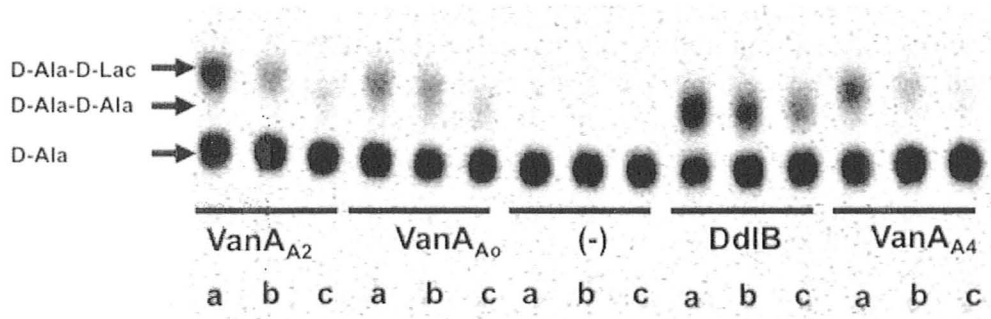


Fig. S13. Substrate specificity of D-Ala-D-X ligase $VanA_{A2}$ and $VanA_{A4}$. Substrate specificity was examined using [U-^{14}C]-D-Ala and either D-Ala or D-Lac. The products of each reaction were separated by TLC and exposed to a phospor-storage screen. All reactions contain 0.1 μCi [U-^{14}C]-D-Ala, and a) 2 mM D-Lac, b) 1 mM D-Ala, 2 mM D-Lac or c) 3 mM D-Ala. Based on migrations patterns, $VanA_{A2}$ and $VanA_{A4}$ exhibit activity like that of $VanA_{Ao}$, the D-Ala-D-Lac ligase from the vancomycin producer *Amycolatopsis orientalis*. DdlB, the D-Ala-D-Ala ligase from *E. coli* only produces D-Ala-D-Ala while the (-) reactions contain no enzyme and represent unreacted D-Alanine.

7.e) X-Ray Crystallography

Purification of $VanA_{A2}$ for crystallization

$VanA_{A2}$ was purified via Ni-NTA affinity chromatography following the standard protocol (described above). To remove the $VanA_{A2}$ histidine-tag, thrombin digestion was performed following protein purification. $VanA_{A2}$ was dialyzed into digestion buffer (20 mM Tris (pH 8.0), 150 mM NaCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, and 0.1% (v/v) β-mercaptoethanol) at 4°C for 2 h. Thrombin (from bovine plasma, Sigma-Aldrich, St. Louis, MO, USA) was then added directly into dialysis tubing at a ratio of 100:1 (mg/mL) $VanA_{A2}$ to thrombin, and the digestion was carried out for 48 h. To remove uncleaved protein, Ni-NTA was performed for the second time with the retention of flow-through rather than the elution fractions. The cleaved $VanA_{A2}$ was further purified via size exclusion chromatography using HiLoad 26/60 Superdex 200 column (GE Healthcare, Little Chalfont, UK) with a buffer containing 100 mM KCl and 20 mM HEPES (pH 7.5). The eluted protein was concentrated using an Amicon Ultra Centrifugal Filter (Millipore, Billerica, MA, USA) to 20 mg/mL as per the Bio-Rad Protein assay. The crystallization trials were performed within 48 h of protein purification.

$VanA_{A2}$ crystallization

$VanA_{A2}$ was crystallized using the vapor-diffusion hanging drop method. The crystallization condition, obtained from Nextal PEGs crystal screen (Qiagen, Hilden, Germany), consisted of 0.1M sodium acetate (pH 4.6) and 25% PEG3000. Prior to crystallization, $VanA_{A2}$ (at 7 mg/mL) was incubated with 2 mM ATP for 30 min, and then 1.5 μL of protein solution was combined with 1.5 μL of the crystallization solution; 1 M ammonium sulfate was used as well solution. For data collection the crystals were cryo-protected in the mother liquor with addition of 17% glycerol and frozen in liquid nitrogen.

X-ray data collection and processing

Data were collected at the National Synchrotron Light Source, Brookhaven National Laboratory (USA), beam-line X25. For data collection, 1 sec exposure and 1° oscillation were used, with crystal-to-detector distance set at 300 mm. The data were processed to 3 Å resolution using the HKL2000 program (*43*). $VanA_{A2}$ crystallized in the space group $P2_12_12_1$ with the unit cell parameters: a=84.13 Å, b=136.03 Å, c=178.09 Å, 6 molecules per asymmetric unit, and solvent content of 46.5%. Data collection statistics are listed in Table S11.

Structure determination and refinement

$VanA_{A2}$ structure (Figs. 3C-D) was determined by molecular replacement using Phaser from the CCP4i program suite (*44*). The dimer of *Enterococcus faecium* VanA (PDB ID: 1E4E) was used as the search model to locate three dimers of $VanA_{A2}$. The $VanA_{A2}$ model was refined via cycles of Refmac5 TLS/restrained refinement in CCP4i, and manual model building in COOT (*45*). Non-crystallographic symmetry (NCS) restraints were used in automated refinement, and NCS maps were used for manual model building. Due to structural disorder, residues 233-248 were not included in the final model. A molecule of ATP and two Mg^{2+} ions were modeled in the active site of $VanA_{A2}$ based on a clear electron density observed in the Fo-Fc difference map. Refinement statistics are listed in Table S11.

Table S11. Data processing and refinement statistics for $VanA_{A2}$/ATP binary complex

Data Collection	
Resolution (Å)	50-3.07
Space group	$P2_12_12_1$
Cell parameters (Å)	a=84.13 b=136.03 c=178.09
Unique reflections	38823
I/σ	14.5(3.6)
R_{sym}	13.4 (49.9)
Redundancy	6.2 (6.2)
Completeness (%)	99.8 (99.4)

Data Refinement	
R (%)	19.3
R_{free} (%)	26.3
Reflections in refinement	36901
Reflections in test set	1963
No of protein atoms	14659
No of water atoms	234
RMSD:	
bond (Å)	0.004
angle (°)	1.2
Mean B factor (Å^2)	20.2
Ramachandran plot:	
most favored (%)	94.8
additionally allowed (%)	5.2

Comparison of modern and ancient VanA structures

The tertiary and quaternary structure of VanA is very well conserved; most of the variation, as expected, is in the loop regions. The most dramatic shift of 13 Å is observed for the ordered part of the omega loop: residues 248-255 ($VanA_{A2}$ numbering): Asp251 (Glu 254 in modern VanA) swings into the active site with the carboxylate moiety of this residue occupying the location and mimicking interactions of the substrate carboxylate. In modern VanA structure, the inhibitor carboxylate forms hydrogen bonds with conserved Ser and Arg, while in $VanA_{A2}$ these are Ser313 and Arg314 forming the same hydrogen bonds with Asp251. Therefore, it appears that in the absence of the substrate the enzyme could maintain the positioning of key binding pocket functionalities through mimicry of the substrate functional group. The difference in the omega loop conformation also results in Ser250 forming interactions with the γ-phosphate of ATP unseen in the modern VanA structure. All the other interactions with the nucleotide are strictly conserved between the two structures.

8. OTHER RESISTANCE GENES

8.a) Amplification and Sequencing

We designed short qPCR assays for three additional antibiotic resistance determinants: *erm*, *bla*, and *aac(3)*. To generate qPCR standards, the fragments of interest (Table S3) were amplified from *Streptomyces coelicolor* (*erm*) or *Streptomyces fradiae* (*bla*, *aac(3)*) genomic DNA, cloned, and amplified from the constructs using standard M13 primers. Purification and quantification were carried out as described above. As before, the assays were optimized and the sensitivity evaluated.

Only the *bla* assay yielded amplification products of the correct size, which were isolated from an agarose gel and re-adenylated. Low concentration PCR products were also subjected to an additional round of amplification prior to insertion into the pCR®II-TOPO cloning vector (Invitrogen, Carlsbad, CA, USA). Ligations were transformed into *E. coli* TOP10 electrocompetent cells (Invitrogen, Carlsbad, CA, USA), and constructs of interest were sequenced at the MobixLab (McMaster University).

8.b) *bla* Sequence Analysis

From thirteen putative *bla* clones, we obtained 9 unique amino sequences, each with unique BLAST hits and E values <0.001. Because of the diversity of the sequences obtained, these were divided into two subsets of data and analyzed using MrBayes, as described above. The results are shown in the main text, Fig. 2A and Fig. S14.

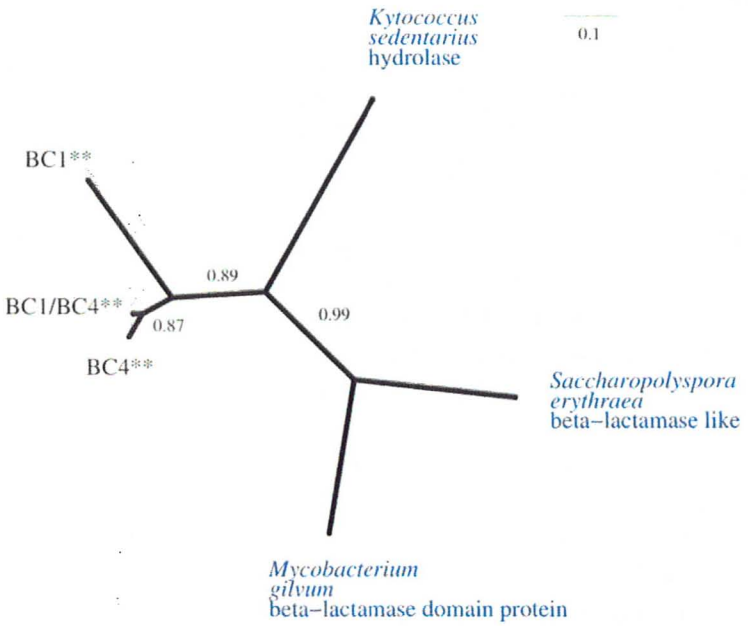


Fig. S14. Genetic diversity of permafrost Bla sequences (subset 2). Unrooted phylogeny of translated sequences. Blue denotes enzymes predicted to belong to the Bla family or having closely related functions (*54*). Permafrost-derived sequences are labeled by the originating core name and sequences containing putative stop codons but homology throughout with a double asterisk (**). The scale bar represents 0.1 substitutions per site. Posterior support values are indicated at the corresponding nodes.

9. REPORTS OF ANCIENT BACTERIA AND RESISTANCE IN PERMAFROST

In recent years, several studies claim to have isolated viable bacteria from ancient permafrost, ice, and other sources; the authenticity of these results remains contentious however, as they have not been independently replicated by other labs (sometimes even contradicting other studies), often lack both biological and experimental replicates, and are generally missing necessary controls to unambiguously confirm claims of ancient origin and lack of contemporary contamination (*18*). Such culture-dependent methods have been employed in reports of ancient bacteria resistant to antibiotics (*46-50*) and to mercury (*50-53*), which may be associated with the acquisition of antibiotic resistance. Although these studies attempt to monitor contamination introduced during coring, none offer sufficient downstream controls to monitor contamination during culturing, including air controls in the work areas and media blanks (*18*). Given the ubiquitous nature of microbial contaminants, it is essential that these controls be performed, and especially pertinent when working with environmental samples for which there is no simple distinction between endogenous and contaminant material. Without appropriate levels of replication and controls it is therefore impossible to rule out contamination in these reports.

10. REFERENCES

1. D. G. Froese, G. D. Zazula, A. V. Reyes, *Quat. Sci. Rev.* **25**, 1542 (2006).
2. D. G. Froese *et al.*, *GSA Today* **19**, 5 (2009).
3. E. Kotler, C. R. Burn, *Can. J. Earth Sci.* **37**, 849 (2000).
4. G. D. Zazula, D. G. Froese, S. A. Elias, S. Kuzmina, R. W. Mathewes, *Quat. Sci. Rev.* **26**, 979 (2007).
5. F. Brock, D. G. Froese, R. G. Roberts, *Quat. Geochronol.* **5**, 625 (2010).
6. M. Demuro *et al.*, *Quat. Geochronol.* **3**, 346 (2008).
7. J. R. Mackay, *Can. Geotech. J.* **11**, 230 (1974).
8. C. R. Burn, Ph.D. thesis, Carleton University (1986).
9. H. French, Y. Shur, *Earth Sci. Rev.* **101**, 190 (2010).
10. P. J. Williams, M. W. Smith, *The Frozen Earth: Fundamentals of Geocryology*. C. Archer *et al.*, Eds., Studies in Polar Research (Cambridge University Press, Cambridge, 1995), pp. 306.
11. J.-M. Konrad, N. R. Morgenstern, *Can. Geotech. J.* **17**, 473 (1980).
12. J.-M. Konrad, N. R. Morgenstern, *Can. Geotech. J.* **18**, 482 (1981).
13. R. D. Miller, in *Applications of Soil Physics,* D. Hillel, Ed. (Harcourt Brace Jovanovich, New York, 1980), pp. 254-299.
14. E. D. Yershov, in *General Geocryology*. (Cambridge University Press, Cambridge, 1998), pp. 346-372.
15. D. F. Juck *et al.*, *Appl. Environ. Microbiol.* **71**, 1035 (2005).
16. G. M. Khlebnikova, D. A. Gilichinsky, D. G. Fedorov-Davydov, E. A. Vorob'eva, *Microbiology* **59**, 106 (1990).
17. T. Shi, R. H. Reeves, D. A. Gilichinsky, E. I. Friedmann, *Microbial Ecol.* **33**, 169 (1997).
18. E. Willerslev, A. J. Hansen, H. N. Poinar, *Trends Ecol. Evol.* **19**, 141 (2004).
19. E. Willerslev *et al.*, *Science* **300**, 791 (2003).
20. R. Boom *et al.*, *J Clin Microbiol* **28**, 495 (1990).
21. N. Rohland, M. Hofreiter, *Biotechniques* **42**, 343 (2007).
22. C. E. King, R. Debruyne, M. Kuch, C. Schwarz, H. N. Poinar, *Biotechniques* **47**, 941 (2009).
23. O. Handt, M. Krings, R. H. Ward, S. Pääbo, *Am. J. Hum. Genet.* **59**, 368 (1996).
24. D. Doyle, K. J. McDowall, M. J. Butler, I. S. Hunter, *Mol. Microbiol.* **5**, 2923 (1991).
25. C. G. Marshall, G. Broadhead, B. K. Leskiw, G. D. Wright, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6480 (1997).
26. C. G. Marshall, I. A. D. Lessard, I.-S. Park, G. D. Wright, *Antimicrob. Agents Chemother.* **42**, 2215 (1998).
27. M. Margulies *et al.*, *Nature* **437**, 376 (2005).
28. R. D. Guthrie, *Frozen Fauna of the Mammoth Steppe: The Story of Blue Babe*. (University of Chicago Press, Chicago, 1990), pp. 338.
29. H. N. Poinar *et al.*, *Science* **311**, 392 (2006).
30. C. R. Harington, F. V. Clulow, *Can. J. Earth Sci.* **10**, 697 (1973).
31. J. A. Leonard *et al.*, *J. Archaeol. Sci.* **34**, 1361 (2007).

32. E. Willerslev *et al.*, *Curr. Biol.* **14**, R9 (2004).
33. Y.-S. Cui *et al.*, *Int. J. Syst. Evol. Microbiol.* **57**, 687 (2007).
34. V. M. D'Costa, K. M. McGrann, D. W. Hughes, G. D. Wright, *Science* **311**, 374 (2006).
35. E. S. Miller, C. R. Woese, S. Brenner, *Int. J. Syst. Bacteriol.* **41**, 363 (1991).
36. S. Qin *et al.*, *Appl. Environ. Microbiol.* **75**, 6176 (2009).
37. J.-H. Yoon, C.-H. Lee, T.-K. Oh, *Int. J. Syst. Evol. Microbiol.* **55**, 2171 (2005).
38. R. Sanchez-Pescador, J. T. Brown, M. Roberts, M. S. Urdea, *Nucleic Acids Res.* **16**, 1218 (1988).
39. J. P. Huelsenbeck, F. Ronquist, *Bioinformatics* **17**, 754 (2001).
40. L. E. Zawadzke, T. D. Bugg, C. T. Walsh, *Biochemistry* **30**, 1673 (1991).
41. D. E. Wampler, E. W. Westhead, *Biochemistry* **7**, 1661 (1968).
42. R. J. Leatherbarrow. (Erithacus Software Ltd, Horely, UK, 1992).
43. Z. Otwinowski, W. Minor, *Methods Enzymol.* **276**, 307 (1997).
44. C. C. Project, *Acta Crystallogr.* **50**, 760 (1994).
45. P. Emsley, K. Cowtan, *Acta Crystallogr.* **60**, 2126 (2004).
46. M. A. Ponder *et al.*, *FEMS Microbiol. Ecol.* **53**, 103 (2005).
47. T. A. Vishnivetskaya *et al.*, *Astrobiology* **6**, 400 (2006).
48. S. Z. Mindlin, V. S. Soina, M. A. Petrova, Z. M. Gorlenko, *Russ. J. Genet.* **44**, 27 (2008).
49. M. A. Petrova, Z. M. Gorlenko, V. S. Soina, S. Z. Mindlin, *Russ. J. Genet.* **44**, 1116 (2008).
50. M. Petrova, Z. Gorlenko, S. Mindlin, *FEMS Microbiol. Lett.* **296**, 190 (2009).
51. M. A. Petrova *et al.*, *Russ. J. Genet.* **38**, 1330 (2002).
52. G. Kholodii, S. Mindlin, M. Petrova, S. Minakhina, *FEMS Microbiol. Lett.* **226**, 251 (2003).
53. S. Mindlin *et al.*, *Res. Microbiol.* **156**, 994 (2005).
54. I. Massova, S. Mobashery, *Antimicrob Agents Chemother* **42**, 1 (1998).

CONCLUSION

With the success of this pilot study, we have demonstrated both the feasibility and value of ancient permafrost metagenomics. In addition to establishing the authenticity of the data by way of rigorous contamination precautions and controls, our experimental design also addressed the challenges involved in working with degraded DNA and in collecting taxonomically informative sequence data.

In choosing a targeted, amplicon sequencing approach, it was necessary to monitor and overcome PCR inhibition, which may otherwise limit our recovery of the diverse sequences present in a given sample. These qPCR-based tests enabled us to maximize DNA yields but also revealed limitations in the existing methods of detecting PCR inhibition, discussed in Part I. Most importantly, we observed many highly efficient reactions in which amplification of an internal positive control was greatly delayed (i.e. a shift in the expected C_q value); thus, inhibition may not be detectable using measurements of amplification efficiency alone – a positive control of known copy number/C_q value must be included. On the other hand, measurement of C_q shift alone is also insufficient, as the accuracy of this value depends on the amplification efficiency. With these observations in mind, our report expands on the recommendations for inhibition testing presented in the MIQE guidelines [5], with special consideration for low copy number samples. These results also highlight the different mechanisms of PCR inhibition, which may aid future research in the identification of inhibitor-specific PCR facilitators and resistant polymerases.

Our subsequent metagenomic assessment revealed an abundance of Late Pleistocene flora and fauna with no evidence of DNA leaching from younger strata; these results, coupled with extensive experimental controls, enabled us to explore the natural reservoir of antibiotic resistance in an environment unaffected by modern antibiotic use. As discussed in Part II, we not only showed that diverse antibiotic resistant determinants were present 30,000 years ago, but also verified their functionality. This report emphasizes the breadth of the antibiotic resistome, which must be considered in the development and use of antibiotics.

Although not the focus of the reports presented here, the wealth of paleoecological information we obtained from just milligrams of permafrost soil must not be understated. It is clear that our analyses of ancient ecosystems need no longer rely on macrofossil remains, or be limited by their spatial and temporal distribution. This offers tremendous potential for biogeographic research, by enabling us to reconstruct prehistoric chronologies of climate change and diversity, range dynamics, speciation, and extinction.

REFERENCES

1. Schweger, C.E., *Primary production and the Pleistocene ungulates - the productivity paradox*, in *Paleoecology of Beringia*, D.M. Hopkins, et al., Editors. 1982, Academic Press. p. 219-221.
2. Willerslev, E., et al., *Diverse plant and animal genetic records from Holocene and Pleistocene sediments.* Science, 2003. **300**(5620): p. 791-795.
3. Lydolph, M.C., et al., *Beringian paleoecology inferred from permafrost-preserved fungal DNA.* Appl. Environ. Microbiol., 2005. **71**(2): p. 1012-1017.
4. Margulies, M., et al., *Genome sequencing in microfabricated high-density picolitre reactors.* Nature, 2005. **437**(7057): p. 376-380.
5. Ramakers, C., et al., *Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data.* Neurosci. Lett., 2003. **339**(1): p. 62-66.
6. Bustin, S.A., et al., *The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments.* Clin. Chem., 2009. **55**(4): p. 611-622.
7. Liu, W. and D.A. Saint, *A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics.* Anal. Biochem., 2002. **302**(1): p. 52-59.
8. Liu, W. and D.A. Saint, *Validation of a quantitative method for real time PCR kinetics.* Biochem. Bioph. Res. Co., 2002. **294**(2): p. 347-353.
9. Tichopad, A., A. Dzidic, and M.W. Pfaffl, *Improving quantitative real-time RT-PCR reproducibility by boosting primer-linked amplification efficiency.* Biotechnol. Lett., 2002. **24**(24): p. 2053-2056.
10. Bar, T., et al., *Kinetic Outlier Detection (KOD) in real-time PCR.* Nucleic Acids Res., 2003. **31**(17): p. e105.
11. Peirson, S.N., J.N. Butler, and R.G. Foster, *Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis.* Nucleic Acids Res., 2003. **31**(14): p. e73.
12. Tichopad, A., et al., *Standardized determination of real-time PCR efficiency from a single reaction set-up.* Nucleic Acids Res., 2003. **31**(20): p. e122.
13. Rutledge, R.G., *Sigmoidal curve-fitting redefines quantitative real-time PCR with the prospective of developing automated high-throughput applications.* Nucleic Acids Res., 2004. **32**(22): p. e178.
14. Kontanis, E.J. and F.A. Reed, *Evaluation of real-time PCR amplification efficiencies to detect PCR inhibitors.* J. Forensic Sci., 2006. **51**(4): p. 795-804.
15. Guescini, M., et al., *A new real-time PCR method to overcome significant quantitative inaccuracy due to slight amplification inhibition.* BMC Bioinformatics, 2008. **9**: p. 326.
16. Rutledge, R.G. and D. Stewart, *Critical evaluation of methods used to determine amplification efficiency refutes the exponential character of real-time PCR.* BMC Mol. Biol., 2008. **9**(1): p. 96.
17. Shain, E.B. and J.M. Clemens, *A new method for robust quantitative and qualitative analysis of real-time PCR.* Nucleic Acids Res., 2008. **36**(14): p. e91.

18. Swillens, S., B. Dessars, and H.E. Housni, *Revisiting the sigmoidal curve fitting applied to quantitative real-time PCR data.* Anal. Biochem., 2008. **373**(2): p. 370-376.
19. Ruijter, J.M., et al., *Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data.* Nucleic Acids Res., 2009. **37**: p. e45.
20. Cloud, J.L., et al., *Description of a multiplex Bordetella pertussis and Bordetella parapertussis LightCycler PCR assay with inhibition control.* Diagn. Micr. Infec. Dis., 2003. **46**(3): p. 189-195.
21. Sutlović, D., et al., *Taq polymerase reverses inhibition of quantitative real time polymerase chain reaction by humic acid.* Croat. Med. J., 2005. **46**(4): p. 556-562.
22. Deagle, B.E., J.P. Eveson, and S.N. Jarman, *Quantification of damage in DNA recovered from highly degraded samples–a case study on DNA in faeces.* Front. Zool., 2006. **3**: p. 11.
23. Nolan, T., et al., *SPUD: a quantitative PCR assay for the detection of inhibitors in nucleic acid preparations.* Anal. Biochem., 2006. **351**(2): p. 308-310.
24. Swango, K.L., et al., *A quantitative PCR assay for the assessment of DNA degradation in forensic samples.* Forensic Sci. Int., 2006. **158**(1): p. 14-26.
25. Niederstätter, H., et al., *A modular real-time PCR concept for determining the quantity and quality of human nuclear and mitochondrial DNA.* Forensic Sci. Int.-Gen., 2007. **1**(1): p. 29-34.
26. Wright, G.D., *The antibiotic resistome: the nexus of chemical and genetic diversity.* Nature Rev. Microbiol., 2007. **5**(3): p. 175-186.
27. Wright, G.D., *The antibiotic resistome.* Expert Opin. Drug Dis., 2010. **5**(8): p. 779-788.
28. D'Costa, V.M., et al., *Sampling the antibiotic resistome.* Science, 2006. **311**(5759): p. 374-377.
29. Dantas, G., et al., *Bacteria subsisting on antibiotics.* Science, 2008. **320**(5872): p. 100-103.
30. Allen, H.K., et al., *Call of the wild: antibiotic resistance genes in natural environments.* Nature Rev. Microbiol., 2010. **8**(4): p. 251-259.
31. Allen, H.K., et al., *Functional metagenomics reveals diverse β-lactamases in a remote Alaskan soil.* ISME J., 2009. **3**(2): p. 243-251.
32. Harington, C.R. and F.V. Clulow, *Pleistocene mammals from Gold Run Creek, Yukon Territory.* Can. J. Earth Sci., 1973. **10**(5): p. 697-759.
33. Zazula, G.D., et al., *Plants, bugs, and a giant mammoth tusk: Paleoecology of Last Chance Creek, Yukon Territory*, in *Yukon Exploration and Geology 2002*, D.S. Edmond and L.L. Lewis, Editors. 2003. p. 251-258.
34. Zazula, G.D., et al., *Palaeobotany: Ice-age steppe vegetation in east Beringia.* Nature, 2003. **423**(6940): p. 603.
35. Zazula, G.D., et al., *Vegetation buried under Dawson tephra (25,300 ^{14}C years BP) and locally diverse late Pleistocene paleoenvironments of Goldbottom Creek, Yukon, Canada.* Palaeogeogr. Palaeocl., 2006. **242**(3-4): p. 253-286.
36. Froese, D.G., et al., *The Klondike goldfields and Pleistocene environments of Beringia.* GSA Today, 2009. **19**(8): p. 4-10.

37. Ponder, M.A., et al., *Characterization of potential stress responses in ancient Siberian permafrost psychroactive bacteria.* FEMS Microbiol. Ecol., 2005. **53**(1): p. 103-115.
38. Vishnivetskaya, T.A., et al., *Bacterial community in ancient Siberian permafrost as characterized by culture and culture-independent methods.* Astrobiology, 2006. **6**(3): p. 400-414.
39. Mindlin, S.Z., et al., *Isolation of antibiotic resistance bacterial strains from Eastern Siberia permafrost sediments.* Russ. J. Genet., 2008. **44**(1): p. 27-34.
40. Petrova, M.A., et al., *Association of the strA-strB genes with plasmids and transposons in the present-day bacteria and in bacterial strains from permafrost.* Russ. J. Genet., 2008. **44**(9): p. 1116-1120.
41. Petrova, M., Z. Gorlenko, and S. Mindlin, *Molecular structure and translocation of a multiple antibiotic resistance region of a Psychrobacter psychrophilus permafrost strain.* FEMS Microbiol. Lett., 2009. **296**(2): p. 190-197.
42. Willerslev, E., A.J. Hansen, and H.N. Poinar, *Isolation of nucleic acids and cultures from fossil ice and permafrost.* Trends Ecol. Evol., 2004. **19**(3): p. 141-147.
43. Sommer, M.O.A., G. Dantas, and G.M. Church, *Functional characterization of the antibiotic resistance reservoir in the human microflora.* Science, 2009. **325**(5944): p. 1128-1131.