



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Low coverage sequencing of two Asian elephant (*Elephas maximus*) genomes

Citation for published version:

Dastjerdi, A, Robert, C & Watson, M 2014, 'Low coverage sequencing of two Asian elephant (*Elephas maximus*) genomes', *GigaScience*, vol. 3, no. 1, pp. 12. <https://doi.org/10.1186/2047-217X-3-12>

Digital Object Identifier (DOI):

[10.1186/2047-217X-3-12](https://doi.org/10.1186/2047-217X-3-12)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

GigaScience

Publisher Rights Statement:

© 2014 Dastjerdi et al.; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



DATA NOTE

Open Access

Low coverage sequencing of two Asian elephant (*Elephas maximus*) genomes

Akbar Dastjerdi^{1†}, Christelle Robert^{2†} and Mick Watson^{3*}

Abstract

Background: There are three species of elephant that exist, the Asian elephant (*Elephas maximus*) and two species of African elephant (*Loxodonta africana* and *Loxodonta cyclotis*). The populations of all three species are dwindling, and are under threat due to factors, such as habitat destruction and ivory hunting. The species differ in many respects, including in their morphology and response to disease. The availability of elephant genome sequence data from all three elephant species will complement studies of behaviour, genetic diversity, evolution and disease resistance.

Findings: We present low-coverage Illumina sequence data from two Asian elephants, representing approximately 5X and 2.5X coverage respectively. Both raw and aligned data are available, using the African elephant (*L. africana*) genome as a reference.

Conclusions: The data presented here are an important addition to the available genetic and genomic information on Asian and African elephants.

Keywords: Elephant, Genome, Sequencing, Variation

Data description

Background

Three species of elephant exist, the Asian elephant (*Elephas maximus*) [1] and two species of African elephant (*Loxodonta africana*, *Loxodonta cyclotis*) [2,3]. Elephants are the Earth's largest land mammal, with the Asian elephant (*E. maximus*) being slightly smaller than African elephants (*L. africana* and *L. cyclotis*). Elephants can reach a height of 4 m, a weight of over 10,000 kg, and adults can have a lifespan as long as 60–70 years. Asian elephants have the following distinctive characteristics: smaller rounded ears, arched back, hemispherical bulges on the head, differing number of nails on the legs, and finger-like features on the trunk. Asian elephant males and both sexes of African elephants can grow tusks over 2 m in length and 45 kg in weight. Asian elephants also have fewer ribs, more molar plates and a small intestine twice as long as that in African elephants [1,2,4]. The elephant species also differ in their nutritional needs [5] and susceptibility to infectious diseases [6].

Characteristic features of elephants make these animals unique. Elephant trunks have many functions (smelling, breathing, trumpeting, drinking and grabbing) and contain approximately 150,000 different muscle fascicles [7], the largest number in any single body part. Elephants have also a longer gestation period than any other mammal, almost 22 months [8]. At birth, elephant calves already weigh some 90 kg and stand about 1 m tall.

Both *E. maximus* and *L. africana* are listed as endangered species by the International Union for Conservation of Nature (IUCN) [9]. Wild populations are restricted to increasingly limited areas of land and are under pressure with regard to habitat loss, poaching, human invasion and various diseases.

The genome of the African elephant, *L. africana*, has been sequenced to 6.8X coverage by The Broad Institute [10] (Genbank Assembly ID: GCA_000001905.1; BioProject accession ID: PRJNA12569) and has been annotated by Ensembl [11]. Here, we present whole-genome sequencing datasets from two Asian elephants (*E. maximus*), both of which are an important addition to the existing public data. The availability and a better understanding of elephant genomes will facilitate efforts to conserve populations of these animals by enabling studies of behaviour,

* Correspondence: mick.watson@roslin.ed.ac.uk

†Equal contributors

³Edinburgh Genomics, University of Edinburgh, EH25 9RG Easter Bush, UK
Full list of author information is available at the end of the article

social organisation, population structure, genetic diversity and phylogeny. In addition, the genome sequences will facilitate development of novel tools to aid in combating the illegal trade of wild elephant and ivory, as well as deciphering their body physiology and immunity.

Original purpose

Two samples, 577_1_Emelia and 577_2_Raman, were originally sequenced to enable the characterisation of elephant endotheliotropic herpesvirus (EEHV) [12], which is associated with a life-threatening haemorrhagic disease in Asian elephants [6]. Despite the virus representing only a tiny fraction of the reads (0.169% and 0.038% respectively), full genome sequences of both EEHV1A and EEHV1B were recovered. A small number of the reads were also used in the resolution of the type specimen of the Asian elephant (*E. maximus*), which was instead found to originate from an elephant of the *Loxodonta* genus [13].

Sample treatment and sequencing

Sample collection, treatment and sequencing have been previously described [12]. Briefly, post-mortem samples were collected from two infant Asian elephants, Emelia (female) [14] and Raman (male) [15], from two UK zoos. The individuals were a similar age at death: 2 years 8 months 11 days (Raman) and 2 years 9 months (Emelia). DNA was extracted from heart (Raman) and tongue (Emelia). From each sample, 5 µg of DNA was treated with NEBNext double-stranded DNA fragmentase (New England BioLabs, Ipswich, MA). Sequencing libraries were constructed using an Illumina TruSeq DNA sample preparation kit. The libraries were sequenced on an Illumina HiSeq 2000 (Illumina, San Diego, CA) instrument at Edinburgh Genomics (University of Edinburgh), resulting in 76 bp paired-end datasets.

Bioinformatics, data description and availability

The data are deposited in the European Nucleotide Archive (ENA) under accession [EMBL:ERP004241], and consist of 97.6 million (577_1_Emelia) and 53.8 million (577_2_Raman) paired 76 bp reads. Assuming a genome size of 3.1Gb, the sequence data described here represent 4.8X and 2.6X theoretical coverage respectively. The insert size for each library is approximately 350 bp. The reads have been aligned to the African elephant genome (Loxofr3.0, INSDC Assembly GCA_000001905.1, Jul 2009) using bwa [16] and converted to BAM by Samtools [17]. The vast majority of the Asian elephant sequence reads align to the African elephant genome: 94.66% (577_1_Emelia) and 93.82% (577_2_Raman). The percentage of bases in the *L. africana* assembly covered by at least one read from the Asian samples are 94.18% and 84.76% for 577_1_Emelia, and 577_2_Raman respectively. These coverage percentages were calculated using the mpileup

tool from Samtools [17] package (version 0.1.18) with the option `-C` set to 50.

Availability of supporting data

The datasets supporting the results of this article are available in the European Nucleotide Archive (ENA) under accession [EMBL:ERP004241], as well as the *GigaScience*, GigaDB repository [18].

Abbreviations

Bp: Base-pairs; EEHV: Elephant endotheliotropic herpesvirus; Gb: Gigabases; IUCN: International Union for Conservation of Nature; kg: Kilograms; m: Metres.

Competing interests

The authors declare they have no competing interests.

Authors' contributions

AD conceived the study, provided samples and helped write the paper; CR carried out bioinformatics analysis and helped write the paper; MW conceived the study, carried out bioinformatics analysis and helped write the paper. All authors read and approved the final manuscript.

Acknowledgements

The work was enabled by funding from the Biotechnology and Biological Sciences Research Council including Institute Strategic Programme and National Capability grants (BBSRC; BBS/E/D/20310000, BB/J004243/1). DNA sequencing was provided by Edinburgh Genomics (<http://genomics.ed.ac.uk>). The authors thank the staff at Chester and Whipsnade Zoos, UK for the elephant tissues and Willem Schaftenaar, Rotterdam zoo (The Netherlands) for his advice on the manuscript.

Author details

¹Virology Department, Animal Health and Veterinary Laboratories Agency, Weybridge, KT15 3NB Addlestone, Surrey, UK. ²The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, EH25 9RG Easter Bush, UK. ³Edinburgh Genomics, University of Edinburgh, EH25 9RG Easter Bush, UK.

Received: 7 May 2014 Accepted: 8 July 2014

Published: 10 July 2014

References

1. Shoshani J, Eisenberg JF: *Elephas maximus*. *Mamm Species* 1982, (182):1–8.
2. Laursen L, Bekoff M: *Loxodonta africana*. *Mamm Species* 1978, (92):1–8.
3. Roca AL, Georgiadis N, Pecon-Slatery J, O'Brien SJ: **Genetic evidence for two species of elephant in Africa**. *Science* 2001, **293**(5534):1473–1477.
4. Sukumar R: *The living elephants: evolutionary ecology, behavior, and conservation*. Oxford: Oxford University Press; 2003.
5. van Sonsbeek GR, van der Kolk JH, van Leeuwen JP, Everts H, Marais J, Schaftenaar W: **Effect of calcium and cholecalciferol supplementation on several parameters of calcium status in plasma and urine of captive Asian (*Elephas maximus*) and African elephants (*Loxodonta africana*)**. *J Zoo Wildl Med* 2013, **44**(3):529–540.
6. Ossent P, Guscelli F, Metzler AE, Lang EM, Rubel A, Hauser B: **Acute and fatal herpesvirus infection in a young Asian elephant (*Elephas maximus*)**. *Vet Pathol* 1990, **27**(2):131–133.
7. Shoshani J: **Understanding proboscidean evolution: a formidable task**. *Trends Ecol Evol* 1998, **13**(12):480–487.
8. Hildebrandt TB, GÖritz F, Hermes R, Reid C, Dehnhard M, Brown JL: **Aspects of the reproductive biology and breeding management of Asian and African elephants *Elephas maximus* and *Loxodonta africana***. *International Zoo Yearbook* 2006, **40**(1):20–40.
9. **The IUCN Red List of Threatened Species**. [<http://www.iucnredlist.org/>]
10. **Elephant genome project**. [<https://www.broadinstitute.org/scientific-community/science/projects/mammals-models/elephant/elephant-genome-project>]
11. Flicek P, Amodè MR, Barrell D, Beal K, Billis K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fitzgerald S, Gil L, Giron CG, Gordon L, Hourlier T,

- Hunt S, Johnson N, Juettemann T, Kahari AK, Keenan S, Kulesha E, Martin FJ, Maurel T, McLaren WM, Murphy DN, Nag R, Overduin B, Pignatelli M, Pritchard B, Pritchard E, Riat HS, *et al*: **Ensembl** 2014. *Nucleic Acids Res* 2014, **42**(Database issue):D749–755.
12. Wilkie GS, Davison AJ, Watson M, Kerr K, Sanderson S, Bouts T, Steinbach F, Dastjerdi A: **Complete genome sequences of elephant endotheliotropic herpesviruses 1A and 1B determined directly from fatal cases.** *J Virol* 2013, **87**(12):6700–6712.
 13. Cappellini E, Gentry A, Palkopoulou E, Ishida Y, Cram D, Roos A-M, Watson M, Johansson US, Fernholm B, Agnelli P, Barbagli F, Littlewood DTJ, Kelstrup CD, Olsen JV, Lister AM, Roca AL, Dalén L, Gilbert MTP: **Resolution of the type material of the Asian elephant, *Elephas maximus* Linnaeus, 1758 (Proboscidea, Elephantidae).** *Zool J Linn Soc* 2014, **170**(1):222–232.
 14. **Elephant database (Emelia).** [http://www.elephant.se/database2.php?elephant_id=779]
 15. **Elephant database (Raman).** [http://www.elephant.se/database2.php?elephant_id=581]
 16. Li H, Durbin R: **Fast and accurate short read alignment with Burrows-Wheeler transform.** *Bioinformatics* 2009, **25**(14):1754–1760.
 17. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R: **The Sequence Alignment/Map format and SAMtools.** *Bioinformatics* 2009, **25**(16):2078–2079.
 18. Dastjerdi A, Robert C, Watson M: **Supporting data for "Low coverage sequencing of two Asian elephant (*Elephas maximus*) genomes".** *GigaSci* 2014, Database <http://dx.doi.org/10.5524/100095>.

doi:10.1186/2047-217X-3-12

Cite this article as: Dastjerdi *et al.*: Low coverage sequencing of two Asian elephant (*Elephas maximus*) genomes. *GigaScience* 2014 **3**:12.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

