

# Substitutions in woolly mammoth hemoglobin confer biochemical properties adaptive for cold tolerance

Kevin L Campbell<sup>1</sup>, Jason E E Roberts<sup>1</sup>, Laura N Watson<sup>2</sup>, Jörg Stetefeld<sup>3</sup>, Angela M Sloan<sup>1</sup>, Anthony V Signore<sup>1</sup>, Jesse W Howatt<sup>1</sup>, Jeremy R H Tame<sup>4</sup>, Nadin Rohland<sup>5,8</sup>, Tong-Jian Shen<sup>6</sup>, Jeremy J Austin<sup>2</sup>, Michael Hofreiter<sup>5,9</sup>, Chien Ho<sup>6</sup>, Roy E Weber<sup>7,10</sup> & Alan Cooper<sup>2,10</sup>

**We have genetically retrieved, resurrected and performed detailed structure-function analyses on authentic woolly mammoth hemoglobin to reveal for the first time both the evolutionary origins and the structural underpinnings of a key adaptive physiochemical trait in an extinct species. Hemoglobin binds and carries O<sub>2</sub>; however, its ability to offload O<sub>2</sub> to respiring cells is hampered at low temperatures, as heme deoxygenation is inherently endothermic (that is, hemoglobin-O<sub>2</sub> affinity increases as temperature decreases). We identify amino acid substitutions with large phenotypic effect on the chimeric β/δ-globin subunit of mammoth hemoglobin that provide a unique solution to this problem and thereby minimize energetically costly heat loss. This biochemical specialization may have been involved in the exploitation of high-latitude environments by this African-derived elephantid lineage during the Pleistocene period. This powerful new approach to directly analyze the genetic and structural basis of physiological adaptations in an extinct species adds an important new dimension to the study of natural selection.**

Understanding the genetic basis of molecular adaptation is of fundamental importance for the study of evolutionary biology. However, many physiological innovations evolved in organisms and paleoenvironments that are long extinct. The functional analysis of inferred ancestral gene products reconstructed from genomes of contemporary organisms has provided a unique means to link key adaptive evolutionary events with specific episodes in Earth's history<sup>1</sup>. For example, the duplication and rapid subfunctionalization of genes encoding pancreatic RNases early in the evolution of ruminant mammals has been associated with the rise of grassland ecosystems initiated by abrupt cooling in the early Oligocene era<sup>2,3</sup>. Unfortunately, this approach cannot be used to study the molecular evolution of lineages with no living descendants, for example,

recently extinct large Ice Age mammal species, although these are likely to have evolved under strong selection pressures.

Elephantids are an ideal model system in this regard, as it is firmly established that the three recent genera (*Loxodonta*, *Mammuthus* and *Elephas*) originated in warm equatorial Africa ~6.7–7.6 million years (Myr) ago<sup>4</sup>, with members of only the mammoth lineage successfully colonizing high latitudes 1.2–2.0 Myr ago<sup>5–7</sup>. Importantly, this incursion by ancestral mammoths into high latitudes coincided with abrupt climatic changes and an intensified cooling in the Arctic<sup>7</sup>. In contrast, most other taxa living in high-Arctic environments (such as the musk-ox and reindeer) do not have a clear evolutionary history of recent movement from tropical to arctic environments; they also do not have close, warm-adapted sister species for phylogenetically informative three-species comparisons.

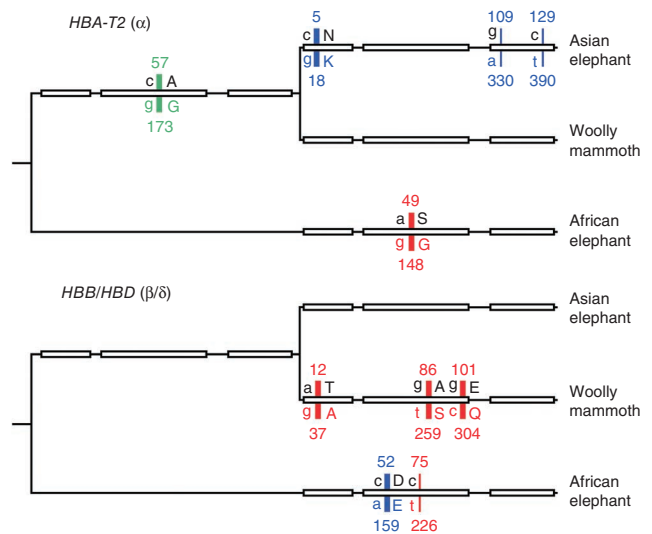
Woolly mammoths exhibited numerous adaptations for heat conservation in their adopted Arctic habitat, including small ears and tails<sup>7</sup> and a thick fur pelage with associated sebaceous glands<sup>8</sup>. Although studies have revealed amino acid replacements within specific mammoth proteins<sup>9,10</sup>, it has not been possible to link these changes with physiological specializations to the Arctic environment. To investigate this issue, we sequenced and synthesized authentic mammoth hemoglobin and identified amino-acid substitutions that confer mammoth hemoglobin with unique biochemical properties and underlie their adaptation to the cold.

We isolated DNA and mRNA from African (*Loxodonta africana*) and Asian (*Elaphas maximus*) elephant blood and amplified their adult-expressed α- and β-like globin genes. Translated gene products encoded 141- and 146-amino-acid polypeptides that precisely match the α-globin and β-globin chains determined for these species, respectively<sup>11,12</sup>. The transcribed α-like globin genes correspond to a locus we designated *HBA-T2* within the α-globin cluster of the draft *Loxodonta* genome (**Supplementary Note**), whereas the β-like genes match a locus we designated *HBB/HBD*, which is a chimeric β/δ fusion gene that originated via an ancient unequal crossover event

<sup>1</sup>Department of Biological Sciences, University of Manitoba, Winnipeg, Canada. <sup>2</sup>Australian Centre for Ancient DNA, School of Earth and Environmental Sciences, University of Adelaide, Adelaide, Australia. <sup>3</sup>Department of Chemistry, University of Manitoba, Winnipeg, Canada. <sup>4</sup>Protein Design Laboratory, Yokohama City University, Yokohama, Japan. <sup>5</sup>Junior Research Group Molecular Ecology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. <sup>6</sup>Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA. <sup>7</sup>Zoophysiology, Institute of Biological Sciences, University of Aarhus, Aarhus, Denmark. <sup>8</sup>Present address: Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA. <sup>9</sup>Present address: Department of Biology, University of York, York, UK. <sup>10</sup>These authors contributed equally to this work. Correspondence should be addressed to K.L.C. (campbelk@cc.umanitoba.ca) or A.C. (alan.cooper@adelaide.edu.au).

Received 18 December 2009; accepted 31 March 2010; published online 2 May 2010; doi:10.1038/ng.574

**Figure 1** Evolution of the genes encoding the single adult-expressed hemoglobin component of three members of the Elephantidae family. The position of nucleotide and amino acid changes are shown within the three coding regions (exons) of the *HBA-T2* and *HBB/HBD* globin genes (shown as open horizontal boxes along each branch) superimposed on the Elephantidae phylogeny<sup>4</sup>. Branch lengths are not proportional to geologic time. Ancestral nucleotide and amino acid residues are shown above, and derived nucleotide and amino acid residues are shown below the exons. The numbers above and the letters to the right of the vertical lines denote the amino acid residue, whereas the numbers below and the letters to the left of each vertical line indicate nucleotide position relative to the ATG initiation codon. Thick vertical lines with bold characters indicate nonsynonymous substitutions, and thin vertical lines represent synonymous substitutions, with red, green and blue characters and bars representing replacements at codon positions 1, 2 and 3, respectively. We employed the  $\alpha$ -globin and  $\beta$ -globin chain sequences of other afrotherian mammals (*Echinops telfairi* (GenBank P24291, P24292), *Procavia habessinica* (P01957, P02086) and *Trichechus inunguis* (P07414, P07415)) to deduce the direction of amino acid substitutions.



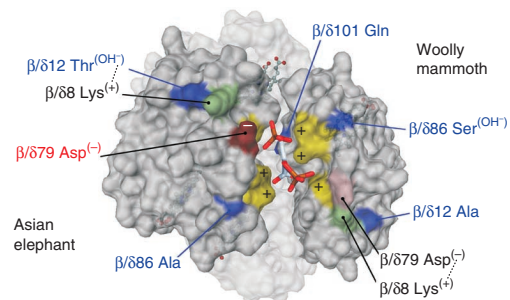
between the *HBD* and *HBB* genes (both of which are now silent) before the diversification of elephants, sea cows and hyraxes<sup>13</sup>.

Using DNA extracted from a ~43,000 year-old Siberian mammoth specimen<sup>9</sup>, we followed stringent ancient DNA methods to amplify the coding regions of the woolly mammoth *HBA-T2* (534 bp) and *HBB/HBD* (780 bp) genes (Supplementary Figs. 1–3 and Supplementary Tables 1 and 2). The mammoth polypeptides differ from the orthologous  $\alpha$ - and  $\beta/\delta$ -globin chains of the Asian elephant at 1 and 3 positions, respectively, and at 2 and 4 positions, respectively, compared to African elephants (Fig. 1). Although sequences of the mammoth *HBA-T2* gene have not changed in the ~6.7 Myr since the divergence of this line from *Elephas*<sup>4</sup>, the woolly mammoth  $\beta/\delta$ -globin chain has acquired three amino acid substitutions: T12A, A86S and E101Q (Fig. 1). Using an approach that targets specific SNPs<sup>9</sup>, we independently verified each position using the original specimen and two additional woolly mammoth samples collected from northern Siberia (Supplementary Figs. 4–8 and Supplementary Table 3). To test if the unusual signature (two transversions and one transition) and high nonsynonymous-to-synonymous nucleotide replacement ratio (3:0) in the mammoth *HBB/HBD* gene were indicative of positive selection, we performed a binomial test. However, given the small number of observations and relatively high proportion of coding sites that are potentially nonsynonymous (264 of 438 = 0.603), this pattern could not be distinguished from a neutral process ( $P = 0.216$ ). Consequently, we conducted empirical structure-functional analyses to investigate whether the observed nucleotide changes alter protein behavior in a manner consistent with positive selection.

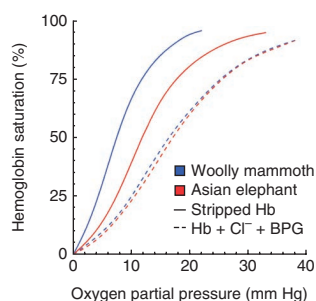
The location and nature of the altered residues are rare among mammalian  $\beta$ -type chains (Supplementary Fig. 9), with all three alterations fundamentally changing the physicochemical properties at each site. The mutations lie on the same side of the protein, with both the T12A and A86S substitutions occurring in exposed surface positions, whereas the E101Q substitution is located in the highly conserved sliding interface between the two rigid  $\alpha_1\beta_2$  dimer subunits (Fig. 2). The latter replacement is structurally important, as changes in the ligation state of the heme iron cause extensive conformational modifications (Supplementary Fig. 10) along this boundary as the  $\alpha\beta$  dimers slide and rotate in relation to one another<sup>14,15</sup>. All five known human hemoglobin variants with amino acid replacements at  $\beta 101$  possess higher intrinsic  $O_2$  affinities than the native protein<sup>15–18</sup>. However, only hemoglobin Rush, the human  $\beta$ -chain mutant displaying the equivalent amino acid substitution (E101Q) to that detected in the  $\beta/\delta$ -chain of mammoths, has an altered sensitivity to red-cell effectors<sup>15–17</sup>. This key attribute arises from the formation of two additional proton-linked  $Cl^-$  binding sites that counterbalances the increase in  $O_2$  affinity arising from the  $\beta 101$  alteration. Notably, the preferential binding of these additional ligands to the deoxy-state Rush protein ‘donates’ heat required for  $O_2$  offloading, thereby lowering the oxygenation enthalpy ( $\Delta H$ ) of the Rush molecule ( $-21.4 \text{ kJ mol}^{-1} O_2$  at pH 6.5) compared to normal human hemoglobin and the other  $\beta 101$  variants ( $-45.2 \text{ kJ mol}^{-1}$ )<sup>15–17</sup>.

A large negative enthalpy of hemoglobin oxygenation (where relatively small increases in temperature cause large decreases in

**Figure 2** Surface model of a chimeric Asian elephant (left) and mammoth (right) deoxyhemoglobin molecule bound to 2,3-bisphosphoglycerate (BPG). The locations of the three mammoth-specific amino acid substitutions are highlighted in blue, and the positions of each heme group are denoted by ball-and-stick diagrams. Regions highlighted in yellow denote positively charged residues (Lys82 of the  $\beta/\delta$ -chain (hereafter denoted  $\beta/\delta 82$  Lys),  $\beta/\delta 143$  His and the amino group of  $\beta/\delta 1$  Val) implicated in the binding of BPG to elephant deoxyhemoglobin<sup>11,12,14</sup>. Note that because the polar hydroxyl side chain of  $\beta/\delta 12$  Thr of Asian elephant deoxyhemoglobin (and human hemoglobin<sup>25</sup>) forms a hydrogen bond with the carbonyl group of  $\beta/\delta 88$  Lys (green), the negatively charged side chain of  $\beta/\delta 79$  Asp (red) is free to project into the BPG binding pocket, where it would tend to repel this anion. Conversely, the methyl side chain of  $\beta/\delta 12$  Ala in mammoth hemoglobin cannot bond with  $\beta/\delta 88$  Lys, allowing the lysyl side chain to form an ionic interaction with  $\beta/\delta 79$  Asp of the E helix and neutralizing its charge (light red). The mammoth-specific  $\beta/\delta 101$  Gln residue is spatially distant from this charged cluster and cannot contribute to BPG binding<sup>15–17</sup>. However, this central cavity residue alters electrostatic interactions at the sliding interface of the molecule that both destabilizes the low-affinity deoxy-state protein and creates additional proton-linked chloride binding sites in mammoth hemoglobin (see main text for details).



**Figure 3** Oxygen equilibrium curves of woolly mammoth (blue) and Asian elephant hemoglobin (red) at 37 °C and pH 7.0. In the absence of allosteric effectors (solid lines), the mammoth  $\beta/\delta$ -chain E101Q substitution destabilizes the tense-state (deoxy) conformation, leading to a protein phenotype with an intrinsic affinity nearly two times higher (curve is shifted to the left). This radical increase in  $O_2$  affinity (which would drastically impair tissue  $O_2$  offloading) is almost precisely compensated by enhanced  $H^+$ ,  $Cl^-$  and 2,3-BPG binding to mammoth hemoglobin that right-shifts the curve more strongly than in Asian elephant hemoglobin. As a result, the overall  $O_2$  affinity of mammoth hemoglobin in the presence of red cell effectors is nearly identical to that of Asian elephants at 37 °C (red dashed line). However, the increased effector binding to mammoth hemoglobin lowers the effect of temperature on  $O_2$  affinity, facilitating the release of  $O_2$  at cold temperatures in relation to Asian elephant hemoglobin.



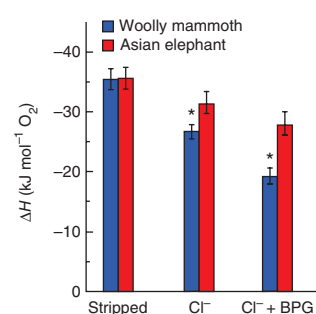
hemoglobin- $O_2$  affinity) is generally considered to be beneficial for mammals, as it promotes site-specific  $O_2$  offloading from the blood to warmer exercising muscles. In contrast, hemoglobins with numerically low  $\Delta H$  values (where changes in temperature have a relatively small effect on hemoglobin- $O_2$  affinity) are recognized to be energetically advantageous for cold-tolerant Arctic mammals (such as reindeer and musk-ox) because they help maintain adequate  $O_2$  delivery to the sparsely insulated limbs and distal appendages<sup>19,20</sup>, where temperatures may decrease to near 0 °C<sup>21</sup> (hence minimizing the thermal gradient for heat loss). The low  $\Delta H$  of reindeer and musk-ox hemoglobins ( $-14.0 \text{ kJ mol}^{-1}$  and  $-15.0 \text{ kJ mol}^{-1}$ , respectively, at pH 7.4) has been attributed to increased  $Cl^-$  binding<sup>19,20</sup>. Three basic residues (Lys8 of the  $\beta$ -chain (hereafter denoted  $\beta 8$  Lys),  $\beta 76$  Lys and  $\beta 77$  His) underlie this trait by forming a cationic cavity between helices A and E of each  $\beta$ -chain, which can be bridged by  $Cl^-$  in the deoxy state<sup>22</sup>. This structural motif, and a correspondingly low  $\Delta H$ <sup>19,20</sup>, is found in the hemoglobin of nearly all ruminant mammals examined (Supplementary Fig. 9), indicating it originated in an early Oligocene-era common ancestor. It is therefore possible that the low  $\Delta H$  values found for the hemoglobins of present day Arctic ruminants arose as an adaptive response to the pronounced cooling that initiated the development of grassland ecosystems ~35 Myr ago.

It is noteworthy that the  $\beta$ -type chains of all three recent elephantid genera possess this same positively charged  $Cl^-$ -binding cluster (Supplementary Fig. 9). To test whether these amino acids provide elephantid hemoglobin with similar reductions in  $\Delta H$ , and to assess the functional effects of the mammoth-specific residue changes, we first inserted Asian elephant *HBA-T2* and *HBB/HBD* cDNA into a hemoglobin expression vector and expressed this protein complex in *Escherichia coli*<sup>23</sup>. We used site-directed mutagenesis to introduce the mammoth-specific substitutions into the Asian elephant plasmid (Supplementary Table 4) and synthesized woolly mammoth hemoglobin *de novo*<sup>23</sup>. We then employed a thin-film technique<sup>24</sup> to measure the oxygen-binding characteristics of the elephant and mammoth proteins and their interaction with naturally occurring red-cell ligands ( $O_2$ , 2,3-bisphosphoglycerate (BPG),  $Cl^-$  and  $H^+$ ) at 10 °C, 25 °C and 37 °C. As predicted from the E101Q substitution in the mammoth  $\beta/\delta$  chain, oxygen equilibrium curves revealed striking functional differences between the two species (Fig. 3), with 'stripped' (allosteric-cofactor free) mammoth hemoglobin possessing

a higher  $O_2$  affinity at all three temperatures. However, mammoth hemoglobin also exhibited higher  $H^+$  and  $Cl^-$  effects, so that in the presence of red-cell effectors, the hemoglobin- $O_2$  affinity of both species was nearly identical at 37 °C (Fig. 3). In addition to their central role in lowering the  $O_2$  affinity of woolly mammoth hemoglobin, the additive exothermic contributions of  $Cl^-$  and BPG binding each significantly lowered the effect of temperature on the  $O_2$ -equilibrium properties of mammoth hemoglobin relative to Asian elephants (Fig. 4; Student's unpaired *t*-test;  $P = 0.0198$  and  $P = 0.0168$  for  $Cl^-$  and BPG, respectively). As a result, the mean overall  $\Delta H$  of mammoth hemoglobin ( $-19.3 \text{ kJ mol}^{-1}$ ) was reduced compared to that of Asian elephant hemoglobin ( $-28.1 \text{ kJ mol}^{-1}$ ). Consequently, the  $O_2$  affinity of mammoth blood is less affected by temperature than that of Asian elephants. This attribute may have been of fundamental adaptive importance for mammoths, whose blood presumably experienced large and rapid temperature changes when perfusing the limbs and other extremities in cold Arctic temperatures.

To identify the structural basis for the prominent functional differences between the two elephantids, we used high-resolution crystal structures of oxy and deoxy human hemoglobin<sup>25</sup> to construct molecular models of woolly mammoth and Asian elephant hemoglobin. In elephant hemoglobin, the carboxyl group of  $\beta/\delta 101$  Glu interacts closely with the guanidino group of  $\beta/\delta 104$  Arg of the same chain (Supplementary Fig. 10a). This interaction is also found in human hemoglobin<sup>25,26</sup>, and it has thus been proposed that the  $\beta$ -chain E101Q substitution of hemoglobin Rush allows the positive charge of  $\beta 104$  Arg to form a  $Cl^-$  binding site within the  $\alpha_1\beta_2$  interface of the deoxy-state molecule<sup>15,16</sup>. However, this mechanism seems unlikely, as the other human hemoglobin mutants that possess polar (lysine and aspartate) or nonpolar (alanine and glycine) residues at  $\beta 101$  did not exhibit altered  $Cl^-$  binding<sup>15-17</sup>. Indeed, our model illustrates that the uncharged  $\beta/\delta 101$  Gln residue of mammoth deoxyhemoglobin can still hydrogen bond with

**Figure 4** Mean enthalpy of oxygenation ( $\Delta H$ ;  $\text{kJ mol}^{-1} O_2$ ) values of woolly mammoth (blue columns) and Asian elephant (red columns) hemoglobin in the absence and presence of effector molecules. Error bars for each treatment ('stripped', 0.1 M  $Cl^-$ , and 0.1 M  $Cl^-$  plus saturating levels of 2,3-bisphosphoglycerate ( $Cl^-$  + BPG)) are  $\pm$  s.e.m. of four calculated  $\Delta H$  values: one from  $O_2$  equilibria measured at 10 °C and



25 °C at pH 7.0; one from measurements at 25 °C and 37 °C at pH 7.0; one from measurements at 10 °C and 25 °C at pH 7.4; and one from measurements at 25 °C and 37 °C at pH 7.4. The temperature dependence of the oxygenation process is governed by the associated overall  $\Delta H$  of this reaction<sup>19</sup>, where numerically low  $\Delta H$  values correspond to small effects of temperature on hemoglobin- $O_2$  affinity. Student's unpaired *t*-tests ( $\alpha = 0.05$ ,  $n = 4$ ) illustrate that the intrinsic thermal sensitivity of mammoth hemoglobin is not different from that of Asian elephant hemoglobin ( $P = 0.9174$ ). Conversely, as denoted by asterisks, the endothermic dissociation of  $Cl^-$  ( $P = 0.0198$ ) and BPG ( $P = 0.0168$ ) each independently lower the oxygenation enthalpy of mammoth hemoglobin to significantly greater degrees than for Asian elephant hemoglobin, as predicted by the E101Q and T12A substitutions on the mammoth  $\beta/\delta$ -globin chain, respectively (see text for details). The  $\Delta H$  of mammoth hemoglobin was independent of pH under all conditions employed here, illustrating that the binding of Bohr protons does not directly contribute to lowering the  $\Delta H$  value.

$\beta/\delta$ 104 Arg (Supplementary Fig. 10b), as has been suggested for hemoglobin Rush<sup>26</sup>. Our physiological data also demonstrates that, as is the case with hemoglobin Rush, deoxygenation of mammoth blood involves the binding of two additional ions ( $H^+$  and  $Cl^-$ ) per dimer as compared to elephant hemoglobin. Consequently, the potentially maladaptive increase in hemoglobin- $O_2$  affinity that accompanies the E101Q replacement on the mammoth  $\beta/\delta$ -chain (Fig. 3) is countered by the formation of these proton-linked  $Cl^-$  binding sites. Notably, the increased  $Cl^-$  binding also contributes an additional endothermic component that lowers the  $\Delta H$  of mammoth hemoglobin (Fig. 4). It must be emphasized that the effect of position  $\beta$ 101 on the functional properties of human hemoglobin is mediated by the size<sup>15</sup> and specific side-chain chemistry of the substituted amino acid residue<sup>27</sup>, meaning that only glutamine can reduce the  $\Delta H$ . Any substitution other than E101Q in the mammoth  $\beta/\delta$ -chain would almost certainly have been maladaptive (for example, the rare heterozygous human  $\beta$ -chain variants British Columbia (E101K), Alberta (E101G) and Potomac (E101D) all produce radically increased blood  $O_2$  affinities and, consequently, clinical erythrocytosis<sup>17</sup>; the  $\Delta H$  of these hemoglobins is also unchanged from that of normal human hemoglobin<sup>15</sup>).

The  $\beta/\delta$ 101 Gln residue of mammoth hemoglobin is physically separate from the BPG binding pocket<sup>15–17</sup> (Supplementary Fig. 11); thus, the BPG-induced reduction in the overall  $\Delta H$  of the protein (Fig. 4) must arise from another residue replacement. Our molecular model illustrates that the T12A substitution of the mammoth  $\beta/\delta$ -chain mediates this alteration by drawing  $\beta/\delta$ 79 Asp away from the cationic BPG binding cavity of mammoth deoxyhemoglobin (Fig. 2). Therefore, two (E101Q and T12A) of the three  $\beta/\delta$ -chain substitutions observed in mammoths independently lower the temperature sensitivity of the mammoth's blood.

Consistent with the additional  $Cl^-$  binding site hypothesis<sup>19,20,22</sup>, the  $\Delta H$  of Asian elephant hemoglobin in the presence of 0.1 M  $Cl^-$  ( $-31.5$  kJ mol<sup>-1</sup>; Fig. 4) is lower than that of human hemoglobin ( $-41.0$  kJ mol<sup>-1</sup>)<sup>20</sup>, which lacks the  $\beta$ 8 Lys- $\beta$ 76 Lys- $\beta$ 77 His motif (Supplementary Fig. 9). However, Asian elephant hemoglobin has a noticeably lower  $Cl^-$  sensitivity than bovine and human hemoglobin (that is, it binds fewer  $Cl^-$  ions; Supplementary Fig. 12). This finding demonstrates that although these three residues form a  $Cl^-$  binding site in Arctic ruminant hemoglobins<sup>20,22</sup>, they do not bind  $Cl^-$ , nor do they contribute to lowering the  $\Delta H$  of elephantid hemoglobins (presumably because of structural differences in this region of the protein). This illustrates the fact that acquisition of equivalent substitutions in independent evolutionary lineages may have different functional consequences on protein behavior and indicates that the elevated  $Cl^-$  sensitivity and attendant  $\Delta H$  reduction of mammoth hemoglobin must arise from unique molecular mechanisms (the  $\beta/\delta$ -chain T12A and E101Q substitutions) that so far have not been described in other living mammals. By using an empirical structure-function analysis of a resurrected gene product, we have identified physiological properties of woolly mammoth hemoglobin that may have played an important role in the adaptation of this African-derived lineage to Arctic environments during the Pleistocene era. This powerful paleogenetic approach adds a new dimension to the investigation of the evolutionary processes and constraints that shaped the adaptive physiological phenotypes of both extinct and extant species.

## METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/naturegenetics/>.

**Accession numbers.** All gene sequences have been deposited in GenBank with the accession numbers FJ716079–FJ716094.

**Requests for materials.** [campbelk@cc.umanitoba.ca](mailto:campbelk@cc.umanitoba.ca).

*Note: Supplementary information is available on the Nature Genetics website.*

## ACKNOWLEDGMENTS

We thank T. Kuznetsova for the mammoth samples, W. Korver for providing Asian and African elephant blood and A. Bang, T.C. Tam, N. Ho, J. Hare, J. da Silva and M. Pagel for technical assistance. Financial support was provided by the National Sciences and Engineering Research Council (NSERC) of Canada (K.L.C. and J.S.), Winnipeg Foundation (K.L.C.), University of Manitoba Research Grant Program (K.L.C.), Japan Society for the Promotion of Science (J.R.H.T.), Max Planck Society (M.H. and N.R.), US National Institutes of Health grant R01GM-084614; C.H.), Danish Natural Science Research Council and the Carlsberg Foundation (R.E.W.) and the Australian Research Council (A.C. and L.N.W.). J.E.E.R., J.W.H. and A.V.S. were supported by NSERC Undergraduate Research Awards, and A.M.S. was supported by a University of Manitoba Graduate Fellowship.

## AUTHOR CONTRIBUTIONS

K.L.C. conceived the research. K.L.C., J.S., M.H., J.J.A., T.-J.S., C.H., R.E.W. and A.C. designed the experiments. K.L.C., J.E.E.R., L.N.W., A.M.S., A.V.S., J.W.H., N.R., T.-J.S., R.E.W. and J.J.A. conducted the experiments. K.L.C., J.S., A.V.S., J.R.H.T., R.E.W. and A.C. analyzed the data. K.L.C. and A.C. drafted the manuscript, and K.L.C., M.H., J.R.H.T., C.H., R.E.W. and A.C. contributed to the final manuscript writing and its revisions.

## COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/naturegenetics/>.

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>.

- Benner, S.A., Caraco, M.D., Thomson, M. & Gaucher, E.A. Planetary biology—paleontological, geological, and molecular histories of life. *Science* **296**, 864–868 (2002).
- Jermann, T.M., Optiz, J.G., Stackhouse, J. & Benner, S.A. Reconstructing the evolutionary history of the artiodactyl ribonuclease superfamily. *Nature* **374**, 57–59 (1995).
- Beintema, J.J., Schuller, C., Irie, M. & Carsana, A. Molecular evolution of the ribonuclease superfamily. *Prog. Biophys. Mol. Biol.* **51**, 165–192 (1988).
- Rohland, N. *et al.* Proboscidean mitogenomics: chronology and mode of elephant evolution using mastodon as outgroup. *PLoS Biol.* **5**, e207 (2007).
- Lister, A.M., Sher, A.V., van Essen, H. & Wei, G. The pattern and process of mammoth evolution in Eurasia. *Quat. Int.* **126–128**, 49–64 (2005).
- Lister, A.M. The impact of Quaternary Ice Ages on mammalian evolution. *Phil. Trans. R. Soc. Lond. B* **359**, 221–241 (2004).
- Lister, A. & Bahn, P.G. *Mammoths: Giants of the Ice Age* (University of California Press, Berkeley, California, USA, 2007).
- Repin, V.E., Taranov, O.S., Ryabchikova, E.I., Tikhonov, A.N. & Pugachev, V.G. Sebaceous glands of the woolly mammoth, *Mammuthus primigenius* Blum: histological evidence. *Dokl. Biol. Sci.* **398**, 382–384 (2004).
- Römpler, H. *et al.* Nuclear gene indicates coat-color polymorphism in mammoths. *Science* **313**, 62 (2006).
- Miller, W. *et al.* Sequencing the nuclear genome of the extinct woolly mammoth. *Nature* **456**, 387–390 (2008).
- Braunitzer, G., Jelkmann, W., Stangl, A., Schrank, B. & Krombach, C. Hemoglobins, XLVIII: the primary structure of hemoglobin of the Indian elephant (*Elephas maximus*, Proboscidea): beta2=Asn. *Hoppe-Seyler's Z. Physiol. Chem.* **363**, 683–691 (1982).
- Braunitzer, G., Stangl, A., Schrank, B., Krombach, C. & Wiesner, H. Phosphate-haemoglobin interaction. The primary structure of the haemoglobin of the African elephant (*Loxodonta africana*, Proboscidea): asparagine in position 2 of the beta-chain. *Hoppe-Seyler's Z. Physiol. Chem.* **365**, 743–749 (1984).
- Opazo, J.C., Sloan, A.M., Campbell, K.L. & Storz, J.F. Origin and ascendancy of a chimeric fusion gene: the beta/delta-globin gene of paenungulate mammals. *Mol. Biol. Evol.* **26**, 1469–1478 (2009).
- Perutz, M.F. Species adaptation in a protein molecule. *Mol. Biol. Evol.* **1**, 1–28 (1983).
- Shih, D.T., Jones, R.T., Imai, K. & Tyuma, I. Involvement of Glu G3(101)beta in the function of hemoglobin. Comparative  $O_2$  equilibrium studies of human mutant hemoglobins. *J. Biol. Chem.* **260**, 5919–5924 (1985).
- Baudin, V. *et al.* Functional consequences of mutations at the allosteric interface in hetero- and homo-hemoglobin tetramers. *Protein Sci.* **2**, 1320–1330 (1993).

17. Jones, R.T. & Shih, T.B. Hemoglobin variants with altered oxygen affinity. *Hemoglobin* **4**, 243–261 (1980).
18. Nakatsuji, T., Shimizu, K. & Huisman, T.H. Hb F-La Grange or  $\alpha_2\gamma_2101(\text{G3})\text{Glu}\rightarrow\text{Lys}$ ; 751Ile; 136Gly: A high oxygen affinity fetal haemoglobin variant observed in a caucasian newborn. *Biochim. Biophys. Acta* **789**, 224–228 (2004).
19. Clementi, M.E., Condò, S.G., Castagnola, M. & Giardina, B. Hemoglobin function under extreme life conditions. *Eur. J. Biochem.* **223**, 309–317 (1994).
20. De Rosa, M.C., Castagnola, M., Bertonati, C., Galtier, A. & Giardina, B. From the Arctic to fetal life: physiological importance and structural basis of an 'additional' chloride-binding site in haemoglobin. *Biochem. J.* **380**, 889–896 (2004).
21. Irving, L. & Krog, J. Temperature of skin in the arctic as a regulator of heat. *J. Appl. Physiol.* **7**, 355–364 (1955).
22. Fronticelli, C. *et al.* Allosteric modulation by tertiary structure in mammalian hemoglobins. Introduction of the functional characteristics of bovine hemoglobin into human hemoglobin by five amino acid substitutions. *J. Biol. Chem.* **270**, 30588–30592 (1995).
23. Shen, T.J. *et al.* Production of unmodified human adult hemoglobin in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **90**, 8108–8112 (1993).
24. Weber, R.E. Use of ionic and zwitterionic (Tris/BisTris and HEPES) buffers in studies on hemoglobin function. *J. Appl. Physiol.* **72**, 1611–1615 (1992).
25. Park, S.-Y., Yokoyama, T., Shibayama, N., Shiro, Y. & Tame, J.R.H. 1.25 Å resolution crystal structures of human haemoglobin in the oxy, deoxy and carbonmonoxy forms. *J. Mol. Biol.* **360**, 690–701 (2006).
26. Abraham, D.J., Peascoe, R.A., Randad, R.S. & Panikker, J. X-ray diffraction study of di and tetra-ligated T-state hemoglobin from high salt crystals. *J. Mol. Biol.* **227**, 480–492 (1992).
27. Turner, G.J. *et al.* Mutagenic dissection of hemoglobin cooperativity: Effects of amino acid alteration on subunit assembly of oxy and deoxy tetramers. *Prot. Str. Funct. Gen.* **14**, 333–350 (1992).

## ONLINE METHODS

### Extraction, amplification, cloning and sequencing of elephant DNA/RNA.

To avoid the risk of cross-contamination, all studies of modern elephant samples were performed at the Department of Biological Sciences, University of Manitoba, Winnipeg.

Subsamples of freshly collected African ( $n = 2$ ) and Asian ( $n = 2$ ) elephant whole blood were immediately stored at  $-70\text{ }^{\circ}\text{C}$  or transferred to 10 volumes of RNA/DNA stabilization reagent for blood/bone marrow (Roche Molecular Biochemicals), mixed and frozen. DNA was extracted from untreated blood following addition of phosphate-buffered saline and proteinase K using the Qiagen DNeasy extraction kit. Isolation of mRNA was conducted on 10 ml of blood lysate samples using the Roche mRNA purification kit, and cDNA was synthesized with the Superscript II RT cDNA kit using Oligo (dT) primers (Invitrogen).

DNA amplifications consisted of a 30-cycle protocol ( $94\text{ }^{\circ}\text{C}$  for 30 s;  $50\text{ }^{\circ}\text{C}$  for 15 s;  $72\text{ }^{\circ}\text{C}$  for 30–75 s) followed by 10 min at  $72\text{ }^{\circ}\text{C}$ ; this was then followed by nested PCR using an annealing temperature gradient of  $50\text{ }^{\circ}\text{C}$  to  $60\text{ }^{\circ}\text{C}$ . Target bands were excised and purified using the MinElute Gel Extraction Kit (Qiagen) and either sequenced directly (see below) or ligated using a Qiagen PCR Cloningplus Kit, and incubated on agar plates. Obtained nucleotide sequences were used to design primers to amplify the flanking regions of each gene via a walking reaction (APAgene Genome Walking Kit; Bio S&T Inc.). Plasmids containing target bands were primed with a BigDye Sequencing Kit using the universal sequencing primers, M-13(F)-40 and M-13(R), and sequenced in both directions with a 3730 ABI PRISM Genetic Analyzer (Applied Biosystems).

Flanking sequences of the *Elephas* and *Loxodonta* *HBA-T2* and *HBB/HBD* genes were used to design primers for cDNA templates using the following cycling conditions:  $96\text{ }^{\circ}\text{C}$  for 3 min, (35–40 cycles of  $96\text{ }^{\circ}\text{C}$  for 30 s,  $55\text{--}65\text{ }^{\circ}\text{C}$  for 15 s,  $72\text{ }^{\circ}\text{C}$  for 30–40 s), followed by  $72\text{ }^{\circ}\text{C}$  for 6 min. Target bands were excised and sequenced directly in both directions with internal nested primers on an Applied Biosystems 3130 Genetic Analyzer.

Consensus gene alignments were constructed for each species using Sequencher (Version 4.2.2) software. Contigs for individual *HBA-T2* and *HBB/HBD* genes from each species incorporated between 11 and 24 overlapping fragments amplified from five to nine separate PCRs, respectively.

**Mammoth DNA amplification, amplicon separation, purification, cloning and sequencing.** *Adelaide.* Amplification and sequencing of the complete coding regions of the mammoth *HBA-T2* and *HBB/HBD* genes were performed in a dedicated ancient DNA facility in Adelaide, where modern elephant DNA has never been present. DNA previously extracted from a mammoth specimen (SP1349 (KIA27805)) recovered from the Kolyma Lowland, Yakutia ( $70^{\circ}\text{ N}$ ,  $151^{\circ}\text{ E}$ ), and dated  $\sim 43,000$  years old<sup>9</sup> was used to amplify the three coding exons of both the mammoth *HBA-T2* and *HBB/HBD* globin genes and the *HBB/HBD* 5' promoter region. Primers (Supplementary Tables 1 and 2) were designed from the alignment of the *Elephas* and *Loxodonta* *HBA-T2* and *HBB/HBD* sequences to target regions within the functional genes that differed significantly from paralogous genes. The mammoth *HBA-T2* and *HBB/HBD* exons and the 5' upstream promoter region of *HBB/HBD* were amplified using a combination of singleplex reactions (*HBB/HBD* 5' promoter), singleplex reactions with DMSO (*HBA-T2* fragments F1–F6) and multiplex and nested singleplex reactions (*HBB/HBD* fragments F1–F5).

Multiplex PCRs were performed in a 25  $\mu\text{l}$  final volume, with 1.25 U of Platinum Taq DNA Polymerase High Fidelity; 200  $\mu\text{M}$  each of dNTP, 4 mM  $\text{MgSO}_4$ , 1 $\times$  high fidelity PCR buffer, 1 mg  $\text{ml}^{-1}$  BSA buffer; 0.15  $\mu\text{M}$  of each primer; and 1  $\mu\text{l}$  of an ancient DNA extract diluted 1:10 with DNA-free water. Thermocycling conditions were  $94\text{ }^{\circ}\text{C}$  for 2 min, 40 cycles of  $94\text{ }^{\circ}\text{C}$  for 15 s,  $55\text{ }^{\circ}\text{C}$  for 15 s,  $68\text{ }^{\circ}\text{C}$  for 30 s and a final extension at  $68\text{ }^{\circ}\text{C}$  for 10 min. Singleplex PCRs were performed in 25  $\mu\text{l}$  final volume, with 0.2 U of Hotmaster Taq, 200  $\mu\text{M}$  of each dNTP, 0.4  $\mu\text{M}$  of each primer, and 1 $\times$  Hotmaster buffer (2.5 mM  $\text{Mg}^{2+}$ ). Templates were either 1  $\mu\text{l}$  of the multiplex PCR reaction diluted 1:40 in DNA-free water, or 1  $\mu\text{l}$  of ancient DNA extract diluted 1:10 in DNA-free water. Thermocycling conditions were  $94\text{ }^{\circ}\text{C}$  for 2 min, 30 cycles of  $94\text{ }^{\circ}\text{C}$  for 20 s,  $55\text{ }^{\circ}\text{C}$  for 10 s,  $65\text{ }^{\circ}\text{C}$  for 30 s and a final extension of  $65\text{ }^{\circ}\text{C}$  for 2 min. Because of the high GC content of the *HBA-T2* gene, we added 5% DMSO to the standard singleplex reaction mixes described above. All PCR reactions were purified with AMPure (Agencourt) according to the manufacturer's instructions.

Unique nucleotide changes on the *HBB/HBD* gene were confirmed by cloning with the TOPO TA Cloning Kit (Invitrogen) according to the manufacturer's instructions. Colonies were picked, stored in 10 mM Tris, and cells were lysed by heating to  $96\text{ }^{\circ}\text{C}$  for 10 min. Between 10 and 14 colonies were sequenced for each cloned fragment (see Supplementary Figs. 1, 4 and 5). Colony PCRs were performed in 25  $\mu\text{l}$  final volume with 0.2 U of Hotmaster Taq, 200  $\mu\text{M}$  of each dNTP, 0.2  $\mu\text{M}$  of each primer (M13F-20 and M13R), 1 $\times$  Hotmaster buffer (2.5 mM  $\text{Mg}^{2+}$ ) and 2  $\mu\text{l}$  of lysed colony as template. Thermocycling conditions were  $94\text{ }^{\circ}\text{C}$  for 2 min, 35 cycles of  $94\text{ }^{\circ}\text{C}$  for 20 s,  $55\text{ }^{\circ}\text{C}$  for 10 s,  $65\text{ }^{\circ}\text{C}$  for 45 s and a final extension of  $65\text{ }^{\circ}\text{C}$  for 10 min. PCR products were sequenced in both directions using BigDye Terminator v3.1 (Applied Biosystems) and the M13 reverse primer only. Reactions were carried out in a 20  $\mu\text{l}$  final volume with 7.5  $\mu\text{l}$  of 2.5 $\times$  reaction buffer, 0.16  $\mu\text{M}$  of primer and 0.5  $\mu\text{l}$  BigDye enzyme. BigDye thermocycling conditions were  $96\text{ }^{\circ}\text{C}$  for 1 min, followed by 25 cycles of  $96\text{ }^{\circ}\text{C}$  for 10 s,  $50\text{ }^{\circ}\text{C}$  for 5 s and  $60\text{ }^{\circ}\text{C}$  for 4 min. BigDye reactions were purified using CleanSEQ (Agencourt) according to manufacturer's instructions and sequenced on an ABI 3730 sequencer.

*Leipzig.* To replicate the Adelaide *HBB/HBD* results, at least two independent multiplex PCRs<sup>28</sup> were performed on woolly mammoth specimen SP1349 (KIA 27805) and two additional Siberian mammoth samples (specimen NS-OgK-O 271 (SP1421) collected at  $72.4^{\circ}\text{ N}$ ,  $143.4^{\circ}\text{ E}$  and specimen MaK-0 101 (SP1419) collected at  $73.6^{\circ}\text{ N}$ ,  $117.2^{\circ}\text{ E}$ ; for details see ref. 9). Separate studies in the Leipzig lab have shown that specimen SP1349 falls into mitochondrial clade 2, whereas SP 1419 and SP1421 fall into clade 1 (ref. 29). Multiplex reactions were performed using three primer pairs targeting the three nonsynonymous SNPs (Supplementary Table 3). The same primers were then used separately for each target. PCR products were visualized on agarose gels, were cloned directly, and a minimum of three clones per SNP for each multiplex PCR were sequenced on an ABI 3730 sequencer (Supplementary Figs. 6–8).

**Expression of recombinant Asian elephant and woolly mammoth hemoglobins.** The  $\alpha$ - and  $\beta$ -globin genes in the normal human adult hemoglobin-expression plasmid pHE2<sup>23</sup> were replaced by Asian elephant *HBA-T2* and *HBB/HBD* coding sequences to form expression plasmid pHE27E. The mammoth hemoglobin expression plasmid pHE27M was created by introducing the  $\alpha$ -chain (K5N) and  $\beta/\delta$ -chain substitutions T12A, A86S and E101Q into the pHE27E plasmid via site-directed mutagenesis using a QuikChange II XL mutagenesis kit (Stratagene). The procedures for expression, isolation and purification of Asian elephant and mammoth recombinant hemoglobin followed those previously described<sup>23</sup>.

**Oxygen binding measurements.** Immediately before  $\text{O}_2$  equilibrium determinations, appropriate volumes of water, HEPES buffer and, when applicable, standard KCl and 2,3-BPG solutions were added to small aliquots of each hemoglobin solution (the final heme and buffer concentrations were 0.21 mM and 0.1 M, respectively). Oxygen-binding equilibria for each sample exposed to stepwise increases in  $\text{O}_2$  tension were measured at  $10\text{ }^{\circ}\text{C}$ ,  $25\text{ }^{\circ}\text{C}$  and  $37\text{ }^{\circ}\text{C}$  at both pH 7.0 and pH 7.4 using a modified diffusion chamber<sup>24</sup>. Oxygen half-saturation pressures ( $P_{50}$ ) and cooperativity coefficients ( $n_{50}$ ) interpolated from Hill plots were calculated from at least four equilibration steps between 30% and 70% saturation for each trial<sup>24</sup>.  $\text{Cl}^-$  concentration for each sample was assessed using a Sherwood MKII Model 926S Chloride Analyzer (Sherwood Scientific Ltd.), and pH was measured in oxygenated subsamples equilibrated to each of the three experimental temperatures using a Radiometer BMS2 Mk2 Blood Micro System and PHM 64 Research pH meter (Radiometer). Stock solutions of BPG added to the hemoglobin samples were assayed using Sigma enzymatic test chemicals. The overall enthalpy of oxygenation ( $\Delta H$ , kJ  $\text{mol}^{-1}\text{ O}_2$ ) values were calculated as

$$-2.303 \times R \times \Delta \log P_{50} / (T_1 - T_2)$$

where  $T_1$  and  $T_2$  are two absolute temperatures and  $R$  is the gas constant. All  $\Delta H$  values were corrected to exclude the heat of solvation of  $\text{O}_2$  ( $-12.55\text{ kJ mol}^{-1}$ )<sup>19</sup>.

**Molecular modeling.** We separately introduced all amino acid substitutions occurring in the woolly mammoth, African and Asian elephant hemoglobin into the 3D oxy (PDB 2DN1) and deoxy (PDB 2DN2) models of the recently

refined 1.25 Å structures of human hemoglobin<sup>25</sup> using the MODELLER function of the Insight II program package version 97.2 (BioSym Technologies). The strain energy in the vicinity of each substitution was generated in the GROMOS force field using the 53A6 parameter set optimized for molecular dynamics simulations<sup>30</sup>. The strain energy was subsequently minimized using the GROMACS package (version 3.3; see URLs). This involved a brief steepest-descents run that employed a maximum step-size protocol of 1 Å and a maximum tolerance of 1,000 kJ mol<sup>-1</sup> nm<sup>-1</sup>. This was followed by a more extensive conjugate-gradients minimization with a tolerance of 100 kJ mol<sup>-1</sup> nm<sup>-1</sup>. A Morse oscillator model was used to represent covalent bonding in the conjugate-gradients minimization step and a harmonic oscillator approximation

was used for the steepest-descents protocol. Three-dimensional molecular representations were visualized with DINO version 0.9.1 (see URLs).

**URLs.** GROMACS, <http://www.gromacs.org/>; DINO, <http://www.dino3d.org/>.

28. Römler, H. *et al.* Multiplex amplification of ancient DNA. *Nat. Protoc.* **1**, 720–728 (2006).
29. Barnes, I. *et al.* Genetic structure and extinction of the woolly mammoth, *Mammuthus primigenius*. *Curr. Biol.* **17**, 1072–1075 (2007).
30. Oostenbrink, C., Villa, A., Mark, A.E. & van Gunsteren, W.F. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **25**, 1656–1676 (2004).