Genome assembly and gene expression in the American black bear provides new insights into the renal response to hibernation

Anuj Srivastava¹, Vishal Kumar Sarsani¹, Ian Fiddes², Susan M. Sheehan¹, Rita L. Seger³, Mary E. Barter¹, Selena Neptune-Bear¹, Charlotte Lindqvist⁴, and Ron Korstanje ¹*

¹The Jackson Laboratory, Bar Harbor, ME 04609, USA, ²Genomics Institute, University of California, Santa Cruz, CA 95064, USA, ³Animal and Veterinary Sciences Program, University of Maine, Orono, ME 04469, USA, and ⁴Department of Biological Sciences, University at Buffalo, Buffalo, NY 14260, USA

*To whom correspondence should be addressed. Tel. 207288 6000. Fax. 207 288 6078. Email: ron.korstanje@jax.org

Received 27 May 2018; Editorial decision 1 October 2018; Accepted 4 October 2018

Abstract

The prevalence of chronic kidney disease (CKD) is rising worldwide and 10–15% of the global population currently suffers from CKD and its complications. Given the increasing prevalence of CKD there is an urgent need to find novel treatment options. The American black bear (Ursus americanus) copes with months of lowered kidney function and metabolism during hibernation without the devastating effects on metabolism and other consequences observed in humans. In a biomimetic approach to better understand kidney adaptations and physiology in hibernating black bears, we established a high-quality genome assembly. Subsequent RNA-Seq analysis of kidneys comparing gene expression profiles in black bears entering (late fall) and emerging (early spring) from hibernation identified 169 protein-coding genes that were differentially expressed. Of these, 101 genes were downregulated and 68 genes were upregulated after hibernation. Fold changes ranged from 1.8-fold downregulation (RTN4RL2) to 2.4-fold upregulation (CISH). Most notable was the upregulation of cytokine suppression genes (SOCS2, CISH, and SERPINC1) and the lack of increased expression of cytokines and genes involved in inflammation. The identification of these differences in gene expression in the black bear kidney may provide new insights in the prevention and treatment of CKD.

Key words: black bear, genome assembly, RNA-Seq, kidney, hibernation

1. Introduction

The prevalence of chronic kidney disease (CKD) is rising worldwide and currently 10–15% of the global population suffer from CKD and its devastating complications.¹ Although the adult human kidney has some ability to recover from acute kidney injury through cellular proliferation of the damaged intrarenal tissues, regenerating nephrons through de novo nephron development is considered unlikely, as the formation of new nephrons in humans is terminated at the embryonic stages. However, renal regeneration through nephron neogenesis in the event of renal injury has been described in some fish species and the possibility of this happening in mammalian species has not been excluded. Members of the bear family (Ursidae)
might be such species and studying how they deal with periods of decreased kidney function during hibernation (biomimicry) could be a new approach to understand kidney disease and develop new treatments.\(^2\)

The American black bear hibernates for up to 7 months annually. During this period, they do not eat, drink, urinate, or defecate. Bear hibernation is a state similar to prolonged sleep during which body temperature is reduced by 1–8\(^\circ\)C.\(^3\) There is a 20–50% reduction in metabolic rate with a depressed heart rate,\(^4\) and the volume of urine produced is reduced by 95%.\(^5\) The small volumes of urine and urea that enter the bladder during hibernation are reabsorbed across the bladder epithelium,\(^6\) and the urea is recycled for production of new protein.\(^7\) Throughout hibernation the kidney continues to concentrate urine and produces renin,\(^8\) erythropoietin,\(^9\) and vitamin D 1α-hydroxylase.\(^10\) Hibernating bears have the ability to prevent azotemia (high levels of nitrogen-containing compounds in the blood, common in human patients with renal function), but the mechanism is unknown.

Understanding these processes could lead to creating novel therapies for treating human conditions related to resistance to the complications of CKD and recovery from acute kidney injury. The studies aforementioned suggest unique kidney features in the American black bear allowing them to endure lower functioning during hibernation and recovery soon after hibernation. They are likely in part encoded in the genome sequence and gene expression patterns unique to the bear. To address this, we have performed high-throughput sequencing of genomic DNA and RNA isolated from kidneys of wild black bears. We generated a \textit{de novo} assembly and annotation of the complete genome and compared transcription profiles of kidneys collected in the spring (within weeks after hibernation) and in the fall (before hibernation).

2. Materials and methods

2.1. Sample collection, library preparation and sequencing

Bear samples were obtained by hunters during the hunting seasons in Maine. Hunters were asked to participate on a voluntary base and no bears were killed for the specific purpose of this study. All methods were carried out in accordance with relevant guidelines and regulations. DNA was isolated using the DNeasy Blood & Tissue Kit (Qiagen). The whole-genome library was prepared using the KAPA Hyper Prep Kit (Kapa Biosystems, Inc., Wilmington, MA) with a bead-based size selection to select for inserts with an average size of 400 bp and 6 cycles of PCR. Sequencing was done on two \(2 \times 125\) bp Illumina 2500 lanes. The mate pair library was prepared using the Illumina Nexera Mate Pair Kit (Illumina, San Diego, CA, USA) with a gel-based size selection to select for inserts with an average size of 10 kb and 14 cycles of PCR. Sequencing was done on two \(2 \times 100\) bp Illumina HiSeq 2500 lanes. The PacBio library was prepared using the Pacific Biosciences SMRTbell Template Prep Kit 1.0 (Pacific Biosciences, Menlo Park, CA, USA) using the ‘20-kb Template Preparation Using BluePippin Size-Selection System (15-kb Size cutoff)\(^1\) protocol obtained from PacBio SampleNet. The BluePippin was set to collect from 7 to 50 kb. Sequencing was done on 23 PacBio SMRT cells. For RNA-Seq, hunters collected kidneys and stored them in liquid nitrogen. Kidneys were first homogenized in Trizol (Invitrogen) and an aliquot was used for RNA isolation using the miRNeasy Mini kit (Qiagen), according to manufacturers’ protocols, including the optional DNase digest step. Quality was assessed using an Agilent 2100 Bioanalyzer instrument and RNA 6000 Nano LabChip assay (Agilent). We prepared total RNA for sequencing using the Illumina TruSeq methodology (TruSeq Stranded Total RNA LT Sample Prep Kit with Ribo-Zero Gold). The first step involves the removal of ribosomal RNA (rRNA) using biotinylated, target-specific oligos, combined with Ribo-Zero rRNA removal beads. After individual samples were bar-coded they were pooled and sequencing was done on two \(2 \times 75\) bp Illumina NextSeq 2500 lanes. All raw data have been submitted to the NCBI’s Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra) under accession number SRR075217.

2.2. Sequence assembly

All data were subjected to quality control check\(^9\) and samples with base qualities greater \(\geq 30\) over 70% of read length were used in the downstream analysis. KmerFreq_HA v2.0 and Corrector_HA v2.0\(^10\) tools of SOAP\(^{2,01}\) tools were further used to perform K-mer frequency generation and error correction of paired-end and mate-pair data, respectively. SOAP\(^{12,7mer}\) v2.04\(^10\) was used to perform the contig and scaffold assembly from paired end and mate-pair libraries (avg. ins = 350 and 10 K for paired-end and mate-pair, respectively). After the assembly, gaps in scaffolds were closed with the GapCloser tool (v 1.12)\(^10\) (with \(-1 125\)) of the soap module. Smallmap, perfectFrombam and pipeline tools of the reapr module (v 1.0.17)\(^11\) was used to recognize the errors in the assembly by re-mapping of paired-end and mate-pair data to the \textit{de novo} assembled genome. The assembly was broken at potential misassembled points. The broken assembly was further used as an input for the jelly tool of the PBSuite\(^12\) (v15.2.20) with blasr (v1.3.1) at parameters \([-\text{minMatch} 8, \text{minPctIdentity} 80, \text{bestn} 1, \text{nCandidates} 20, \text{maxScore} -500\]). This tool was used to upgrade the existing Illumina assembly with low-pass PacBio data. Repeat masking was performed by downloading the Repbase TE library from the repbase server (http://www.girinst.org/server/RepBase/). To identify known TE elements, we used repeat masker and repeat protein mask software in the Repeat Masker package (www.repeatmasker.org), which identifies TEs by aligning the genome sequence to a defined TE database. Tandem repeats were predicted using TRF\(^13\) by using the default parameters ‘\text{Match} = 2, \text{Mismatch} = 7, \text{Delta} = 7, \text{PM} = 80, \text{PI} = 10, \text{Minscore} = 50, \text{and MaxPeriod} = 12’. The completeness of the assembly was estimated by using CEGMA.\(^14\)\(^14\) The assembly was screened against a collection of 248 universal eukaryotic single-copy genes. Core eukaryotic gene datasets were downloaded and a blast database was made from the assembly before running CEGMA. This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession LNZR01000000. Version LNZR01000000 is described in this work. It is also available at ftp://ftp.jax.org/maine/blackbear_project/.

2.3. Annotation

The black bear genome assembly was annotated using Comparative Annotation Toolkit (CAT—https://github.com/ComparativeGenomics Toolkit/Comparative-Annotation-Toolkit). CAT uses whole genome alignments generated by progressiveCactus\(^15\) to project gene annotations from a high quality reference genome on to one or more target genomes. This process leverages previously curated annotation sets to rapidly construct a set of orthologues in the target genomes. After transcript projection,\(^16\) the protein coding projections are provided to the \textit{ab-initio} gene finding tool AUGUSTUS\(^17\) with additional extrinsic information derived from RNA-Seq. AUGUSTUS then enforces a coding
gene model that allows for shifts in splice sites to maintain frame. This process can also rescue exons that did not align in the whole genome alignment. CAT also performs true ab-initio gene prediction by using a new method of running AUGUSTUS called Comparative Augustus, or AugustusCGP. 18 This parameterization simultaneously predicts genes in all species in a progressive Cactus alignment, using RNA-Seq data in one or more species to help guide the annotation process. After projection, clean-up and ab-initio prediction, CAT combines these separate transcript sets into one final gene set through a consensus finding process. For orthologous transcripts, the transcript with the best fidelity to the reference with the best extrinsic support is chosen. In the case where multiple paralogous projections are found, CAT selects the most likely orthologue through a combination of splice junction fidelity, synteny and alignment identity. Finally, ab-initio predictions from AugustusCGP are evaluated for providing new information. If significant overlap with an orthologous projection is found, and the transcript provides a new splice junction or exon, then it is included as a new iso-form of the orthologue. Otherwise, the locus is considered a candidate for a novel gene, often a gene family expansion. In addition to an annotation set on target genomes, CAT produces a UCSC Assembly Hub. 19 This assembly hub has tracks for the raw transMap projections, the post-filtering projections, the various modes of AUGUSTUS employed, the final consensus annotation set, as well as the input RNA-Seq information including a filtered splice junction track. All annotations are stored in a modified bigBed format with a wide variety of additional features annotated on each entry, which can be accessed from the details page for that entry. Additional features include RNA-Seq support for specific splice junctions and binary classifiers such as having a frame-shifting indel relative to the source transcript. In addition to the genome-specific tracks, an alignment track (snake track) shows the cactus alignment between the current species and other species. Browsing this assembly hub provides the opportunity to visualize the relationship between all aligned species and the various ways the transcript projections were filtered and combined with ab-initio predictions. Assembly hub can be visualized by visiting UCSC ‘Track Data Hubs’ (https://genome.ucsc.edu/index.html) and then adding URL ftp://ftp.jax.org/hub can be visualized by visiting UCSC ‘Track Data Hubs’ (https://genome.ucsc.edu/index.html) and then adding URL ftp://ftp.jax.org/
2500 platform and single molecule sequencing using the PacBio RSII system (Supplementary Tables S2 and S3). Raw reads were assembled into 113,759 scaffolds and contigs ≥ 1,000 bases with an N50 length of 190 kb, totalling 2.59 Gb in length (Supplementary Table S4). This is slightly larger than the estimated size of the panda bear assembly (2.46 Gb)\(^{2,4}\) and the polar bear assembly (2.53 Gb)\(^{2,3}\).

We estimated the quality of the assembled sequence by mapping all paired-end reads back to the assembled genome with Burrows–Wheeler Aligner\(^{25}\) to determine the mappability and median coverage of the assembled genome. Approximately 90% of reads mapped back to the assembly with a mapping quality of \(\geq 30\) and with a median coverage of \(43\). The peak sequencing depth was \(50\), and more than 20 reads covered over 85% of the assembled sequences (Fig. 1a). The completeness of our assembly was estimated using CEGMA by screening against 248 highly conserved core eukaryotic genes. Our black bear assembly covers 212 out of 248 genes completely and 241 out of 248 genes partially. Figure 1b shows this 85.5% completeness compared with 88.7% for the panda bear assembly and 90.7% for the polar bear assembly, which are all comparable. GC content in mammals is correlated with a number of genomic features that are functionally relevant, for example, gene density, transposable element distribution, and methylation rate. The GC-content distribution in the black bear covers a narrower range compared with the panda bear and the polar bear (Fig. 1c). This might be a consequence of the lower amount of repetitive sequences that is present in the black bear genome (Fig. 2).

### 3.2. Annotation of the black bear genome

Similar to the panda genome,\(^{2,6}\) our black bear genome assembly was annotated using the Comparative Annotation Toolkit (CAT). CAT identified 29,624 genes (18,091 coding) in our black bear assembly, representing 88% of the genes present in the dog annotation. In addition, 2,995 transcripts with at least 2 splices were predicted by AugustusCGP, which also had at least one intron junction supported by RNA-Seq and not supported by transMap. Of these, 1,730 were associated with a transMapped gene and assigned as a novel isoform of that gene. The remainder were evaluated for being putatively novel loci. Six hundred and fifty-three loci were identified as being possibly paralogous, which is defined as overlapping a paralogous transMap projection, which was dropped in paralogue resolution. Twenty-eight loci were identified as possible false fusions, defined as having \(>80\%\) overlap with more than one transMap projection. Four hundred and fifty-eight loci were identified as poor alignments, which are predictions in regions, which have alignment to a known gene but did not pass transMap filtering. This left 126 putatively novel loci. Of the 22,081

---

**Figure 1.** Sequencing depth and comparison of several characteristics of the black bear, panda, and polar bear genome assemblies. (A) Distribution of sequencing depth of the assembled genome (B) Completeness and contiguity of the assembly was estimated using CEGMA by screening against 248 universal eukaryotic single-copy genes. Results for the black bear (85.5% completeness) were comparable with the panda (88.7%) and polar bear (90.7%). (C) Density plot of the %GC content in the bear genomes comparing the black bear (solid line) with the panda (dashed line) and polar bear (dotted line). The narrower distribution in the black bear might indicate fewer repetitive sequences.

**Figure 2.** Comparison of different types of repeats between the black bear and several mammalian species shows a lower number of repeat sequences in the black bear.
orthologous protein coding transcripts identified, pairwise codon-aware alignments of coding sequences identified frame-shifting indels in 7,848 transcripts. A slight enrichment of frame-shifting deletions was seen relative to insertions, suggesting a systematic bias in the assembly process.

3.3. High-throughput sequencing of renal RNA and differential expression between fall and spring

Kidneys were collected in the fall (animals JAX118, JAX134, JAX141, JAX146, JAX160, and JAX165), before hibernation, and in the spring (animals JAX101, JAX 103, JAX106, JAX113, JAX136, and JAX140), shortly after the bears emerged from hibernation (three males and three females for each time point). The timing of sample collection was (in part) driven by practicality. It was only possible to collect samples during a few specific periods, which were the late fall, before hibernation, and early spring, approximately 1 or 2 weeks after the animals come out of hibernation. Approximately 60 million RNA-Seq reads were obtained from each sample before quality control (Supplementary Table S5). After appropriate quality control and correction for batch effect, we performed a PCA (Supplementary Fig. S1a) and determined that sample 101 (from a male in the spring) was an outlier (see Supplementary Fig. S3 and Table S9). After removing this sample and repeating the analysis (Supplementary Fig. S1b) we compared the spring samples to the fall samples. We identified 169 differentially expressed protein-coding genes with an adjusted P-value below 0.05, of which CISH, SLCO1C1, and NDST3 showed a >2-fold change (Fig. 3 and Supplementary Table S6). Of these, 101 genes were downregulated and 68 genes were upregulated after hibernation. Fold changes ranged only from 1.8-fold downregulation (RTN4RL2) to 2.4-fold upregulation (CISH). Figure 4 shows a heatmap for all 11 bears with the 50 most significant genes (Supplementary Fig. S2 shows all 169 genes). Pathway enrichment analysis (Ingenuity Pathway Analysis) of the 169 genes did not identify any specific pathways. However, most notable was the upregulation of three cytokine suppression genes (SOCS2, CISH, and SERPINC1) and the lack of increased cytokine expression (e.g. IL6, CCL2, CCL6) and damage markers (LCN2 and HAVCR1) normally seen in lower functioning or recovering kidneys of other species.27

Because of the lack in inflammatory response, we asked whether the differentially expressed genes matched a particular development stage of the kidney. We used a method previously employed to place human cancers on a developmental landscape.28 First, we obtained gene expression profiles from mouse kidneys at different developmental stages (from E12.5 through E16.5) that were deposited in NCBI GEO and plotted the developmental expression profile in their first two temporal principal components (PC1-2). The high-dimensional expression profiles are simplified into a developmental timeline, ordering the genes in a linear array based on their temporal pattern of expression. Early genes are localized on the left end, genes with no bias towards early or late expression centre in the middle and late genes localize the right end. Thus, the unique order of genes on the timeline represents a summary of early and late states for each developmental process. Next, we matched the differentially expressed gene sets with their mouse orthologues (153 out of 169 genes) and plotted their location on the developmental timeline for the genes that were expressed in the developing kidney (109 genes). We observed clustering of down-regulated genes with the early mouse genes (E12.5) and clustering of upregulated genes with the late mouse genes (16.5) (Fig. 5). The processes that take place in the E12.5 kidney are focused on progenitor cell dynamics, FGF, and Wnt signaling, which we know are pro-fibrotic in the adult. In contrast, E16.5 is very different as it involves glomerulogenesis, tubulogenesis and maturation.

Taken together, despite the many months of reduced renal function and downregulated metabolism during hibernation, the gene expression differences between spring and fall bears is limited both in number of genes that are differentially expressed, fold-change, and the lack of upregulation of genes that are typically seen in other species after a period of low function or damage. In fact, the expression profile of bears coming out of hibernation resembles a developmental stage.

3.4. Identification of RNA-editing in the black bear kidney

Our RNA-Seq data identified possible RNA-editing in a small set of genes when aligning the reads to our assembled genome sequence. To rule out the possibility of variation in the transcripts due to single nucleotide polymorphisms (SNPs) in the genomes of the bears for which we performed the RNA-Seq, we sequenced the genomes of bears 101, 103, 113, 118, 134, and 160 at a low coverage (Supplementary Table S7). After comparing with the genome sequences and filtering for protein-coding transcripts we identified 38 transcripts containing a different variant (between 10% and 95% of the total reads) from the genome sequence (Supplementary Table S8). Of these 38 transcripts, 30 (79%) show A-to-I editing, while the other sites have C-to-U editing. Almost half of the edited sites are in the coding region and five are predicted to lead to an amino acid change.

According to the REDIportal (http://srv00.recas.ba.infn.it/atlas/) RNA-editing has been observed in four of the genes in the human kidney (although not at the same base pair) and for one gene, FLNB, RNA-editing also occurs both in the human and the mouse kidney. Interestingly, RNA-editing in 37 of the transcripts is observed in most or all 12 bears, but for ZNF688 we find editing in all 6 spring bears and no editing in the 6 fall bears.
To confirm our observations in the high-throughput sequencing data, we randomly selected eight editing sites and designed primers flanking these sites that could discriminate between RNA and possible genomic DNA contamination. Subsequent Sanger sequencing confirmed all eight editing sites (Supplementary Table S8).

4. Discussion

Previous studies in bears suggest they have unique features in the kidney that allow them to endure lower renal functioning during hibernation and recovery soon after hibernation. As these features are likely in part encoded in the genome sequence and gene expression patterns unique to the bear, we set out to explore the genome sequence and renal gene expression of the American black bear (*Ursus americanus*).

A first important step was to establish the complete annotated genome sequence. To accomplish this, we used three approaches for sequencing a single male bear with 100x coverage before filtering. This resulted in a final, high quality, assembly. This current genome is comparable in quality and size to the published panda genome24 and polar bear genome.23 A comparison between the three genomes shows a narrower range in GC-content in the black bear (Fig. 1c), likely the result of fewer repetitive sequences caused by transposon integration (Fig. 2).

Figure 4. Cluster analysis of the gene expression profiles showing the 50 genes with the lowest P-value.
We expected we would find a signature of these changes and possible function.

Novel loci that warrant further investigation as to their uniqueness allowed us to establish a high-quality annotation of the genome with 22,081 protein coding transcripts and identification of 126 putative orthologues of genes that are downregulated after hibernation in blue and in their first two temporal principal components (PC1-2) with the mouse kidney developmental expression profile at E12.5–E16.5 (Figure 5). The new Comparative Annotation Toolkit (CAT) pipeline allowed us to establish a high-quality annotation of the genome with 22,081 protein coding transcripts and identification of 126 putative novel loci that warrant further investigation as to their uniqueness and possible function.

Based on the reported changes in the bear kidney function during hibernation, we expected we would find a signature of these changes in the gene expression of kidneys collected soon after hibernation when compared with kidneys collected before hibernation. We therefore collected samples from 3 male and 3 female bears (samples were collected within 10 days from each other) soon after hibernation and from 3 male and 3 female bears before hibernation. One limitation is that the time between waking up from hibernation and sample collection is unknown and we assumed it was similar for all animals. Deviation from this assumption is expected to lead to more variation in the gene expression profile. RNA-Seq analysis identified 169 protein coding genes that were differentially expressed between the two time points. Of these, 101 genes were downregulated and 68 genes were upregulated after hibernation. Fold changes only ranged from 1.8-fold to 805-fold. Of these, 101 genes were downregulated and 68 genes were upregulated after hibernation cluster with a developmental stage in which glomerulogenesis, tubulogenesis and maturation are important. In patients with CKD, systemic inflammation is a common finding that promotes premature aging and predicts poor outcome.32 Although this is the first study in which gene expression is analysed in kidney tissue, it is not the first in which gene expression in bears has been examined. Fedorov et al. analysed samples from captive bears that were either kept in hibernating or non-hibernating conditions with a custom microarray. They analysed heart and liver samples and found 245 differentially expressed genes in the heart and 319 genes in the liver.31 We were able to map the majority (85.2%) of their ESTs onto our genome assembly and match them to our transcripts. However, none of the genes that were found to be differentially expressed in heart and liver overlapped with differentially expressed genes in the kidney. The lack of overlap could be due to difference in experimental conditions such as the timing of tissue harvest. Alternatively, it may indicate that while nearly every tissue adapts to hibernation, the ensemble of genes involved is highly tissue-specific.

Overall, both the number of differentially expressed genes and the fold-changes are low, but are in line with the findings of Fedorov et al. Although it has been shown that American black bears maintain a reduced metabolic rate for up to 3 weeks after emerging from their dens, a possible explanation for our results is that the kidneys already completely recovered from hibernation within the short time between emerging from hibernation and sample collection. Another explanation is that the reduction in metabolic rate and urine production has a much smaller impact on the kidney than would be predicted. Our RNA-Seq data also identified the presence of RNA editing, which was confirmed for all sites that were tested with Sanger sequencing. Several of the genes that are edited also show RNA-editing in mouse and human, but at different positions (Supplementary Table S8). The most interesting is the RNA editing in the 3’ UTR of the transcriptional regulator ZNF688, which only seems to happen in the spring. This suggests season-dependent RNA-editing and a possible mechanism through which gene expression is altered.

In conclusion, we established a high-quality and well-annotated genome sequence of the black bear (U. americanus). RNA-Seq of kidney samples before and after hibernation shows RNA-editing and the modest differential expression of a set of 169 genes that might be involved in the unique stress response due to hibernation in the bear. Our results suggest that during hibernation, changes in gene expression favour anti-inflammatory pathways. Further studies are needed to understand the effects of these expression differences in the kidney.

**Acknowledgements**

The authors thank Heidi Munger, Ryan Lynch, Edith Schriever, Lucy Rowe, and Xiaoan Ruan, for technical assistance, and Carol Bult, Peter Stemvinkel, Leif Oxburgh, and Paul Shiels for discussion and critical review of the
manuscript. A special thanks to the Penobscot Indian Nation and Passamaquoddy Tribe for providing bear kidney samples.

Funding
This work was supported by the NIDDK Diabetic Complications Consortium [grant DK076169, Subaward 25034-70] and the National Cancer Institute [P30CA034196].

Supplementary data
Supplementary data are available at DNARES online.

Conflict of interest
None declared.

References
1. Levin, A., Tonelli, M., Bonventre, J., et al. 2017, Global kidney health 2017 and beyond: a roadmap for closing gaps in care, research, and policy, Lancet, 390, 1888–917.
2. Stenvinkel, P. and Johnson, R.J. 2013, Kidney biomimicry—a rediscov-
ered scientific field that could provide hope to patients with kidney dis-
ease, Arch. Med. Res., 44, 584–90.
3. Tøien, Ø., Blake, J., Edgar, D.M., Grahn, D.A., Heller, H.C. and Barnes, B.M. 2011, Hibernation in black bears: independence of metabolic sup-
pression from body temperature, Science, 331, 906–9.
4. Hellgren, E.C. 1998, Physiology of hibernation in bears, Ursus, 10, 467–77.
5. Nelson, R.A., Wahnert, H.W., Jones, J.D., Ellefson, R.D. and Zollman, P.E. 1973, Metabolism of bears before, during, and after winter sleep, Am. J. Physiol., 224, 491–6.
6. Barboza, P.S., Farley, S.D. and Robbins, C.T. 1997, Whole-body urea cy-
cling and protein turnover during hyperphagia and dormancy in growing bears (Ursus arctos and U. americanus), Can. J. Zool., 75, 2129–36.
7. Brown, D.C., Mulhausen, R.O., Andrew, D.J. and Seal, U.S. 1971, Renal func-
tion in anesthetized dormant and active bears, Am. J. Physiol., 220, 293–8.
8. Seger, R.L., Cross, R.A., Rosen, C.J., et al. 2011, Investigating the mecha-
nism for maintaining eucalcaemia despite immobility and anuria in the hi-
bernating American black bear (Ursus americanus), Bone, 49, 1205–12.
9. Patel, R.K. and Jain, M. 2012, NGS QC Toolkit: a toolkit for quality con-
trol of next generation sequencing data, PLoS One, 7, e30619.
10. Luo, R., Liu, B., Xie, Y., et al. 2012, SOAPdenovo2: an empirically im-
proved memory-efficient short-read de novo assembler, Gigascience, 1, 18.
11. Hunt, M., Kikuchi, T., Sanders, M., Newbold, C., Berriman, M. and Otto, T.D. 2013, REAPR: a universal tool for genome assembly evalua-
tion, Genome Biol., 14, R47.
12. English, A.C., Richards, S., Han, Y., et al. 2012, Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology, PLoS One, 7, e47768.
13. Benson, G. 1999, Tandem repeats finder: a program to analyze DNA sequences, Nucleic Acids Res., 27, 573–80.
14. Parra, G., Bradnam, K. and Korf, I. 2007, CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes, Bioinformatics, 23, 1061–7.
15. Paten, B., Earl, D., Nguyen, N., Diekhans, M., Zerbino, D. and Haussler, D. 2011, Cactus: algorithms for genome multiple sequence alignment, Genome Res., 21, 1512–28.
16. Stanke, M., Diekhans, M., Baertsch, R. and Haussler, D. 2008, Using na-
tive and syntenically mapped cDNA alignments to improve de novo gene finding, Bioinformatics, 24, 637–44.
17. Stanke, M., Schöffmann, O., Morgenstern, B. and Waack, S. 2006, Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources, BMC Bioinform., 7, 62.
18. König, S., Romoth, L.W., Gentlicher, L. and Stanke, M. 2016, Simultaneous gene finding in multiple genomes, Bioinformatics, 32, 3388–95.
19. Nguyen, N., Hickey, G., Raney, R.J., et al. 2014, Comparative assembly hubs: web-accessible browsers for comparative genomics, Bioinformatics, 30, 3293–301.
20. Schwab, K., Hartman, H.A., Liang, H.-C., Aronow, B.J., Patterson, L.T. and Potter, S.S. 2006, Comprehensive microarray analysis of Hoxa11/Hoxd11 mutant kidney development, Dev. Biol., 293, 540–54.
21. Gautier, L., Cope, L., Bolstad, B.M. and Irizarry, R.A. 2004, affy—an analysis of Affymetrix GeneChip data at the probe level, Bioinformatics, 20, 307–15.
22. Kinsella, R.J., Khāhārī, A., Haider, S., et al. 2011, Ensembl BioMartS: a hub for data retrieval across taxonomic space, Database (Oxford), 2011, bar030.
23. Miller, W., Schuster, S.C., Welch, A.J., et al. 2012, Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change, Proc. Natl. Acad. Sci., 109, E2382–E90.
24. Li, R., Fan, W., Tian, G., et al. 2010, The sequence and de novo assembly of the giant panda genome, Nature, 463, 311–7.
25. Li, H. and Durbin, R. 2009, Fast and accurate short read alignment with Burrows–Wheeler transform, Bioinformatics, 25, 1754–60.
26. Chen, M., Hu, Y., Liu, J., et al. 2015, Improvement of genome assembly completeness and identification of novel full-length protein-coding genes by RNA-Seq in the giant panda genome, Sci. Rep., 5, 18019.
27. Kumar, S., Liu, J. and McMahon, A.P. 2014, Defining the acute kidney injury and repair transcriptome, Semin. Nephrol., 34, 404–17.
28. Kho, A.T., Zhao, Q., Cai, Z., et al. 2004, Conserved mechanisms across development and tumorogenesis revealed by a mouse development per-
spective of human cancers, Genes Dev., 18, 629–40.
29. Kupchella, C.E., Heckman, J.R., Lega, R.D. and Plamp, C.E. 1977, The influence of cold-torpor and water deprivation on plasma and kidney gly-
cosaminoglycans in emydine turtles, Comp. Biochem. Physiol. A Comp. Physiol., 56, 225–9.
30. Liew, H., Roberts, M.A., MacGinley, R. and McMahon, L.P. 2017, Endothelial glycoalyx in health and kidney disease: rising star or false dawn? Nephrology (Carlton), 22, 940–6.
31. Stenvinkel, P., Frobert, O., Anderstam, B., et al. 2013, Metabolic changes in summer active and anuric hibernating free-ranging brown bears (Ursus arctos), PLoS One, 8, e72934.
32. Kooman, J.P., Dekker, M.J., Usvyat, L.A., et al. 2017, Inflammation and mate change, Database (Oxford), 2017, bar030.