Genome analysis reveals insights into physiology and longevity of the Brandt's bat *Myotis brandtii*

Inge Seim[1,2,*], Xiaodong Fang[3,4,*], Zhiqiang Xiong[3], Alexey V. Lobanov[1], Zhiyong Huang[3], Siming Ma[1], Yue Feng[3], Anton A. Turanov[1], Yabing Zhu[3], Tobias L. Lenz[1], Maxim V. Gerashchenko[1,5], Dingding Fan[3], Sun Hee Yim[1], Xiaoming Yao[3], Daniel Jordan[1], Yingqi Xiong[3], Yong Ma[3], Andrey N. Lyapunov[6], Guanxing Chen[3], Oksana I. Kulakova[7], Yudong Sun[3], Sang-Goo Lee[2], Roderick T. Bronson[8], Alexey A. Moskalev[7,9,10], Shamil R. Sunyaev[1], Guojie Zhang[3], Anders Krogh[4], Jun Wang[3,4,11] & Vadim N. Gladyshev[1,2]

Bats account for one-fifth of mammalian species, are the only mammals with powered flight, and are among the few animals that echolocate. The insect-eating Brandt's bat (*Myotis brandtii*) is the longest-lived bat species known to date (lifespan exceeds 40 years) and, at 4–8 g adult body weight, is the most extreme mammal with regard to disparity between body mass and longevity. Here we report sequencing and analysis of the Brandt's bat genome and transcriptome, which suggest adaptations consistent with echolocation and hibernation, as well as altered metabolism, reproduction and visual function. Unique sequence changes in growth hormone and insulin-like growth factor 1 receptors are also observed. The data suggest that an altered growth hormone/insulin-like growth factor 1 axis, which may be common to other long-lived bat species, together with adaptations such as hibernation and low reproductive rate, contribute to the exceptional lifespan of the Brandt's bat.

[1] Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. [2] Department of Bioinspired Science, Ewha Womans University, Seoul 120-750, South Korea. [3] BGI-Shenzhen, Shenzhen 518083, China. [4] Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen DK-2200, Denmark. [5] University of Nebraska, Lincoln, Nebraska 68588, USA. [6] Kirov State Center for Distance Education of Children, Kirov 610006, Russia. [7] Institute of Biology, Komi Science Center, Russian Academy of Sciences, Syktyvkar 167982, Russia. [8] Rodent Histopathology Laboratory, Harvard Medical School, Boston, Massachusetts 02115, USA. [9] Syktyvkar State University, Syktyvkar 167001, Russia. [10] Moscow Institute of Physics and Technology, Dolgoprudny 141700, Russia. [11] Department of Biology, University of Copenhagen, Copenhagen DK-2200, Denmark. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.W. (email: wangj@genomics.org.cn) or to V.N.G. (email: vgladyshev@rics.bwh.harvard.edu).

Bats are among the most successful groups of animals. They account for ~20% of mammalian species, are the only mammals that have evolved powered flight, and are among the few animals that echolocate[1]. Bats are also among the smallest of mammals, but are unusually long-lived, thus challenging the observed positive correlation between body mass and maximum lifespan[2]. The Brandt's bat (*Myotis brandtii*) holds the record with regard to lifespan among the bats. Its reported maximal lifespan of at least 41 years also makes it the most extreme mammal with regard to disparity between body mass and longevity[3]. These animals reside in temperate areas throughout Europe and Asia, resting during the day, feeding at night and hibernating during winter. As of yet, however, these highly specialized animals have escaped detailed genome characterization.

In this study, we present a genome and transcriptome of the Brandt's bat. Their analyses revealed adaptations that this exceptionally long-lived animal developed for echolocation, hibernation, reproduction and visual function. We identified unique sequence changes in the growth hormone (GH) receptor (GHR) and insulin-like growth factor 1 (IGF1) receptor (IGF1R), which appear to be common to other long-lived bat species and contribute to the small body size and long lifespan of the Brandt's bat.

Results

Genome sequence assembly and annotation. A whole-genome shotgun strategy was applied to sequence the genome of an adult male Brandt's bat (*M. brandtii*) from the Obvalnaya cave in Russia (Fig. 1 and Supplementary Fig. S1; Supplementary Note 1). We also sequenced liver, kidney and brain transcriptomes of hibernating and summer-active *M. brandtii*. Approximately 156 Gb (or 78×) genome data including 101 Gb (or 51×) short insert size reads were generated, resulting in a high-quality assembly with scaffold N50 of 3.1 Mb and contig N50 of 21 kb (Table 1, Supplementary Tables S1–S3, Supplementary Figs S2,S3). The prevailing theory is that bats and birds have smaller genomes than other vertebrates due to metabolic constraints on cell size and, therefore, genome size[4,5]. Consistent with previous findings[6], the *M. brandtii* genome size was smaller (~2 Gb) than other mammals (Supplementary Fig. S2). Nucleotide diversity (per nucleotide heterozygosity) was estimated at 0.4%, higher than in another long-lived mammal, the naked mole rat[7] (Supplementary Table S4). Known transposon-derived repeats accounted for 22% of the genome, which is lower than in related mammals and may contribute to the small size of the *M. brandtii* genome (Supplementary Table S5). The predicted gene set included 22,256 protein-coding genes (Supplementary Table S6 and Supplementary Fig. S4), of which 21,219 (95.3%) genes were transcribed (based on the RNA-seq data) and 20,744 (93.2%) genes could be functionally annotated by homology approaches (Supplementary Table S7).

Genome evolution. The order of bats (Chiroptera) consists of Yangochiroptera (includes the Brandt's bat and other echolocating bats) and Yinpterochiroptera (includes megabats, as well as some echolocating bat species), and belongs to the group of laurasiatherian mammals (together with Carnivora including dog and cat, Cetartiodactyla including cow and dolphin, Perissodactyla including horse, Eulipotyphla and Pholidota)[1]. To gain insight into the evolutionary relationships of bats, 2,654 single-copy orthologous genes were obtained from *M. brandtii* and eight related mammals and used to construct a phylogenetic tree (Supplementary Fig. S5). High bootstrap evidence supported bats as the sister taxon of the horse (*Equus caballus*), although the remote estimated divergence date of ~81.7 million years ago is close to the time of mammalian radiation (Fig. 2). This inference is also supported by comparing orthologous proteins (Supplementary Fig. S6 and Supplementary Table S8) and by recent findings at the transcript level[8,9]. Interestingly, both *M. brandtii* and *E. caballus* are long-lived species, but differ dramatically in body size. Compared with the common ancestor of bats and horses, *M. brandtii* has 67 significantly expanded and 44 significantly contracted gene families (Fig. 2 and Supplementary Fig. S7). Many expanded gene families in the Brandt's bat are involved in the immune response, for example, IPR003599 immunoglobulin subtype and IPR003596 immunoglobulin V-set subgroup (Supplementary Data 1), which is consistent with previous studies on bats in general.

Table 1 | Global statistics of the *M. brandtii* genome.

	Insert size (bp)	Total data (Gb)	Sequence coverage (X)
(a) Sequencing			
Paired-End library	170–800 bp	101.24	50.6
	$2–40 \times 10^3$	54.48	27.2
	Total	155.72	77.8
	N50 (kb)	**Longest (kb)**	**Size (Gb)**
(b) Assembly			
Contig	21.1	217.9	2.05
Scaffold	3,111.4	22,122.7	2.18
	Number	**Total length (Mb)**	**Percentage of genome**
(c) Annotation			
Repeats	2,350,296	483.06	22.17
Genes	22,256	691.61	31.75
CDS	178,548	31.77	1.46
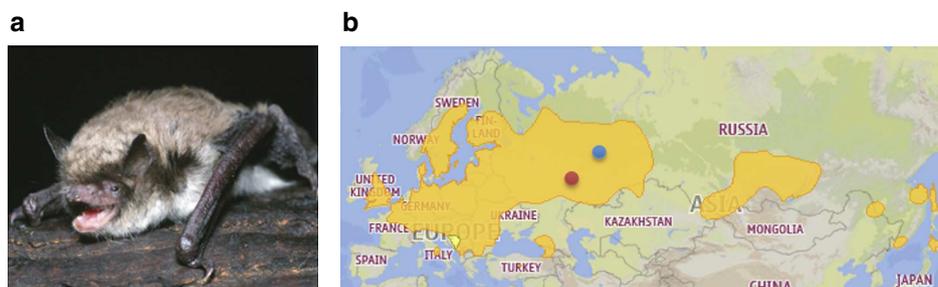
Figure 1 | Species map of *M. brandtii*. (a) *M. brandtii*. **(b)** Occurrence of *M. brandtii* and areas of sample collection. Coloured dots indicate the location of sampled hibernating (December and April; shown in blue) and summer-active (June; shown in red) animals.
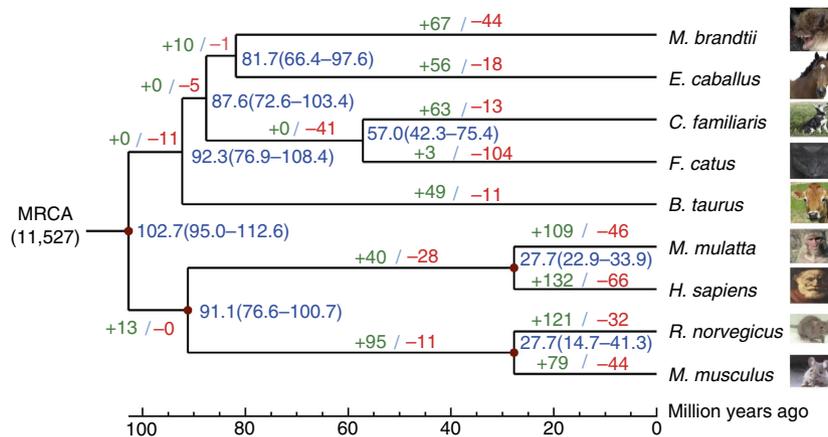
Figure 2 | Relationship of *M. brandtii* to other mammals. Consensus phylogenetic tree shows the relationship among nine mammals based on 2,654 single-copy genes. The divergence times for all nodes were estimated based on the four red nodes with fossil records as calibration times and marked in each node with error range. Gene family expansion events are marked in green, and gene family contraction events in red. MRCA refers to the gene family number of the most recent common ancestor.

We found that many olfactory receptor gene families underwent contraction in the *M. brandtii* lineage (Supplementary Data 2), which may reflect that *M. brandtii* depends primarily on echolocation and vision for sensory perception. Indeed, the olfactory system is aberrant in many bats, including the related *M. lucifugus*[10,11].

Analysis of syntenic regions identified 349 gained and 98 lost *M. brandtii* genes (Supplementary Tables S9,S10), and at least 75.3% of the gained genes showed evidence of transcription. Some of the gained genes were obtained by gene duplication. For example, we detected five copies of *FBXO31* based on synteny analysis, whereas an additional 52 *FBXO31* genes could be identified in the *M. brandtii* genome. This protein mediates DNA damage-induced growth arrest by targeting CDK1 for ubiquitin-mediated degradation. Several studies have reported a relatively recent wave of DNA transposon activity in *M. lucifugus* (but not megabats or other mammals), which may have contributed to the *M. brandtii* gene duplication[12]. We also identified 194 pseudogenes, containing 142 frameshift and 72 premature termination events (Supplementary Data 3). A salient example is the gulonolactone (L-) oxidase (*GULOP*) gene. As is the case for primates, the ability to synthesize vitamin C is lost in bats[13]. In agreement with this observation and a recent study that examined the related *M. lucifugus*[14], we found that this gene is inactivated in *M. brandtii* (Supplementary Fig. S8).

Echolocation. Amino-acid residues that are uniquely altered in the bat lineage alone or in bats and species with similar phenotypes, such as powered flight in bats and birds and echolocation in bats and dolphins, can reveal clues about functional adaptations. We examined amino-acid changes at evolutionary conserved residues of orthologs in 48 vertebrate genomes (Supplementary Table S11) and identified four proteins with unique amino-acid changes that are common to the echolocating mammals the Brandt's bat and the Atlantic bottlenose dolphin, *Tursiops truncatus* (Table 2). Interestingly, two of these genes (*SLC45A2* and *RGS7BP*) are expressed in the inner ear. *SLC45A2* (Supplementary Fig. S9) is a transporter protein that has an important role in melanin synthesis[15] and was recently reported to be one of the most abundantly expressed genes in the inner ear[16]. Melanin is highly sound-absorbing[17] and may function in the inner ear to protect the cochlea from aging-related hearing loss and sound stressors in much the same manner as melanin in

the skin protects against ultraviolet light[18,19,20]. A recent study found that *RGS7BP* (Supplementary Fig. S10) is expressed in vestibular ganglion neurons in the inner ear, which are important for spatial orientation[21]. These findings add to the list of independent gene loci with putative functions in the inner ear of echolocating bats and dolphins, such as prestin (also known as *SLC26A5*)[22], *TMC1*[23], *DFNB59*[23], *SLC45A2* and *RGS7BP*, implicating them in echolocation.

Among the positively selected genes in the Brandt's bat, we found additional genes with potential roles in echolocation (Supplementary Table S12): *TRPV5* (TRPC channels have a role in mechanosensation, including hearing and proprioception) and *NOX3* (highly expressed in the inner ear and involved in the biogenesis of otoconia/otolith—the crystalline structures for perception of gravity). GO categories significantly enriched (Benjamini–Hochberg false discovery rate (FDR) adjusted P-value < 0.05) in the set with evidence for positive selection in the Brandt's bat included genes associated with immunity and fertilization (Supplementary Table S13). Our analysis also revealed that 21 *M. brandtii* genes contained at least one unique amino-acid change (Supplementary Table S14). Notable among them were genes associated with visual function (*CNGB3*), reproduction (*FSHB*, *DDX4*, *HARS2*, *PRSS23* and *PHTF1*), Wnt signalling (*APC*), RNA interference (*ERI3*) and helicase activity (*DDX1*, *DDX4*, *HELB* and *VPS4B*).

Vision. The small eyes of echolocating bats are adapted for use beyond the range of echolocation[24], as well as for short-range navigation in concert with echolocation[25]. Molecular innovation of photopigment opsins, rhodopsin and downstream signalling components may, therefore, confer an advantage to echolocating bats at dawn and dusk (Supplementary Note 2). Indeed, we found changes in photoreceptor genes of the Brandt's bat suggestive of adaptation to dim-light conditions (Supplementary Fig. S11) and UV vision (Supplementary Fig. S12). *CNGB3* (cyclic nucleotide gated channel beta 3) codes for one of the two proteins necessary for sensory transduction of cone photoreceptors. Mutations in this gene result in impaired conversion of light captured by cone opsins to nerve impulses and *CNGB3* mutations have been associated with achromatopsia[26,27], where subjects can only see clearly at low light levels. Although their cone photoreceptor cells remain structurally intact, signalling is blunted and they are either totally or partially colour blind and display reduced visual

capacity in bright light (day-blindness). *M. brandtii*, *M. lucifugus* and *E. fuscus*, which roost in caves during the day, harbour a unique amino-acid change in the transmembrane domain S5 of CNGB3 (Supplementary Fig. S13), where a tyrosine highly conserved in vertebrates is replaced with cysteine. Interestingly, *M. lucifugus* avoids obstacles better under dim-light conditions such as dusk than under bright daylight[25,28]. We hypothesize that the CNGB3 amino-acid change is a loss-of-function mutation that result in day-blindness in strictly nocturnal, echolocating bats. It will be of future interest to obtain the genome or retinal transcriptome data from additional bat species.

Hibernation and reproduction. Hibernation is a hypometabolic state in mammals wherein they enter a state of torpor, typically in response to environmental conditions such as a lack of food (insects in the case of *M. brandtii*) and subzero temperatures[29]. To obtain insights into this adaptation, we examined gene expression in the liver of *M. brandtii* subjected to hibernation for 6 months (Supplementary Table S15 and Supplementary Data 4). This analysis revealed an enrichment (Benjamini–Hochberg FDR-adjusted P-value <0.05) of downregulated genes associated with RNA binding (GO:0003723) and structural integrity of the ribosome (GO:0003735) (Table 3). In addition, downregulated genes have roles in protein synthesis (multiple ribosomal proteins, *NACA*, and *PRPS1*; GO:0003723), as well as glycolysis (*ENO1* and *GAPDH*), splicing (*SNRPC*) and mitochondrial respiration (*NDUFA3*). Enriched gene categories upregulated during hibernation included digestion (metabolism; GO:0007586) and peptidase activity (Table 3). We also observed

genes involved in lipid metabolism (*PLA2G1B*, *CEL*, *UNQ599*, *G0S2*, *CYP4X1* and *CLPS*), coagulation (*CPB1* and *MBL1*), maintenance of oxidative stress (*OSGIN1* and *LYPD3*) and adaptation to starvation (*FGF21* and *ATG101*). Overall, these changes are consistent with the expected[29] shift from carbohydrate to lipid metabolism and the reduction in protein synthesis during the hypometabolic state of hibernation.

In *Myotis*, mating occurs right before or during hibernation, where males can mate with torpid females. The sperm is then stored in the female reproductive tract, with ovulation and fertilization delayed until spring[30]. Follicle-stimulating hormone (FSH) is a glycoprotein hormone consisting of a common α-subunit and a hormone-specific β-subunit. The hormone binds to its cognate G-protein-coupled receptor (FSH receptor; FSHR) in the testis and ovary and, together with luteinizing hormone (LH) and its receptor (LHR), is produced in the anterior pituitary and essential for reproduction. We found a unique amino-acid change (S89R) of potential functional significance in the mature FSH beta-subunit (FSHβ) peptide of hibernating bats (Fig. 3 and Supplementary Fig. S14). This change occurs within the hydrophobic receptor–ligand interface that determines receptor-binding specificity[31], and the replacement of a small amino acid with a large, charged one would be expected to disrupt the FSHβ–FSHR interaction. Interestingly, it has recently been hypothesized[31] that Ser89 may explain why LH cannot bind FSHR. LH, encoded by *LHB*, is a paralogous gene to *FSHB* and the LH beta-subunit (LHβ) is similar to FSHβ but harbours arginine at the site homologous to Ser89. Moreover, site-directed mutagenesis of residues 88–91 of FSHβ (DSDS) to the LHβ sequence (RRST) allows FSHβ to bind LHR[32]. A study of the

---

Table 2 | Proteins with unique residues in *Myotis* and dolphin.

Gene name	Symbol	*M. brandtii* protein	*M. lucifugus* protein	Human RefSeq	AA change	Function
Regulator of G-protein signalling 7 binding protein	RGS7BP	10035863	ENSG00000186479	NM_001029875	D114G	Regulator of GPCR signaling
Solute carrier family 45, member 2	SLC45A2	10029293	ENSG00000164175	NM_016180	Q344R	Membrane transporter
RNA-binding motif protein 46	RBM46	10031311	ENSG00000151962	NM_144979	T249A	Unknown
Structural maintenance of chromosomes 1B	SMC1B	10011712	ENSG00000077935	NM_148674	K1015R	Maintenance of chromosome integrity

AA, amino acid. AA changes correspond to residue locations in the translated human RefSeq mRNA entry (typically the longest RefSeq entry of each gene). *M. lucifugus* proteins correspond to Ensembl IDs. All proteins were manually verified. The term 'dolphin' refers to the Atlantic bottlenose dolphin.

Table 3 | Significant GO categories of hibernating *M. brandtii*.

GO categories	Description	Taxonomy	No. of genes (all)	No. of genes (in list)	P-value (hypergeometric)	Corrected P-value (Benjamini–Hochberg)
Lower during hibernation						
GO:0003723	RNA binding	MF	14	8	8.00E − 04	0.024
GO:0003735	Structural constituent of ribosome	MF	8	6	5.00E − 04	0.024
Higher during hibernation						
GO:0007586	Digestion	BP	9	7	2.00E − 04	0.0358
GO:0070011	Peptidase activity, acting on L-amino-acid peptides	MF	18	10	1.00E − 03	0.0365
GO:0008233	Peptidase activity	MF	18	10	1.00E − 03	0.0365
GO:0017171	Serine hydrolase activity	MF	9	6	3.80E − 03	0.0462
GO:0004252	Serine-type endopeptidase activity	MF	9	6	3.80E − 03	0.0462
GO:0008236	Serine-type peptidase activity	MF	9	6	3.80E − 03	0.0462
GO:0016787	Hydrolase activity	MF	36	15	2.00E − 03	0.0462
GO:0005576	Extracellular region	CC	43	18	3.00E − 04	0.015

BP, biological process; CC, cellular compartment; GO, gene ontology; MF, molecular function. GO categories over-represented among genes differentially expressed in the liver. Benjamini–Hochberg adjusted P-value <0.05.
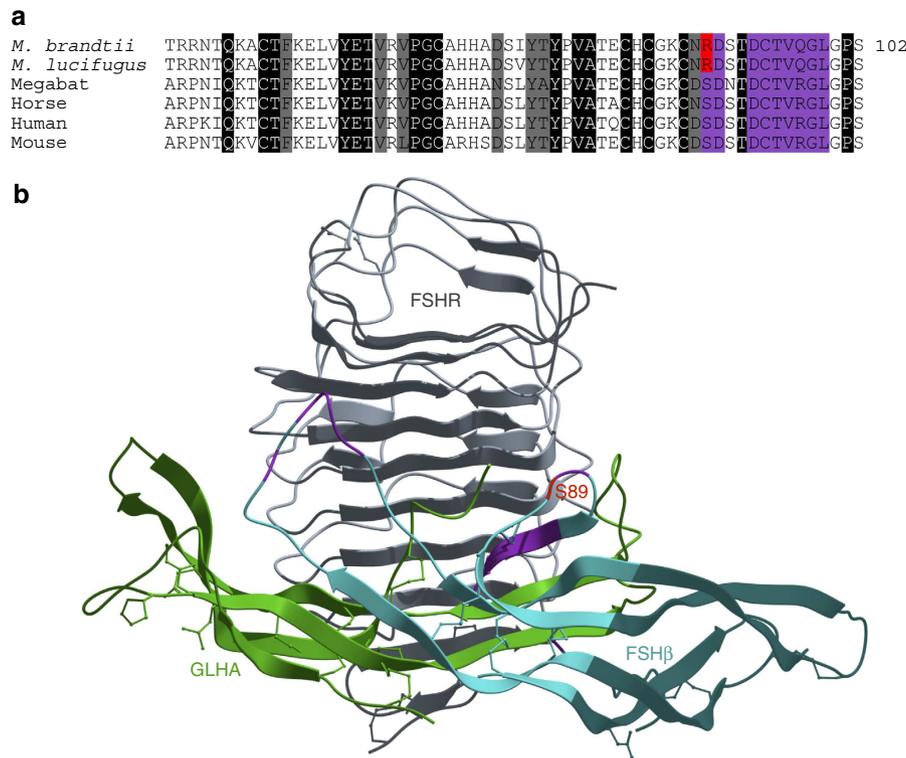
Figure 3 | FSHβ of the Brandt's bat harbours a radical amino-acid substitution in a conserved position. (**a**) Alignment of *FSHB*-encoded peptide sequences. Amino-acid numbering corresponds to the mature FSHβ peptide. (**b**) Structural model of FSHβ (cyan), the follicle-stimulating hormone receptor FSHR (grey) and the glycoprotein hormones α chain GLHA/FSHα (green) (PDB ID code 1xwd). In both subfigures, FSHβ residues buried at the FSHR receptor–ligand interface are highlighted in purple, while a radical amino-acid substitution within the interface in bats with delayed ovulation is shown in red.

Asiatic greater yellow house bat (*Scotophilus heathi*), which is in the same family as *M. brandtii*, suggests that increased responsiveness to LH could be associated with delayed ovulation[33]. Taken together, we hypothesize that an aberrant gonadotropin system, where FSH may have reduced affinity for FSHR and a gain-of-affinity for LHR, may contribute to the delayed ovulation observed in many bat species. Future studies using, for example, knock-in mice expressing a Ser89Arg mutation, should provide interesting insights into the phenotypic consequences of this amino-acid change.

Potential adaptations for extended lifespan in bats. Growth hormone (GH) insensitivity encompasses a range of genetic abnormalities of the GH/IGF1 axis. Mutations in the single-transmembrane receptor GHR, including the transmembrane domain-coding exon 8, have been shown to result in human Laron-type dwarfism (short-stature), a subtype of GH insensitivity[34]. Interestingly, GHR mutations or GHR dysfunctions have been associated with increased resistance to diabetes and cancer in humans and mice[35–37]. We found that Leu284 in the transmembrane domain of GHR, which is highly conserved in tetrapods, is deleted in *M. brandtii*, *M. lucifugus* and two other echolocating bat species, the big brown bat (*Eptesicus fuscus*) and the Mexican free-tailed bat (*Tadarida brasiliensis*) (Fig. 4a and Supplementary Fig. S15). In contrast, the echolocating Jamaican fruit bat (*Artibeus jamaicensis*) and the megabat (*Pteropus vampyrus*) do not harbour this deletion. In addition, the nearby Ile275 is uniquely replaced with methionine in *M. brandtii*, *M. lucifugus*, *E. fuscus* and *T. brasiliensis*. GH binding induces conformational changes thought to be mediated by the GHR transmembrane domain[38]. Furthermore, mutations in the

transmembrane domain of closely related receptors have been observed to alter their function[39]. The GHR transmembrane residue deletion is present in the Vespertilionoidea superfamily, which comprises the largest number of species in the Chiroptera suborder Yangochiroptera. This superfamily includes the Vespertilionidae family, commonly known as Vesper bats (the 4–8 g *M. brandtii*, the 6.8 g *M. lucifugus*, the 16.9 g *E. fuscus*)[40] and the Molossidae family (the 11.2 g *T. brasiliensis*)[40]. The GHR transmembrane residue deletion is not present in the Jamaican fruit bat (the 39.8 g *A. jamaicensis*[40]) in the family Phyllostomidae. These data suggest that the amino-acid deletion appeared after the divergence between Vespertilionoidea and the branch leading to New World leaf-nosed bats (Phyllostomidae) ∼50 million years ago[41].

We also found changes in the *M. brandtii* IGF1R sequence, another key aging-related gene that is tightly intertwined with GHR function. Experiments in *Caenorhabditis elegans* found that knockout of *daf-2*, a functional homologue of IGF1R and the insulin receptor, extends lifespan several-fold[42]. The single-transmembrane domain (amino acids 936–959) of IGF1R is highly conserved in mammals, including the megabat, but in *M. brandtii*, *M. lucifugus* and *E. fuscus* we observed multiple unique changes both in the transmembrane region (M954T) and immediately preceding (F933H) and following (L967M) it (Fig. 4b and Supplementary Fig. S16). The transmembrane region could not be retrieved from transcriptomic data of *T. brasiliensis* and *A. jamaicensis*. This domain is important for signal transduction in response to binding of IGF1, and both mutations within this region and deletion of amino acids immediately preceding it affect IGF1R function[43].

To gain further insights into the evolved GH/IGF1 axis, we compared gene expression between mouse and Brandt's bat livers
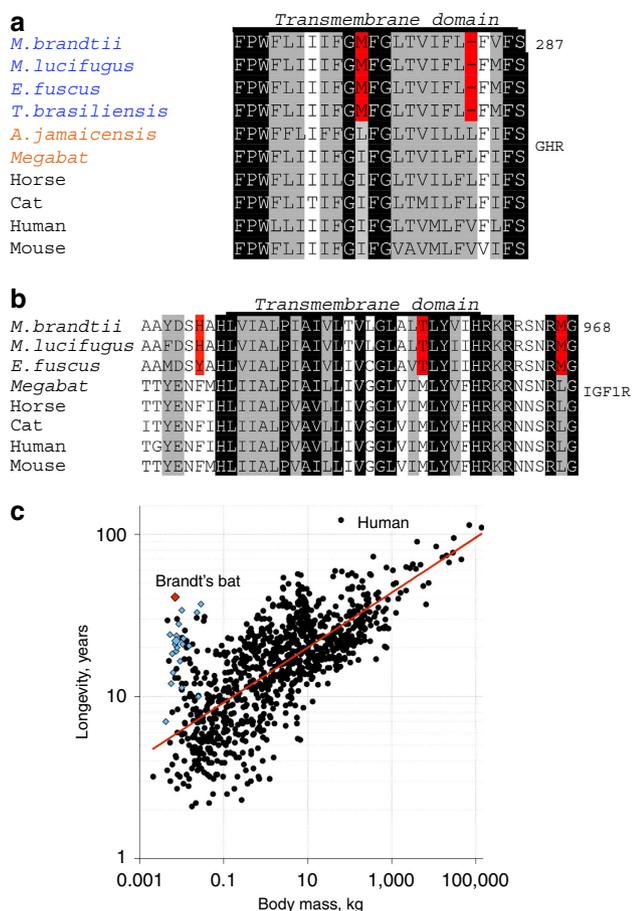
Figure 4 | Adaptations in the GH/IGF1 axis of the Brandt's bat and its much longer lifespan than expected for its body size. (**a**) Alignment of mammalian GHR protein sequences. An amino-acid deletion at position 284 and a methionine at position 275 unique to bats in the Vespertilionoidea superfamily are shown in red. Insect-eating bats are shown in blue and fruit-eating bats in orange. (**b**) Alignment of mammalian IGF1R protein sequences. Unique amino-acid replacements in bats in the Vespertilionoidea superfamily are highlighted in red. Identical residues and similar residues are shaded in black and grey, respectively. (**c**) Correlation between body mass and maximum lifespan in mammals. *Myotis* bats are shown as blue diamonds, and other mammals as grey circles. The Brandt's bat is indicated by a red diamond.

(Supplementary Note 3). If *M. brandtii GHR* is dysfunctional and bats display a phenotype corresponding to that of long-lived $GHR^{-/-}$ mice, one would expect insulin signalling in the liver of the Brandt's bat similar to the changes observed in the $GHR^{-/-}$ mouse liver[44]. Several insulin signalling-associated genes, including the major mediator of insulin action, forkhead box protein O1 (*FOXO1*), were expressed at a higher level in both *M. brandtii* and $GHR^{-/-}$ mice compared with wild-type mice (Supplementary Table S16). Dysregulated expression of insulin homeostasis genes in the liver of $GHR^{-/-}$ mice is caused by the ablation of GH signalling in these animals. $GHR^{-/-}$ mice exhibit an altered pancreatic structure, with small islets of Langerhans and decreased β-cell mass[45]. We examined the pancreas of *M. brandtii* (Supplementary Note 4) and observed a similar phenotype (Supplementary Fig. S17). We further compared *M. brandtii* gene expression to those of several long-lived mice[46] (Supplementary Note 5). *M. brandtii* exhibited the same gene expression changes for 11 out of 15 orthologous genes differentially expressed in long-lived mice compared with wild-type strains (Supplementary Table S17).

Finally, we assembled the mitochondrial genome of *M. brandtii* (Supplementary Fig. S18) and analysed the amino-acid content of its genes. It has been hypothesized that a relationship exists between reduced number of cysteine in mitochondrially encoded proteins and longevity[47,48]. We found that cysteine content of mitochondria-encoded proteins was much lower than that expected based on its body mass and slightly higher than expected based on lifespan (Supplementary Fig. S19). In this regard, the Brandt's bat is more similar to large-sized animals (Fig. 4c and Supplementary Fig. S19). However, additional factors (such as phylogenetic bias) may influence this correlation[49,50].

Discussion

The Brandt's bat possesses features attributed to increased longevity of bats[40]: it hibernates, has a low reproductive rate and roosts in caves (thus, avoiding predators and extreme weather conditions). Like larger mammals, this bat exhibits long reproductive cycles, long time to maturity and a long lifespan. The Brandt's bat produces a single pup, which may account for 10–15% of the adult body mass.

It is prudent to ask whether the changes in the GHR/IGF1 axis in the Brand's bat contribute to the animal's long lifespan, its small body size or both. The prevailing evidence is that early mammals, including the earliest horses, had a small body size[51,52]. Although the body size of most mammals and their extinct relatives has since increased, the body size of bats is much more diverse. Recent work supports the hypothesis that the body size of bats correlates with the echolocation and flight traits as well as dietary specializations, where a small body size is advantageous for echolocating animals that feed on small, fast-moving insects[53]. In contrast, a larger body size may be advantageous for bat species, some of which do not echolocate, that feed on bigger, slow-moving insects or fish or fruits. Echolocating insectivorous bats smaller than the Brandt's bat, such as the distantly related 2 g bumblebee bat *Craseonycteris thonglongyai* in the suborder Yinpterochiroptera[54], but with unknown maximum lifespans, exist. Although it is appreciated that other genes have been altered during the ~82 million years of bat evolution, we suggest that the changes observed in GHR and IGF1R contribute to the longevity and dwarfism-like phenotype of the Brandt's bat. Moreover, *M. brandtii* may mirror GHR dysfunction in mice and humans, and possibly insulin signalling in long-lived nematodes.

Taken together, we suggest that the combination of adaptations, such as hibernation, low reproductive rate, cave roosting and an altered GH/IGF1 axis, serve to extend the Brandt's bat lifespan. Our observations are pertinent to the Yangochiroptera suborder of echolocating bats, the second and most numerous suborder of bats (Chiroptera). It will be important for future studies to compare these findings to long-lived, small, echolocating bats of the suborder Yinpterochiroptera (which includes the non-echolocating megabat, *P. vampyrus*, examined in this study). Further analyses, including sequencing and lifespan studies (such as banding schemes) are required to fully determine the role of the GHR/IGF1 axis in the longevity and body size of bats.

Following submission of our Brandt's bat genome study, an article on the genomes of the related David's Myotis (*Myotis davidii*) and the megabat the black flying fox (*Pteropus alecto*) was published[55]. Although the other study examined different aspects of bat biology, the availability of two additional bat genomes should greatly help in further analyses of longevity and other traits reported in our study.

In summary, genome and transcriptome analyses of a mammal with a most exceptional lifespan provide clues to its evolution and several unique traits. The generated data sets can be further

mined and the identified genes characterized, helping to explain the basis for powered flight, echolocation, hibernation, small body size and most importantly, the basis for the remarkable longevity of this animal, and bats in general.

Methods

Animal sampling. Male *M. brandtii* bats were collected at a hibernation roost in the Perm Krai, Russia (GPS coordinates N 58 50.089 E 57 36.716) in December 2011 and April 2012, and at a summer roost in Kirov Oblast, Russia (GPS coordinates N 57 25.000 E 48 54.500) in June 2012 (Fig. 1). Animals were killed on-site and tissues frozen in liquid nitrogen. The sequenced individual bat was a hibernating adult male from the Obvalnaya cave in the Perm Krai. Liver, kidney and brain transcriptomes were from both hibernating and summer-active adult males. All procedures were carried out in accordance with guidelines of the Animal Ethics Committee of the Russian Academy of Sciences.

Genome sequencing and annotation. The *M. brandtii* genome and transcriptomes were sequenced on the Illumina HiSeq 2000 platform. The genome was assembled using the *de Bruijn* graph-based assembler SOAP*denovo*[56]. Full description and associated references for sequencing, assembly and gene annotation are provided as Supplementary Methods.

Identification of proteins with unique amino-acid changes. An in-house Perl pipeline was used to identify orthologous protein sites where types/groups of residues are shared by Brandt's and other species of interest. In addition to the Brandt's bat, 47 vertebrate species were examined. Briefly, two Perl scripts were employed: the first script identified amino-acid residues conserved in all sequences except the species of interest (for example, a valine in *M. brandtii*, and an alanine in all other organisms). As genome-wide comparisons of radical amino-acid changes may be more likely to reflect functional diversification, a second Perl script was employed to scan orthologous protein sites where types/groups of residues were shared by *Myotis* alone or with other species of interest. It is generally assumed that radical amino-acid changes are much more likely to reflect functional diversification, and the second script revealed if *M. brandtii* and *M. lucifugus* (plus other species of interest) harbour residue changes in four groups of amino acids: acidic (ED), basic (KHR), cysteine (C) and 'other' (STYNQGAVLIFPMW). For example, two species harbour K and H, although the other organisms contain either acidic, cysteine or 'other' residues at a particular site. Candidates were manually examined for alignment quality and conservation of the region with amino-acid changes. Raw *M. brandtii* genome and transcriptome data were examined to further exclude gene misprediction and misannotations. See Supplementary Methods for further details.

Identification of genes under positive selection. To detect positive selection, we examined 1:1 orthologous protein-coding sequence alignments of Brandt's bat, the megabat *P. vampyrus*, dog, horse, cow and the dolphin *Tursiops truncatus*. Briefly, the strict branch-site model in the CodeML program of the PAML (phylogenetic analysis using maximum likelihood) package (v4.6; settings: model = 2, NS sites = 2)[57] was employed. The branch-site test for positive selection assumes that the branches in the phylogenetic tree can be divided into foreground branches (where codons are under positive selection) and background branches (where positive selection is absent)[57]. Model A1 (H_0; CodeML settings: fix_omega = 1, omega = 1), where a codon may evolve neutrally or under purifying selection was compared with model A, where codons on the branch of interest can be under positive selection (H_A; CodeML settings: fix_omega = 0, omega = 1). Likelihood ratio test *P*-values were computed assuming that the null distribution was a χ^2-distribution with 1 degree of freedom and adjusted for multiple testing with a FDR threshold of 0.05. If significant, codon sites with a Bayes empirical Bayes probability > 95% were considered to be under positive selection. We manually examined a subset of PAML results significant at a 1% FDR. Large, divergent protein families and proteins with coiled-coil domains were omitted from the analysis. See Supplementary Methods for greater detail.

Transcriptome sequencing and analyses. Total RNA was isolated from livers, kidneys and brains of animals sampled in December 2011 (hibernated for 2 months), April 2012 (hibernated for 6 months) and June 2012 (summer-active animals). RNA sequencing libraries were constructed using the Illumina mRNA-Seq Prep Kit. Briefly, oligo(dT) magnetic beads were used to purify mRNA molecules. The mRNA was further fragmented and randomly primed during the first-strand synthesis by reverse transcription. This procedure was followed by second-strand synthesis with DNA polymerase I to create double-stranded cDNA fragments. The double-stranded cDNA was subjected to end repair by Klenow and T4 DNA polymerases and A-tailed by Klenow lacking exonuclease activity. Ligation to Illumina Paired-End Sequencing adapters, size selection by gel electrophoresis and PCR amplification completed the library preparation. The 200 bp paired-end libraries were sequenced using Illumina HiSeq 2000 (90 bp at each end).

Transcriptome reads were mapped by TopHat[58], and mapped reads were analysed with in-house Perl scripts. Differentially expressed genes were calculated using the GFOLD package[59]. Briefly, GFOLD (generalized fold change) assigns reliable statistics for expression changes based on the posterior distribution of log fold change. Read counts were obtained from TopHat-generated BAM files using SAMtools[60] and differentially expressed genes identified by comparing the gene expression in the liver of hibernating and summer-active bats, exactly as outlined in the GFOLD manual.

References

1. Simmons, N. B. In *Mammal species of the World: A Taxonomic and Geographic Reference* (Johns Hopkins University Press, 2005).
2. Kirkwood, T. B. & Austad, S. N. Why do we age? *Nature* **408,** 233–238 (2000).
3. Podlutsky, A. J., Khritankov, A. M., Ovodov, N. D. & Austad, S. N. A new field record for bat longevity. *J. Gerontol. A Biol. Sci. Med. Sci.* **60,** 1366–1368 (2005).
4. Hughes, A. L. & Hughes, M. K. Small genomes for better flyers. *Nature.* **377,** 391 (1995).
5. Zhang, Q. & Edwards, S. V. The evolution of intron size in amniotes: a role for powered flight? *Genome. Biol. Evol.* **4,** 1033–1043 (2012).
6. Smith, J. D. & Gregory, T. R. The genome sizes of megabats (Chiroptera: Pteropodidae) are remarkably constrained. *Biol. Lett.* **5,** 347–351 (2009).
7. Kim, E. B. *et al.* Genome sequencing reveals insights into physiology and longevity of the naked mole rat. *Nature* **479,** 223–227 (2011).
8. Papenfuss, A. T. *et al.* The immune gene repertoire of an important viral reservoir, the Australian black flying fox. *BMC Genomics* **13,** 261 (2012).
9. Shaw, T. I. *et al.* Transcriptome sequencing and annotation for the Jamaican Fruit Bat (Artibeus jamaicensis). *PLoS ONE* **7,** e48472 (2012).
10. Zhao, H., Xu, D., Zhang, S. & Zhang, J. Widespread losses of vomeronasal signal transduction in bats. *Mol. Biol. Evol.* **28,** 7–12 (2011).
11. Wible, J. R. & Bhatnagar, K. P. Chiropteran vomeronasal complex and the interfamilial relationships of bats. *J. Mamm. Evol.* **3,** 285–314 (1996).
12. Pritham, E. J. & Feschotte, C. Massive amplification of rolling-circle transposons in the lineage of the bat Myotis lucifugus. *Proc. Natl Acad. Sci. USA* **104,** 1895–1900 (2007).
13. Drouin, G., Godin, J. R. & Page, B. The genetics of vitamin C loss in vertebrates. *Curr. Genomics* **12,** 371–378 (2011).
14. Hiller, M. *et al.* "Forward genomics" approach links genotype to phenotype using independent phenotypic losses among related species. *Cell Rep.* **2,** 817–823 (2012).
15. Fukamachi, S., Shimada, A. & Shima, A. Mutations in the gene encoding B, a novel transporter protein, reduce melanin content in medaka. *Nat. Genet.* **28,** 381–385 (2001).
16. Peters, L. M. *et al.* Signatures from tissue-specific MPSS libraries identify transcripts preferentially expressed in the mouse inner ear. *Genomics* **89,** 197–206 (2007).
17. McGinness, J., Corry, P. & Proctor, P. Amorphous semiconductor switching in melanins. *Science* **183,** 853–855 (1974).
18. Conlee, J. W., Abdul-Baqi, K. J., McCandless, G. A. & Creel, D. J. Differential susceptibility to noise-induced permanent threshold shift between albino and pigmented guinea pigs. *Hear. Res.* **23,** 81–91 (1986).
19. Hayashi, H. *et al.* Comparison of the quantity of cochlear melanin in young and old C57BL/6 mice. *Arch. Otolaryngol. Head Neck Surg.* **133,** 151–154 (2007).
20. Murillo-Cuesta, S. *et al.* Melanin precursors prevent premature age-related and noise-induced hearing loss in albino mice. *Pigment. Cell. Melanoma. Res.* **23,** 72–83 (2010).
21. Lu, C. C., Appler, J. M., Houseman, E. A. & Goodrich, L. V. Developmental profiling of spiral ganglion neurons reveals insights into auditory circuit assembly. *J. Neurosci.* **31,** 10903–10918 (2011).
22. Liu, Y. *et al.* Convergent sequence evolution between echolocating bats and dolphins. *Curr. Biol.* **20,** R53–R54 (2010).
23. Davies, K. T., Cotton, J. A., Kirwan, J. D., Teeling, E. C. & Rossiter, S. J. Parallel signatures of sequence evolution among hearing genes in echolocating mammals: an emerging model of genetic convergence. *Heredity* **108,** 480–489 (2012).
24. Suthers, R. A. & Wallis, N. E. Optics of the eyes of echolocating bats. *Vision Res.* **10,** 1165–1173 (1970).
25. Orbach, D. N. & Fenton, B. Vision impairs the abilities of bats to avoid colliding with stationary obstacles. *PLoS One* **5,** e13912 (2010).
26. Sundin, O. H. *et al.* Genetic basis of total colourblindness among the Pingelapese islanders. *Nat. Genet.* **25,** 289–293 (2000).
27. Peng, C., Rich, E. D. & Varnum, M. D. Achromatopsia-associated mutation in the human cone photoreceptor cyclic nucleotide-gated channel CNGB3 subunit alters the ligand sensitivity and pore properties of heteromeric channels. *J. Biol. Chem.* **278,** 34533–34540 (2003).
28. Bradbury, J. W. & Nottebohm, F. The use of vision by the little brown bat, Myotis lucifugus, under controlled conditions. *Anim. Behav.* **17,** 480–485 (1969).
29. Storey, K. B. Mammalian hibernation. Transcriptional and translational controls. *Adv. Exp. Med. Biol.* **543,** 21–38 (2003).

30. Thomas, D. W., Fenton, F. & Barclay, R. M. R. Social behavior of the little brown bat, *Myotis lucifugus*: I. Mating behavior. *Behav. Ecol. Sociobiol.* **6,** 129–136 (1979).

31. Fan, Q. R. & Hendrickson, W. A. Structure of human follicle-stimulating hormone in complex with its receptor. *Nature* **433,** 269–277 (2005).

32. Dias, J. A., Zhang, Y. & Liu, X. Receptor binding and functional properties of chimeric human follitropin prepared by an exchange between a small hydrophilic intercysteine loop of human follitropin and human lutropin. *J. Biol. Chem.* **269,** 25289–25294 (1994).

33. Krishna, A. & Abhilasha Mechanism of delayed ovulation in a vespertilionid bat, scotophilus heathi: role of gonadotropin, insulin, and insulin-like growth factor-1. *Physiol. Biochem. Zool.* **73,** 523–529 (2000).

34. David, A. *et al.* Evidence for a continuum of genetic, phenotypic, and biochemical abnormalities in children with growth hormone insensitivity. *Endocr. Rev.* **32,** 472–497 (2011).

35. Coschigano, K. T., Clemmons, D., Bellush, L. L. & Kopchick, J. J. Assessment of growth parameters and life span of GHR/BP gene-disrupted mice. *Endocrinology* **141,** 2608–2613 (2000).

36. Guevara-Aguirre, J. *et al.* Growth hormone receptor deficiency is associated with a major reduction in pro-aging signalling, cancer and diabetes in humans. *Sci. Transl. Med.* **3,** 70ra13 (2011).

37. Ikeno, Y. *et al.* Reduced incidence and delayed occurrence of fatal neoplastic diseases in growth hormone receptor/binding protein knockout mice. *J. Gerontol. A Biol. Sci. Med. Sci.* **64,** 522–529 (2009).

38. Brown, R. J. *et al.* Model for growth hormone receptor activation based on subunit rotation within a receptor dimer. *Nat. Struct. Mol. Biol.* **12,** 814–821 (2005).

39. Brooks, A. J. & Waters, M. J. The growth hormone receptor: mechanism of activation and clinical implications. *Nat. Rev. Endocrinol.* **6,** 515–525 (2010).

40. Wilkinson, G. S. & South, J. M. Life history, ecology and longevity in bats. *Aging Cell.* **1,** 124–131 (2002).

41. Teeling, E. C. *et al.* A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science* **307,** 580–584 (2005).

42. Kenyon, C., Chang, J., Gensch, E., Rudner, A. & Tabtiang, R. A. C. elegans mutant that lives twice as long as wild type. *Nature* **366,** 461–464 (1993).

43. Takahashi, K., Yonezawa, K. & Nishimoto, I. Insulin-like growth factor I receptor activated by a transmembrane mutation. *J. Biol. Chem.* **270,** 19041–19045 (1995).

44. List, E. O. *et al.* Endocrine parameters and phenotypes of the growth hormone receptor gene disrupted (GHR-/-) mouse. *Endocr. Rev.* **32,** 356–386 (2011).

45. Liu, J. L. *et al.* Disruption of growth hormone receptor gene causes diminished pancreatic islet size and increased insulin sensitivity in mice. *Am. J. Physiol. Endocrinol. Metab.* **287,** E405–E413 (2004).

46. Swindell, W. R. Gene expression profiling of long-lived dwarf mice: longevity-associated genes and relationships with diet, gender and aging. *BMC Genomics* **8,** 353 (2007).

47. Kitazoe, Y. *et al.* Stability of mitochondrial membrane proteins in terrestrial vertebrates predicts aerobic capacity and longevity. *Genome Biol. Evol.* **3,** 1233–1244 (2011).

48. Moosmann, B. & Behl, C. Mitochondrially encoded cysteine predicts animal lifespan. *Aging Cell.* **7,** 32–46 (2008).

49. Aledo, J. C. *et al.* Mitochondrially encoded methionine is inversely related to longevity in mammals. *Aging Cell.* **10,** 198–207 (2011).

50. Aledo, J. C., Valverde, H. & de Magalhaes, J. P. Mutational bias plays an important role in shaping longevity-related amino acid content in mammalian mtDNA-encoded proteins. *J. Mol. Evol.* **74,** 332–341 (2012).

51. Smith, F. A. *et al.* The evolution of maximum body size of terrestrial mammals. *Science* **330,** 1216–1219 (2010).

52. Secord, R. *et al.* Evolution of the earliest horses driven by climate change in the paleocene-eocene thermal maximum. *Science* **335,** 959–962 (2012).

53. Giannini, N. P., Gunnel, G., Habersetzer, J. & Simmons, N. B. *In Evolutionary History of Bats: Fossils, Molecules, and Morphology*. Chapter 16 (Cambridge University Press, 2012).

54. Puechmaille, S. J. *et al.* Population size, distribution, threats and conservation status of two endangered bat species: Craseonycteris thonglongyai and Hipposideros turpis. *Endang. Species Res.* **8,** 15–23 (2009).

55. Zhang, G. *et al.* Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. *Science* **339,** 456–460 (2013).

56. Li, R. *et al.* De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* **20,** 265–272 (2010).

57. Zhang, J., Nielsen, R. & Yang, Z. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol. Biol. Evol.* **22,** 2472–2479 (2005).

58. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25,** 1105–1111 (2009).

59. Feng, J. *et al.* GFOLD: a generalized fold change for ranking differentially expressed genes from RNA-seq data. *Bioinformatics* **28,** 2782–2788 (2012).

60. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* **25,** 2078–2079 (2009).

Author contributions

V.N.G. coordinated the study. A.N.L., O.I.K. and A.A.M. collected and prepared animal samples. A.A.T., S.H.Y. and R.T.B. prepared and analysed biological samples. Z.X., X.F., Z.H., Y.Z., Y.F., D.F., X.Y., Y.M., Y.S., G.C. and Y.X. performed genome sequencing, assembly and annotation. X.F., Z.H., G.Z. and J.W. supervised genome sequencing, assembly and annotation. I.S., Z.X., A.V.L., X.F., S.M., Z.H., M.V.G., Y.Z., Y.F., D.F., X.Y., S.G.L., Y.M., Y.S., G.C., T.L., D.J., S.R.S., A.K. and V.N.G. performed genome and transcriptome analyses. All authors contributed to data interpretation. I.S. and V.N.G. wrote the paper with significant contributions from X.F., Z.X, Z.H. and input from all authors. We thank Ivan Chadin for assistance in specimen collection.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Seim, I. *et al.* Genome analysis reveals insights into physiology and longevity of the Brandt's bat Myotis brandtii. *Nat. Commun.* 4:2212 doi: 10.1038/ncomms3212 (2013).