Transcriptomic analyses reveal tissue-specific selection on genes related to apoptotic processes in the subterranean rodent, *Ctenomys sociabilis*

Andrew Lang[1+], Lauren Kordonowy[1], Eileen Lacey[2], Matthew MacManes[1*]

[1]Department of Molecular, Cellular and Biomedical Sciences
University of New Hampshire
Durham, NH 03824
[1+]al2025@wildcats.unh.edu
[1*]Matthew.MacManes@unh.edu

[2] Museum of Vertebrate Zoology
Department of Integrative Biology

University of California, Berkeley
Berkeley, CA 94706

[*]Corresponding Author

ABSTRACT

Specialization for a subterranean existence is expected to impact multiple aspects of an organism's biology, including behavior, physiology, and genomic structure. While the phenotypic correlates of life underground have been extensively characterized, the genetic bases for these traits are not well understood, due in part to the challenges of generating large, multi-locus data sets using traditional DNA sequencing strategies. To begin exploring the genomic architecture of adaptation to a subterranean existence, we generated high-quality *de novo* transcriptome assemblies for 8 different tissue types (hippocampus, hypothalamus, kidney, liver, spleen, ovary, testis, skin) obtained from the colonial tuco-tuco (*Ctenomys sociabilis),* a group-living species of subterranean rodent that is endemic to southwestern Argentina. From these transcriptomes, we identified genes that are evolving more rapidly in the *C. sociabilis* lineage compared to other subterranean species of rodents. These comparisons suggest that genes associated with immune response, cell-cycle regulation, and heavy metal detoxification have been subject to positive selection in *C. sociabilis*. Comparisons of transcripts from different tissues suggest that the spleen and liver - organs involved in immune function and detoxification - may be particularly important sites for these adaptations, thereby underscoring the importance of including multiple tissue types in analyses of transcriptomic variation. In addition to providing an important resource for future genomic studies of *C. sociabilis*, our analyses generate new insights into the genomic architecture of functionally significant phenotypic traits in free-living mammals.

INTRODUCTION

Convergent traits provide critical opportunities to examine interactions between shared environmental challenges, selection, and the evolution of phenotypic and genotypic variation (Mares, 1975; Muschick, Indermaur & Salzburger, 2012; Parker et al., 2013). One well-characterized suite of convergent phenotypic traits occurs among subterranean rodents, which are defined by their tendency to spend virtually their entire lives in underground burrows (Nevo, 1979; Lacey & Patton, 2000). This designation encompasses more than 120 species representing 6 families and 3 suborders of rodents (Lacey & Patton, 2000; Gardner, Wilson & Reeder, 2005). Shared physiological challenges associated with life underground include the high energetic costs of excavating burrows (Luna & Antinuchi, 2006; Zelová et al., 2011), hypoxia and hypercapnia (Lovegrove, 1986; Buffenstein, 2000), maintenance of circadian patterns of activity (Vasicek et al., 2005; Urrejola et al., 2005; Tomotani et al., 2012), and, at least in some habitats, exposure to heavy metals in soils (Lovegrove, 1986; De Vleeschouwer et al., 2014; Fernández-Cadena et al., 2014). While the convergent phenotypic processes associated with these challenges have been studied in some detail (Nevo, 1979; Buffenstein, 2000; Burda, Šumbera & Begall, 2007), the genetic architecture underlying similar physiological responses to these challenges remains

largely unknown (but see Partha et al., 2017). Determining the proximate mechanisms (*e.g.,* the genetic underpinnings) of adaptations enabling organisms to thrive in such an environment is critical to improving our understanding of how specializations for a subterranean existence arise and are maintained.

The advent of high-throughput transcriptome sequencing has greatly facilitated efforts to relate patterns of gene expression to differences in phenotypic traits, including physiological processes such as metabolism (Devi et al., 2016) and water regulation (Kordonowy & MacManes, 2016; MacManes, 2017). This sequencing strategy has also been used to identify physiologically relevant regions of the genome undergoing positive selection (Zhang, Dyer & Rosenberg, 2000; Swanson et al., 2001; Brodsky et al., 2005; Kosiol et al., 2008; Karn et al., 2008; Gardiner et al., 2008; Kong et al., 2011), thereby generating insights into the evolutionary bases for relationships between gene expression and specialization for specific phenotypic attributes. The marked examples of evolutionary convergence and divergence among burrow-dwelling mammal species offer an ideal opportunity to implement sequencing methods for exploration of the genomic bases for the functional and evolutionary consequences of a shared subterranean lifestyle.

The colonial tuco-tuco *(Ctenomys sociabilis*) is a subterranean rodent that is endemic to Neuquen Province, Argentina (Tammone, Lacey & Relva, 2012). This species has been the subject of extensive research due to its unusual social system; while the majority of ctenomyids are thought to be solitary, *C. sociabilis* is group living, with burrow systems routinely occupied by multiple adult females plus, in many cases, a single adult male (Lacey, Braude & Wieczorek, 1997; Lacey & Wieczorek, 2004). In particular, this species has been studied with respect to not only to behavior, ecology and demography (Lacey, Braude & Wieczorek, 1997; Lacey, 2001; Chan & Hadly, 2011; MacManes & Lacey, 2012), but also neuroendocrinology (Beery, Lacey & Francis, 2008; Woodruff et al., 2013) and population genetic structure (Lacey, 2001; Hambuch & Lacey, 2002; Chan et al., 2005). Compared to other subterranean rodents for which transcriptomic data are available (Malik et al., 2011; Lin et al., 2014), *C. sociabilis* is phylogenetically, geographically, and behaviorally distinct, suggesting that this species is critical to efforts to examine the genomic impacts of adaptation to life underground.

Here we present a high-quality annotated transcriptome generated from eight tissue types (hippocampus, hypothalamus, kidney, liver, spleen, ovary, testis, skin) obtained from *C. sociabilis.* The use of multiple tissues has resulted in a particularly complete transcriptome for htis non-traditional study species. In addition to presenting this annotated assembly, we characterize each tissue type with regard to the most highly abundant transcripts, after which we compare patterns of expression across tissue types. We then

conduct a comparative analysis of coding sequence evolution in *C. sociabilis* based on contrasts with single-tissue transcriptomes from seven other subterranean rodent species. In addition to highlighting the importance of tissue type in determining patterns of transcript abundance, our analyses generate important new insights into the genetic correlates of subterranean life.

METHODS

Sample collection, RNA Extraction & Library Preparation

Tissue samples were obtained from two adult *C. sociabilis* (1 male and 1 female) that were members of a captive population of this species maintained at the University of California, Berkeley. The housing and husbandry of this population have been described previously (MacManes & Lacey, 2012; Woodruff et al., 2013). The animals sampled were euthanized via overdose with Isoflurane followed by decapitation. The hippocampus, hypothalamus, kidney, liver, ovary, skin, and spleen were extracted from the female and the testes were extracted from the male. Each tissue type was placed in a cryotube containing RNAlater (Thermo Fisher Scientific, Waltham, MA) and then flash frozen with liquid nitrogen. The interval between euthanasia and flash freezing of tissues did not exceed five minutes. All tissue samples were stored at -80°C until they were sent to the Broad Institute (Cambridge, MA) for RNA extraction, cDNA library preparation, and 125bp paired-end sequencing on an Illumina 2500 platform. All procedures involving live animals were approved by the Berkeley Animal Care and Use Committee and were consistent with guidelines established by the American Society of Mammalogy for the use of wild mammals in research (Sikes, 2016).

Tissue-Specific Transcriptome Assembly

Tissue-specific Illumina reads (36-49 million paired-end reads per tissue) were obtained for each of the 8 tissue types examined. For each tissue type, read quality was evaluated with SolexaQA++ v3.1.4 (Cox, Peterson & Biggs, 2010) and reads were corrected using Rcorrector v1.0.1 (Song & Florea, 2015). Adaptor sequences and reads falling below the quality threshold PHRED=2 were removed using Trimmomatic (Bolger, Lohse & Usadel, 2014), following the protocol of MacManes (2014). *De novo* transcriptome assemblies were generated using Trinity v2.1.1 (Haas et al., 2013). For each tissue type, two assemblies were generated – a khmer normalized (Crusoe et al., 2015) 100x coverage assembly and a non-normalized assembly. Digital normalization had no detectable effect on either the completeness or

the consistency of the resulting transcriptomes (Table S1) and thus all downstream analyses were conducted using assemblies generated from the non-normalized datasets.

Compiled Transcriptome Assembly, Annotation and Analysis

In addition to tissue-specific transcriptomes, read data from all tissue types were pooled to generate a single, merged transcriptome assembly. We produced 12 alternative merged assemblies through combinations of read subsampling, transrate optimization, and merging algorithms. Each assembly was evaluated for quality using TransRate v1.0.1 (Smith-Unna et al., 2016), which generates both quality metrics and an optimized assembly. In addition, we evaluated each assembly for completeness using the Vertebrata database within BUSCO v1.1b1 (Simão et al., 2015). Based on these analyses, we selected the assembly with the highest quality and completeness. The pipeline for producing this selected assembly is described below. Because previous research has revealed that little information is gained from using datasets above 40M reads (MacManes, 2015), a random subset of 50 million paired-end reads were selected for analyses (seqtk v1.0-r82 (https://github.com/lh3/seqtk)) from the entire dataset (N= 339 million reads). The subset of reads was assembled with both Trinity v2.1.1 and BinPacker (Liu et al., 2016). The resulting two assemblies were merged into a single assembly using Transfuse v0.5.0 (https://github.com/cboursnell/transfuse). This merged assembly was optimized with TransRate to retain only highly-supported contigs. The resulting assembly was annotated with dammit version 0.3.2 (https://github.com/camillescott/dammit) and filtered to retain only annotated transcripts.

Transcript Abundance and Gene Presence/Absence

To generate measures of relative transcript composition across tissue types, the abundance of each annotated transcript in our tissue-specific assemblies was assessed using Kallisto v0.42.4 (Bray et al., 2016). Transcripts with TPM (transcripts-per-million) values of less than 1 were determined to be absent from a given tissue (MacManes et al., 2017; MacManes & Lacey, 2012; MacManes & Eisen, 2014). Transcript presence/absence was compared across all tissues using the UpSetR package (Gehlenborg, 2016), and the 10 most abundant genes were identified within each tissue.

Comparative Analysis with Other Subterranean Taxa

To compare patterns of gene evolution across multiple lineages of subterranean rodents, we downloaded Illumina RNAseq reads for 5 other subterranean species (*Spalax carmeli, Bathyergus suillus,*

Tachyoryctes splendens, Eospalax baileyi, Cryptomys hottentotus pretorian) from the NCBI Sequence Read Archive (accession numbers SRR2016467, SRR2141210, SRR214121, SRR931783, and SRR2141213, respectively). In addition, we downloaded mRNA datasets derived from whole genome sequencing projects for a sixth species of subterranean rodent (*Heterocephalus glaber:* Mole Rat genome v1.7.2 http://gigadb.org/dataset/100022) and for *Mus musculus (Mus* genome vGRCm38); the latter served as the outgroup for these analyses. These mRNA data sets were assembled following the Oyster River Protocol (http://oyster-river-protocol.readthedocs.io/, (MacManes, 2015)). Together with the transcripts for *C. sociabilis* generated here, this comparative data set encompassed 3 families of subterranean rodents (Ctenomyidae, Spalacidae, Bathyergidae), each of which represents a phylogenetically distinct origin of specialization for life in underground burrows.

For each of the species in this comparative data set, coding sequences were identified using TransDecoder v3.0.0 (Haas et al., 2013). Orthologous relationships among these species (including the *M. musculus* outgroup) were identified using the output from BUSCO v2.0 and the associated database of mammalian sequences. The resulting groups of orthologous transcripts were then edited to include only single copy transcripts, which were then aligned using Prank v150803. Sequence alignments were refined using pal2nal v14 (Suyama, Torrents & Bork, 2006) and a gene tree was constructed using RAxML v8.2.8 (Stamatakis, 2014). To explore potential evidence of selection on the genes included in our dataset, we used PAML v4.9a (Yang, 2007), with our gene tree as the phylogenetic framework. Specifically, we tested for positive selection using the M7 versus M8 models in PAML. We then tested for evidence of lineage-specific selection using the PAML branch-site model with *C. sociabilis* as the foreground lineage. We controlled the false discovery rate for multiple comparisons following the procedure of Benjamini and Hockberg (1995). Genes determined to be under positive selection were then examined using the Gene Ontology Consortium Enrichment Analysis (http://geneontology.org/page/go-enrichment-analysis) tool to determine if these loci were grouped according to ontology terms.

To explore potential tissue-specific patterns of gene expression among loci identified as being under positive selection in *C. sociabilis*, we imported gene expression count data generated by Kallisto into the R statistical package v3.3.0 (Team, R C, 2013). To allow comparisons across tissue types, we normalized count data using the TMM method (McCarthy, Chen & Smyth, 2012) as implemented in edgeR v3.1.4 (Robinson, McCarthy & Smyth, 2010). For each transcript under positive selection, we identified the tissue for which the expression level was highest. These maximum count values were then normalized by dividing by the total number of genes expressed in that tissue; this procedure allowed us to identify tissues enriched for positively selected transcripts.

Data and Code Availability

Sequence read files for this study are available on the NCBI Short Read Archive (PRJNA358281). All code used in transcriptome assembly, annotation, analyses, and data visualization is freely available online at (https://github.com/macmanes-lab/tuco_manuscript and https://github.com/macmanes-lab/paml). The tissue-specific assemblies, as well as the final merged *C. sociabilis* transcriptome assembly are available on Dropbox (in fasta format), as are all annotation data files (in gff3 format) and kallisto transcript counts (https://www.dropbox.com/sh/jq98iderelxi9sm/AAAQG6Ex51sG9dcIrb8vK8gPa?dl=0). These files will be uploaded to Dryad upon acceptance of this manuscript for publication.

RESULTS AND DISCUSSION

Tissue-specific Transcriptome Assembly Analysis

Individual tissue-specific transcriptome assemblies were 68-82% complete (mean= 75.87%), with TransRate scores ranging from 0.145 to 0.172 (Table S1). The TransRate optimized assemblies, which included only highly-supported transcripts, contained on average 7% fewer BUSCOs than the original assemblies. Due to this pronounced reduction in completeness, the TransRate optimized assemblies were not used for subsequent analyses. While individual, non-optimized tissue-specific assemblies were of acceptable quality and completeness, they were notably inferior in quality and completeness to the compiled, transfused assembly described below.

Compiled Transcriptome Assembly, Annotation and Analysis

The most complete and highest quality assembly was generated from a 50 million read-pair subsample of the full dataset (Table 1). This assembly was annotated and all non-annotated transcripts were removed to produce the final assembly (annotation_only; Table 1). Removal of unannotated transcripts resulted in minimal reduction of TransRate and BUSCO scores but reduced the number of contigs by ~ 50%; the transcripts removed were likely artifacts of the assembly process (Moreton, Izquierdo & Emes, 2016) and thus this reduction was was not considered problematic. Reads from different tissue types mapped to the final transcriptome at a rate of 86-90% (Table 2). The final assembly contained 96,224 annotated transcripts, with 79,938 search matches to the Uniref90 database, 73,896 matches to OrthoDB (Waterhouse et al., 2013), 46,659 matches to PFAM, and 2,698 matches to RFAM (Griffiths-Jones et al.,

2005). Of the 96,224 transcripts in this final assembly, 78,241 (81.3%) contained open reading frames (ORFs) and 53,711 (55.8%) contained complete ORFs, indicating that these transcripts included the entire protein-coding sequence for the associated locus.

Comparative analysis with Other Subterranean Taxa

Using the output from BUSCO, we identified 2,182 single-copy ortholog groups from the transcriptomes of seven subterranean rodent species and from *Mus musculus*. Of these, 1,951 (89.4%) were successfully aligned and analyzed via PAML software. Branch site analysis identified 50 transcripts as being under positive selection in the lineage leading to *C. sociabilis;* in contrast, only seven were identified using the site-model of positive selection. While the larger set of transcripts identified using the branch-site model for GO enrichment did not reveal statistically significant enrichment of GO terms for *C. sociabilis* genes under positive selection, it did reveal that many of the GO terms identified corresponded to processes of cell proliferation control, DNA damage response, immune response, and ion transport. These findings are intriguing in light of evidence suggesting that burrowing rodents may be exposed to heavy metals or other toxins in the soils that they inhabit (De Vleeschouwer et al., 2014; Fernández-Cadena et al., 2014) and recent studies characterizing the immunogenetics of subterranean rodents (Cutrera et al., 2010; Merlo, Cutrera & Zenuto, 2016; Novikov et al., 2016). Particularly exciting is the identification of transcripts involved in the control of cell proliferation, which has potential ties to susceptibility to cancer (Tian et al., 2013).

For each gene under positive selection, we identified the tissue in which it was most abundant (Figure 1). We then compared the number of positively selected genes per tissue to that expected under a random distribution of these loci across tissue types – that is we divided the 50 genes under positive selection by the number of tissues (N=8) sequenced and then normalized these values according to the overall number of genes expressed in each tissue. This analysis revealed a significantly higher representation of genes under positive selection in the spleen and liver ($\chi 2$test, p-values <0.05), an outcome that is perhaps not surprising given the functional roles of these tissues. Collectively, the preponderance of genes under positive selection in *C. sociabilis* that are associated with response to cell damage and immune response suggests that the environmental physiology of this species deserves further investigation.

Transcript Abundance and Gene Presence/Absence

Filtering of transcripts to remove those for which TPM was less than 1 (Havens & MacManes, 2016; Kordonowy & MacManes, 2016) removed 5,722 (6.0%) of our annotated transcripts. Of the remaining 90,502 transcripts, 21,602 (23.9%) were expressed in all of the tissue types examined. In contrast, 774 (0.9%) of these transcripts were expressed in only a single tissue type. The distribution of these unique transcripts across tissue types was as follows: skin (N = 171), liver (N = 156), testes (N = 140), ovary (N = 93), spleen (N = 92), kidney (N = 77), hypothalamus (N = 23), and hippocampus (N = 22).

Between 81% and 88% of reads mapped to the reference transcriptome. Visual representations of transcript overlap between tissue types are presented in Figures 2, S1, and S2. The 10 most common transcripts unique to each tissue type are shown in Figure 3. While our data set did not allow a statistical comparison of levels of gene expression across tissue types, our assessments of transcript abundance per tissue type provide potential insights into the function of each tissue examined (Table S2). In particular, pairwise comparisons of transcript abundance revealed that tissues with similar functions tended to display similar suites of highly-expressed transcripts. For example, the two brain tissues examined – the hippocampus and the hypothalamus – shared the highest number of transcripts (5,200 out of 62,716 and 66,421 transcripts, respectively). The two reproductive tissues examined – the testes and the ovary – had an overlap of 1,359 out of 66,876 and 67,251 transcripts, respectively. The spleen did not share many transcripts with other tissues; the greatest overlap in spleen transcripts was with the testes (400 of 66,876 transcripts) and the ovary (298 of 67,251 transcripts). The kidney and liver, both associated with detoxification, shared 1,382 of 61,767 and 46,063 transcripts, respectively. Somewhat surprisingly, of the 58,796 transcripts in the skin, this tissue shared 1,397 with the ovary, the largest transcript overlap of any other tissue with the skin.

Tissue Characterization

Each of the tissue types included in this study has been well characterized with respect to its function in mammalian biology. Accordingly, we examined whether functional differences between tissues were reflected in the identities of the most abundant transcripts unique to each tissue. We also assessed loci under positive selection, highlighting aspects we believe may be key factors associated with live in underground burrows. The functions of many of the most abundant transcripts that were unique to a given tissue type have been characterized as part of empirical studies, as described below:

The hippocampus. The hippocampus is integrally involved in neurotransmission (Vianna et al., 2000; Shatz, 2009). In particular, the hippocampus has been studied with regard to spatial memory and

navigation (Bannerman et al., 2002; Eichenbaum, 2017) and as a site for for adult neurogenesis in the mammalian brain (Seri et al., 2001; van Praag et al., 2002). Among the transcripts that were uniquely abundant in the hippocampus in *C. sociabilis* were genes associated with regulating presynaptic density (Neurexin: NRXN1, TPM= 33.72) and synchronous firing of hippocampal pyramidal cells (Carbonic Anhydrase VII: CA7, TPM= 7.78) (Ruusuvuori et al., 2004; Kumar & Thakur, 2015). Loci found to be under positive selection in the hippocampus include genes involved in cell cycle progression and tumor growth, such as BRCA1 Associated Protein 1 (BAP1) and Apoptosis Antagonizing Transcription Factor (AATF) (Bruno et al., 2002; Qin et al., 2015). Both of these genes have been implicated in tumor suppression and cell growth inhibition, with BAP1 functioning by means of deubiquitinating host cell factor-1 (Machida et al., 2009) and AATF acting as an essential cofactor for the p53 gene (Bruno, Iezzi & Fanciulli, 2016).

The hypothalamus. The hypothalamus has been implicated in multiple critical signaling pathways, such as the Hypothalamic-Pituitary-Adrenal (stress) and Hypothalamic-Pituitary-Gonadal (reproductive) axes in vertebrates (Hall et al., 2012; Clément, 2016). Transcripts that were uniquely abundant in the hypothalamus tended to be directly involved in downstream signaling of activities such as feeding and parental or sexual behaviors (Insulin Receptor Substrate 4: IRS4, TPM= 31.63) as well as formation of the diencephalon and prethalamic brain region (FEZ Family Zinc Finger 1: FEZF1, TPM= 24.65) (Numan & Russell, 1999; Shimizu & Hibi, 2009). Genes identified to be under positive selection, similar to those identified for the hippocampus, are implicated in the cell cycle. For example, Prostate Androgen-Regulated Mucin-Like Protein 1 (PARM1) functions in prostate cell androgen dependence, has been linked to apoptotic mechanisms (Bruyninx et al., 1999), and may impart cell immortalisation (Cornet et al., 2003).

The ovary. Ovarian function is highly regulated by hormonal signals that mediate cell proliferation and the production of viable ova (Verga Falzacappa et al., 2009). Transcripts that were uniquely abundant in the ovary included an immunogene (Immunoglobulin Kappa Locus: IGK, TPM= 102.30) as well as genes involved in neuron development (NSMF, TPM= 50.20), and primordial follicle formation (Follistatin: FST, TPM= 23.87) (Brekke & Garrard, 2004; Palevitch et al., 2009; Kimura, Bonomi & Schneyer, 2011). Ovarian genes under positive selection (e.g., Nuclear Mitotic Apparatus Protein 1; NUMA1) tend to function in the structural components of cellular division and mRNA binding. For example, Nuclear Mitotic Apparatus Protein 1 (NUMA1) interacts with proto-oncogene PIM1 during mitosis and regulates p53-mediated transcription (Bhattacharya et al., 2002).

The testis. Similar to the ovaries, testis function is regulated hormonally and results in the production of viable gametes (Alves et al., 2013; O'Shaughnessy, 2014). The uniquely most abundant testis transcripts included an antimicrobial defense immunogene (Beta-defensin: DEFB118, TPM= 86.89), a transcription factor (PAS Domain Containing 1: PASD1, TPM= 45.73), and a gene unique to the testes that has not been fully characterized with regard to structure or function (P Antigen Family, Member 1: PAGE1, TPM= 44.62). Testicular genes under positive selection include known regulators of DNA damage (SprT-Like N-Terminal Domain; SPRTN, Ring Finger and WD Repeat Domain 3; RFWD3) (Fu et al., 2010; Gong & Chen, 2011; Liu et al., 2011; Juhasz et al., 2012) and cell proliferation regulation (Dishevelled Segment Polarity Protein 3, DVL3) (Schlange et al., 2007). Thus, as in the ovary, active testicular genes were generally associated with immune response and cell replication.

The skin. Not surprisingly, the majority of the most abundant transcripts that were uniquely abundant in skin were keratins (Keratin 71 Type II: KRT71 TPM= 5248.41, Keratin Associated Protein 3-1 Type II: KRTAP3-1 TPM= 1696.92, Kerain 83: KRT83 TPM= 941.12, Keratin 73 Type II: KRT73 TPM= 747.05, Keratin 85 Type II: KRT85 TPM= 495.35, Keratin type II cytoskeletal 5: KRT5 TPM= 446.98), the proteins that comprise the protective external layer for epithelial cells (Bragulla & Homberger, 2009; Deek et al., 2016). Highly abundant skin transcripts also include genes involved in muscle movement (Myosin Light Chain 1; MYL1 TPM= 345.10, Troponin T3; TNNT3 TPM= 190.25) (Periasamy et al., 1984; Ling et al., 2010; Wei & Jin, 2016). Genes found to be under positive selection in skin have been associated with tumor suppression (UBS Domain Protein 1; UBXN1) (Wu-Baer, Ludwig & Baer, 2010) and repair of double-stranded DNA (Heterogeneous Nuclear Ribonucleoprotein U Like 1 (HNRNPUL1) (Polo et al., 2012).

The kidney. Two well-documented functions of renal tissue are the transport of nutrients and the secretion of urine (Wang & Giebisch, 2009; Bobulescu & Moe, 2012). Consistent with this, uniquely abundant transcripts identified in the kidney included solute carriers SLC34A1 (TPM= 433.72) and SLC14A2 (TPM= 114.25), which are involved in transport of nutrients and urea (Shayakul, Clémençon & Hediger, 2013; Martovetsky, Bush & Nigam, 2016). Among those genes displaying signatures of positive selection in the kidney were Suppressor of Ty 3 (SUPT3), which binds p53 during DNA repair (Martinez et al., 2001; Gamper & Roeder, 2008) and N-Myc Downstream Regulated 1 (NDRG1), which is involved in suppression of metastasis, particularly under hypoxic conditions (Salnikow et al., 2002; Mao et al., 2013).

The spleen. Uniquely abundant transcripts in the spleen tended to encompass more functional diversity than transcripts identified for the other tissues sampled. Highly abundant spleen-specific transcripts

include proteins involved in nucleotide exchange (ARHGEF17, TPM= 137.25), erythropoiesis (EPOR TPM= 26.56, SPTA1 TPM= 21.82), and GTP hydrolysis (GBP6, TPM= 22.15), as well as at least one kinase (LIMK2, TPM= 10.09) that is associated with immune function (Bernard, 2007; Kim et al., 2011; Lutz et al., 2013; Ponceau et al., 2015; Kuhrt & Wojchowski, 2015). Both erythrocytic activity and immune function are consistent with the functional role of the spleen, which filters blood and recycles blood cells (Cesta, 2006; Scott & Olson, 2007; Droppelmann et al., 2013; Pivkin et al., 2016). Interestingly, the spleen was found to express more genes under positive selection than expected (Fig. 1), suggesting this tissue may be an active target for adaptation. Three of these genes (Sperm Associated Antigen 9; SPAG9, Cell Division Cycle 7; CDC7, and Zinc Finger CCCH-Type Containing 13; ZC3H13) have been previously characterized in humans. Upregulated in cancerous cells, SPAG9 is thought to be an early marker for diagnosis (Baser et al., 2013; Chen et al., 2014). Cell Division Cycle 7 is a DNA replication regulator, and can inactivate tumor suppressor protein p53 when CDC7 is overexpressed in tumor cells (Bonte et al., 2008; Ito et al., 2012). Finally, ZC3H13 is a component of Wilms' tumor associating protein, a splicing regulator potentially required for cell cycle progression (Horiuchi et al., 2013).

The liver. The primary functions of the liver are to produce blood coagulation hormones, detoxify blood, and to metabolize foreign substances (Cheeke, 1994; Wada, Usui & Sakuragawa, 2008; Davidson, Ballinger & Khetani, 2016; Schiöth et al., 2016; Harrall et al., 2016). The two genes that were most uniquely expressed in the liver were associated with these functions, specifically blood clotting (Fibrinogen Alpha Chain: FGA, TPM= 864.02), and drug toxin metabolism (Cytochrome P450 2A11: CYP2A11, TPM= 370.43) (Mosesson, 2005; Yang et al., 2012). Our results suggest that the liver, like the spleen, may also be an active site of adaptation given the number of genes found to be under positive selection in the liver was more than twice that expected by chance (Fig. 1). Of these genes, three are involved in metal ion transport (Solute Carrier Family 30 Member 10 [SLC30A10], Nedd4 Family Interacting Protein 2 [NDFIP2], and Family With Sequence Similarity 21 Member C [FAM21]) (Ohana et al., 2006; Yang et al., 2012; Shusterman et al., 2014; Gallon & Cullen, 2015; Lee, Chang & Blackstone, 2016; Foot et al., 2016), while three others have ontology terms associated with immune response (Signal Peptide Peptidase Like 2A [SPPL2A: Biological Process- regulation of immune response], Ataxin 2 [ATXN2: Biological Process - negative regulation of multicellular organism growth], SET Domain Containing 6 [SETD6: Biological Process- regulation of inflammatory response]). Given the roles that the spleen and liver play in immunological processes and the genes identified to be under positive selection in these tissues, it is possible that both the spleen and liver of the tuco-tuco are particularly involved in adaptation to the subterranean environment.

C. sociabilis is not the first subterranean rodent to provide evidence of possible adaptation to the regulation of cell cycling. The naked mole-rat (*H. glaber)*, has been the subject of numerous studies attempting to discern the source of the cancer resistance reported for this long-lived species (Buffenstein, 2008; Rodriguez et al., 2011; Delaney et al., 2013). Decreased prevalence of cancer in the naked mole rat has been attributed to a heightened sensitivity to contact inhibition (Seluanov et al., 2009) and fibroblast secretion of high-molecular-mass hyaluronan (Tian et al., 2013). Studies have also suggested that the naked mole rat has increased translational fidelity due to a unique 28S ribosomal structure (Azpurua et al., 2013). More recently, cancer has been detected in this species (Delaney et al., 2016), although these examples were based on studies of captive mole-rats not exposed to the natural hypoxic environment for this species, an environmental setting that may have contributed to tumor formation (Welsh & Traum, 2016). Colonial tuco-tucos also presumably occur in hypoxic environments and it is possible that the fourteen apoptotic genes identified as being subject to positive selection in this species also have important regulatory functions in this setting. Gene ontology terms associated with cell cycling/DNA damage response genes comprised over 20% (12 genes of 50) of the genes identified as being under positive selection, with other gene ontology categories comprising a substantially smaller portion of the loci thought to be subject to selection. Collectively, these genes present important candidates for future studies of regulation of cell physiology in subterranean rodents.

Future studies of *C. sociabilis* would benefit from quantifying differential gene expression across multiple individuals to provide a more robust quantitative assessment of tissue-specific patterns of gene expression. Of the highly abundant transcripts identified for each tissue type, many suggest a role in immune function while positively selected genes hint at specializations for cell cycle regulation. Both of these characteristics are seen across the different tissues samples for *C. sociabilis*. Expression patterns can vary greatly among individuals, and thus although our data set does not allow for statistical analyses of patterns of gene expression in *C. sociabilis*, our findings are consistent with those revealed by previous studies of subterranean organisms. Expansion of our analyses to include multiple individuals, as well as additional taxa, will allow for a more comprehensive understanding of the genomic underpinnings of physiological adaptations to subterranean life.

SUMMARY

In this study, we present a high quality and complete transcriptome for the colonial tuco-tuco (*C. sociabilis)*. By characterizing transcriptomes generated from eight tissue types, we provide preliminary

insights into how transcript abundance differs across tissues. Notably, the most abundant transcripts and the genes subject to positive selection were generally consistent with the primary physiological function(s) of the tissues from which they were derived, with a prevalence of transcripts associated with cell proliferation. We also identify a set of genes that appear to be under positive selection; the number of genes subject to selection that were expressed in the liver and spleen were greater than expected, suggesting that these tissues are of particular functional importance to the colonial tuco-tucos. The underlying reasons for enhanced selection of genes in these tissues remains to be determined, providing an intriguing basis for additional studies of genomic evolution in *C. sociabilis* and other subterranean rodents. At the same time, given extensive field data regarding the behavior, ecology, and physiology of *C. sociabilis*, the transcriptomic data presented here represent a critical tool for future studies aimed at clarifying relationships among physiology, selection, and specialization for a subterranean lifestyle.

Assembly	Num. Reads	Num. Contigs	Assembly Size	Transrate Score	BUSCO Metrics
good_compiled_50M_transfuse	50M	157996	240Mb	0.430	C: 88%, D: 64%, M: 8%
annotation_only	50M	96224	227Mb	0.420	C: 88%, D: 64%, M: 8%

Table 1. A comparison of assemblies utilizing metrics for quality and completeness. (Num. Reads = Number of Reads, Num. Contigs = Number of Contigs, Assembly size, TransRate score, and BUSCO Metrics: C = Complete, D = Duplicated, M = Missing BUSCOs). The good_compiled_50M_transfuse assembly was chosen for annotation, and the annotation_only assembly is the transcriptome we present as our finalized assembly.

	good_compiled_50M_transfuse		Annotation_only	
Tissue	% mapped	% prop paired	% mapped	% prop paired
hippocampus	87.51	80.28	85.18	78.51
hypothalamus	86.40	79.10	83.98	77.29
kidney	86.74	77.05	84.53	75.47
liver	87.79	79.68	85.92	78.50
ovary	84.59	74.62	81.57	72.42
testes	85.24	77.00	82.86	75.32
skin	85.82	73.42	83.37	71.71
spleen	90.67	83.53	88.89	82.24
Average	86.85	78.09	84.54	76.43

Table 2. Burrows-Wheeler Aligner mapping statistics comparing the percent mapping and percent properly paired mapping rates of the annotated assembly (annotated good_compiled_50M_transfuse) and the final assembly (Annotation.only).

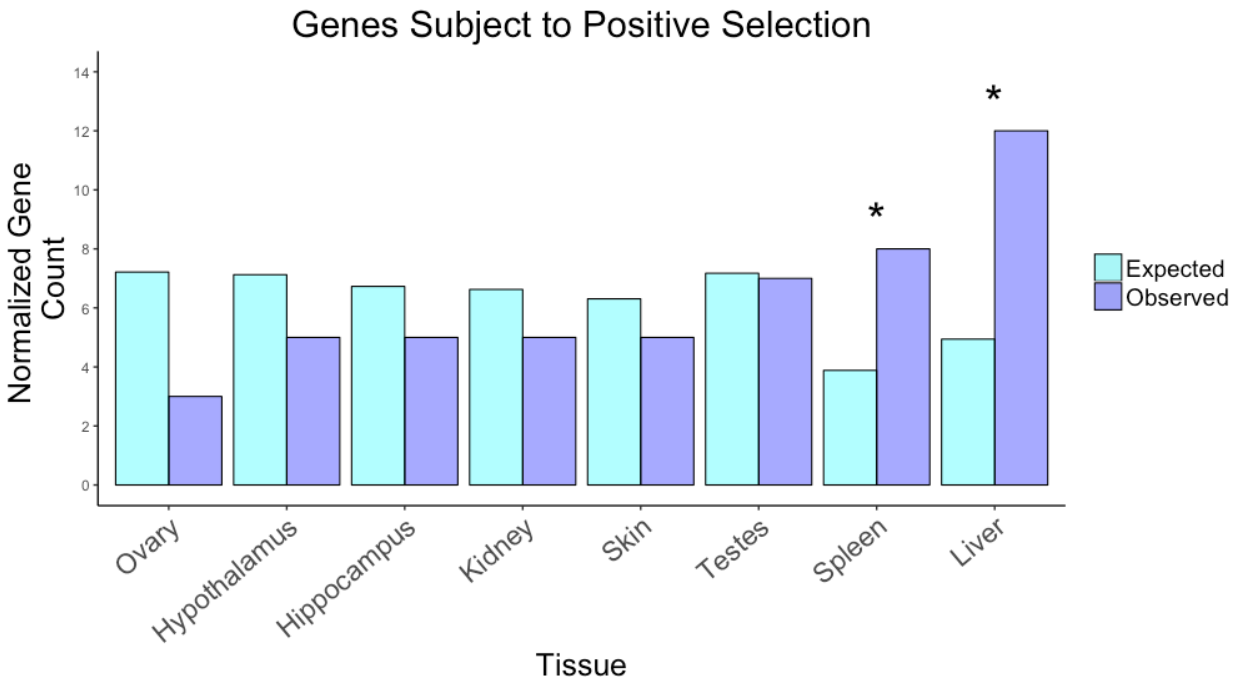

Figure 1. Tissue-specific counts of the 50 positively selected genes detected, normalized by the total number of genes present in each tissue. Tissue types are indicated on the x-axis. Expected abundance of positively selected genes is depicted by light blue bars; observed abundance of positively selected genes is shown in dark blue. Asterisks denote statistically significant differences between expected and observed values (Chi-square tests, $p < 0.05$).

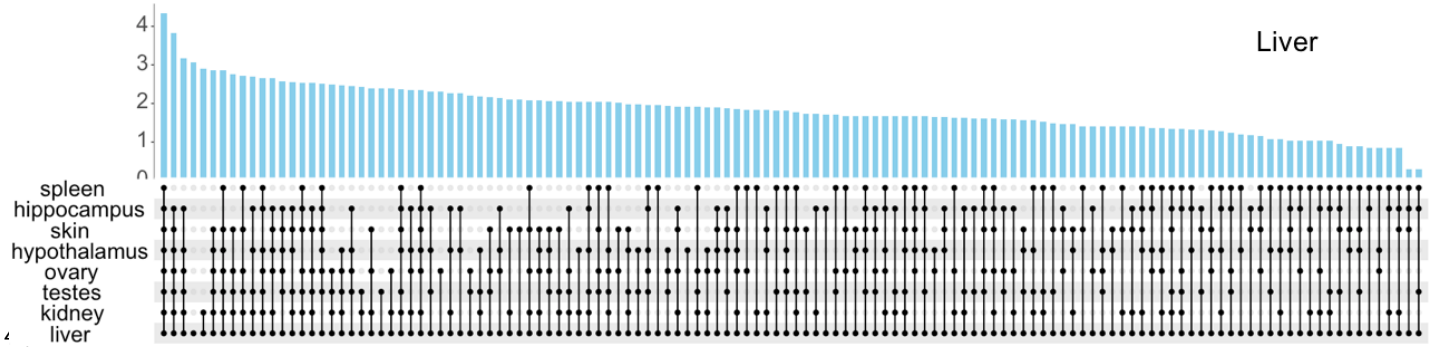

Figure 2. Comparing transcript composition of the liver to other tissues. The x-axis depicts intersections between tissue types, and the y-axis is the $\log_{10}$ transformation of normalized transcript counts. The 128 Intersection groups have been arranged to present groups with the highest transcript counts to the left, and lowest counts to the right. Figures depicting transcript composition of the remaining tissues can be found in supplemental materials (Figures S1 & S2).

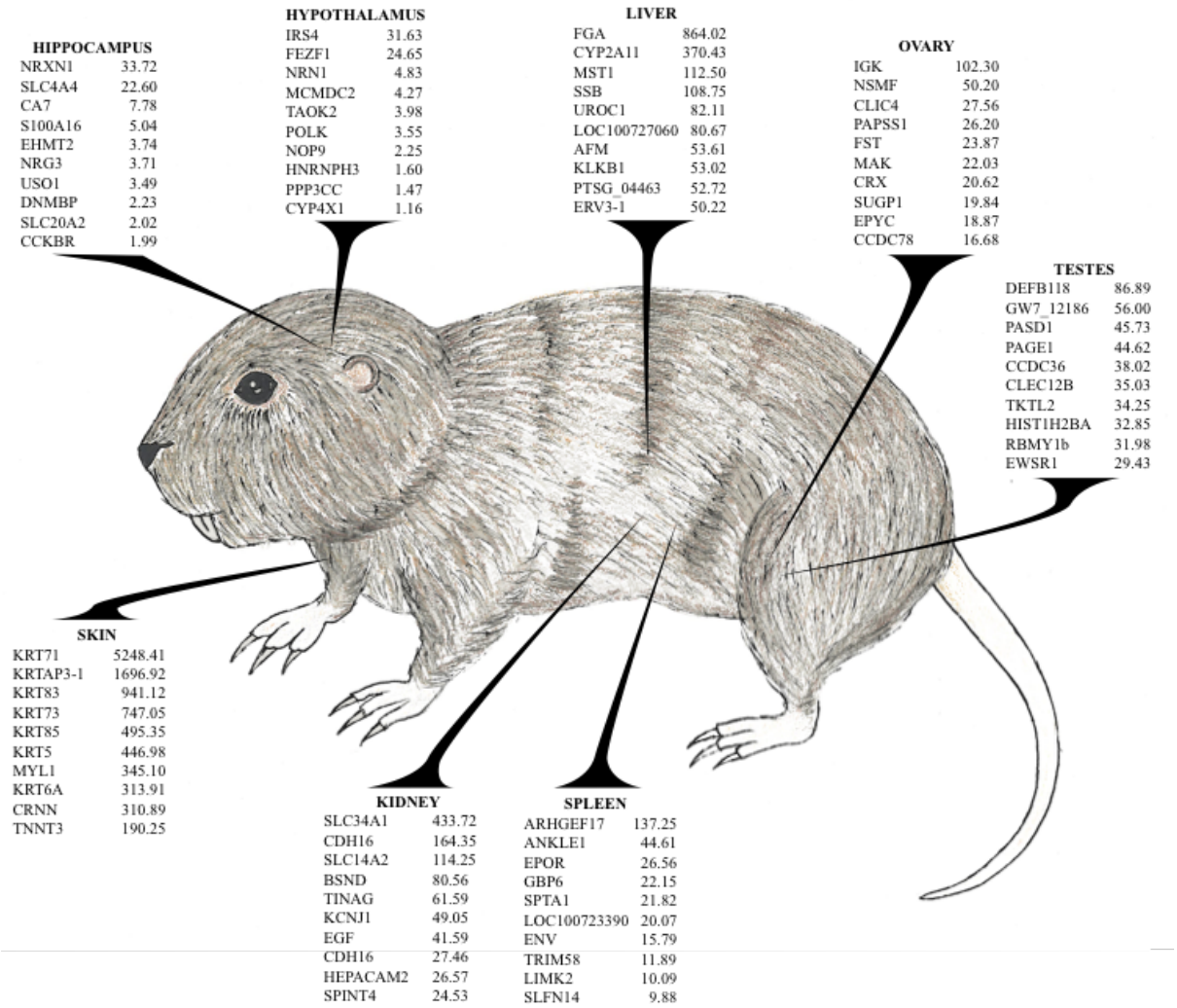

Figure 3. Ten most abundant unique transcripts for each tissue type. For each tissue type, the left column is the gene ID, while the right column contains the associated TPM values.

510 Literature Cited

511 Alves MG., Rato L., Carvalho RA., Moreira PI., Socorro S., Oliveira PF. 2013. Hormonal control of
512 Sertoli cell metabolism regulates spermatogenesis. *Cellular and molecular life sciences: CMLS*
513 70:777–793.
514 Azpurua J., Ke Z., Chen IX., Zhang Q., Ermolenko DN., Zhang ZD., Gorbunova V., Seluanov A. 2013.
515 Naked mole-rat has increased translational fidelity compared with the mouse, as well as a unique
516 28S ribosomal RNA cleavage. *Proceedings of the National Academy of Sciences of the United States*
517 *of America* 110:17350–17355.
518 Bannerman DM., Deacon RMJ., Offen S., Friswell J., Grubb M., Rawlins JNP. 2002. Double dissociation
519 of function within the hippocampus: spatial memory and hyponeophagia. *Behavioral neuroscience*
520 116:884–901.
521 Baser E., Togrul C., Ozgu E., Ayhan S., Caglar M., Erkaya S., Gungor T. 2013. Sperm-associated antigen
522 9 is a promising marker for early diagnosis of endometrial cancer. *Asian Pacific journal of cancer*
523 *prevention: APJCP* 14:7635–7638.
524 Beery AK., Lacey EA., Francis DD. 2008. Oxytocin and vasopressin receptor distributions in a solitary
525 and a social species of tuco-tuco (Ctenomys haigi and Ctenomys sociabilis). *The Journal of*
526 *comparative neurology* 507:1847–1859.
527 Benjamini Y, Hochberg Y 1995. Controlling the False Discovery Rate: A Practical and Powerful
528 Approach to Multiple Testing. *Journal of the Royal Statistical Society* 57:289–300.
529 Bernard O. 2007. Lim kinases, regulators of actin dynamics. *The international journal of biochemistry &*
530 *cell biology* 39:1071–1076.
531 Bhattacharya N., Wang Z., Davitt C., McKenzie IFC., Xing P-X., Magnuson NS. 2002. Pim-1 associates
532 with protein complexes necessary for mitosis. *Chromosoma* 111:80–95.
533 Bobulescu IA., Moe OW. 2012. Renal transport of uric acid: evolving concepts and uncertainties.
534 *Advances in chronic kidney disease* 19:358–371.

535 Bolger AM., Lohse M., Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data.
536 *Bioinformatics* 30:2114–2120.
537 Bonte D., Lindvall C., Liu H., Dykema K., Furge K., Weinreich M. 2008. Cdc7-Dbf4 kinase
538 overexpression in multiple cancers and tumor cell lines is correlated with p53 inactivation. *Neoplasia*
539 10:920–931.
540 Bragulla HH., Homberger DG. 2009. Structure and functions of keratin proteins in simple, stratified,
541 keratinized and cornified epithelia. *Journal of anatomy* 214:516–559.
542 Bray NL., Pimentel H., Melsted P., Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification.
543 *Nature biotechnology* 34:525–527.
544 Brekke KM., Garrard WT. 2004. Assembly and analysis of the mouse immunoglobulin kappa gene
545 sequence. *Immunogenetics* 56:490–505.
546 Brodsky LI., Jacob-Hirsch J., Avivi A., Trakhtenbrot L., Zeligson S., Amariglio N., Paz A., Korol AB.,
547 Band M., Rechavi G., Nevo E. 2005. Evolutionary regulation of the blind subterranean mole rat,
548 Spalax, revealed by genome-wide gene expression. *Proceedings of the National Academy of*
549 *Sciences* 102:17047–17052.
550 Bruno T., De Angelis R., De Nicola F., Barbato C., Di Padova M., Corbi N., Libri V., Benassi B., Mattei
551 E., Chersi A., Soddu S., Floridi A., Passananti C., Fanciulli M. 2002. Che-1 affects cell growth by
552 interfering with the recruitment of HDAC1 by Rb. *Cancer cell* 2:387–399.
553 Bruno T., Iezzi S., Fanciulli M. 2016. Che-1/AATF: A Critical Cofactor for Both Wild-Type- and
554 Mutant-p53 Proteins. *Frontiers in oncology* 6:34.
555 Bruyninx M., Hennuy B., Cornet A., Houssa P., Daukandt M., Reiter E., Poncin J., Closset J., Hennen G.
556 1999. A novel gene overexpressed in the prostate of castrated rats: hormonal regulation, relationship
557 to apoptosis and to acquired prostatic cell androgen independence. *Endocrinology* 140:4789–4799.
558 Buffenstein R. 2000. Ecophysiological responses of subterranean rodents to underground habitats. *Life*
559 *underground: the biology of subterranean rodents (EA Lacey, JL Patton, and GN Cameron, eds.).*
560 *University of Chicago Press, Illinois*:62–110.

561 Buffenstein R. 2008. Negligible senescence in the longest living rodent, the naked mole-rat: insights from
562 a successfully aging species. *Journal of comparative physiology. B, Biochemical, systemic, and*
563 *environmental physiology* 178:439–445.
564 Burda H., Šumbera R., Begall S. 2007. Microclimate in Burrows of Subterranean Rodents — Revisited.
565 In: *Subterranean Rodents*. Springer, Berlin, Heidelberg, 21–33.
566 Cesta MF. 2006. Normal structure, function, and histology of the spleen. *Toxicologic pathology* 34:455–
567 465.
568 Chan YL., Hadly EA. 2011. Genetic variation over 10,000 years in Ctenomys: comparative
569 phylochronology provides a temporal perspective on rarity, environmental change and demography.
570 *Molecular ecology* 20:4592–4605.
571 Chan YL., Lacey EA., Pearson OP., Hadly EA. 2005. Ancient DNA reveals Holocene loss of genetic
572 diversity in a South American rodent. *Biology letters* 1:423–426.
573 Cheeke PR. 1994. A review of the functional and evolutionary roles of the liver in the detoxification of
574 poisonous plants, with special reference to pyrrolizidine alkaloids. *Veterinary and human toxicology*
575 36:240–247.
576 Chen ME., Lin SH., Chung LW., Sikes RA. 1998. Isolation and characterization of PAGE-1 and GAGE-
577 7. New genes expressed in the LNCaP prostate cancer progression model that share homology with
578 melanoma-associated antigens. *The Journal of biological chemistry* 273:17618–17625.
579 Chen F., Lu Z., Deng J., Han X., Bai J., Liu Q., Xi Y., Zheng J. 2014. SPAG9 expression is increased in
580 human prostate cancer and promotes cell motility, invasion and angiogenesis in vitro. *Oncology*
581 *reports* 32:2533–2540.
582 Clément F. 2016. Multiscale mathematical modeling of the hypothalamo-pituitary-gonadal axis.
583 *Theriogenology* 86:11–21.
584 Cooper CDO., Liggins AP., Ait-Tahar K., Roncador G., Banham AH., Pulford K. 2006. PASD1, a
585 DLBCL-associated cancer testis antigen and candidate for lymphoma immunotherapy. *Leukemia*
586 20:2172–2174.

587 Cornet AM., Hanon E., Reiter ER., Bruyninx M., Nguyen VH., Hennuy BR., Hennen GP., Closset JL.
588 2003. Prostatic androgen repressed message-1 (PARM-1) may play a role in prostatic cell
589 immortalisation. *The Prostate* 56:220–230.
590 Cox MP., Peterson DA., Biggs PJ. 2010. SolexaQA: At-a-glance quality assessment of Illumina second-
591 generation sequencing data. *BMC bioinformatics* 11:485.
592 Crusoe MR., Alameldin HF., Awad S., Boucher E., Caldwell A., Cartwright R., Charbonneau A.,
593 Constantinides B., Edvenson G., Fay S., Fenton J., Fenzl T., Fish J., Garcia-Gutierrez L., Garland P.,
594 Gluck J., González I., Guermond S., Guo J., Gupta A., Herr JR., Howe A., Hyer A., Härpfer A., Irber
595 L., Kidd R., Lin D., Lippi J., Mansour T., McA'Nulty P., McDonald E., Mizzi J., Murray KD.,
596 Nahum JR., Nanlohy K., Nederbragt AJ., Ortiz-Zuazaga H., Ory J., Pell J., Pepe-Ranney C., Russ
597 ZN., Schwarz E., Scott C., Seaman J., Sievert S., Simpson J., Skennerton CT., Spencer J., Srinivasan
598 R., Standage D., Stapleton JA., Steinman SR., Stein J., Taylor B., Trimble W., Wiencko HL., Wright
599 M., Wyss B., Zhang Q., Zyme E., Brown CT. 2015. The khmer software package: enabling efficient
600 nucleotide sequence analysis. *F1000Research* 4:900.
601 Cutrera AP., Zenuto RR., Luna F., Antenucci CD. 2010. Mounting a specific immune response increases
602 energy expenditure of the subterranean rodent Ctenomys talarum (tuco-tuco): implications for
603 intraspecific and interspecific variation in immunological traits. *The Journal of experimental biology*
604 213:715–724.
605 Davidson MD., Ballinger KR., Khetani SR. 2016. Long-term exposure to abnormal glucose levels alters
606 drug metabolism pathways and insulin sensitivity in primary human hepatocytes. *Scientific reports*
607 6:28178.
608 Deek J., Hecht F., Rossetti L., Wißmiller K., Bausch AR. 2016. Mechanics of soft epithelial keratin
609 networks depend on modular filament assembly kinetics. *Acta biomaterialia* 43:218–229.
610 Delaney MA., Nagy L., Kinsel MJ., Treuting PM. 2013. Spontaneous histologic lesions of the adult naked
611 mole rat (Heterocephalus glaber): a retrospective survey of lesions in a zoo population. *Veterinary*
612 *pathology* 50:607–621.

613 Delaney MA., Ward JM., Walsh TF., Chinnadurai SK., Kerns K., Kinsel MJ., Treuting PM. 2016. Initial
614 Case Reports of Cancer in Naked Mole-rats (Heterocephalus glaber). *Veterinary pathology* 53:691–
615 696.
616 Devi K., Mishra SK., Sahu J., Panda D., Modi MK., Sen P. 2016. Genome wide transcriptome profiling
617 reveals differential gene expression in secondary metabolite pathway of Cymbopogon winterianus.
618 *Scientific reports* 6:21026.
619 De Vleeschouwer F., Vanneste H., Mauquoy D., Piotrowska N., Torrejón F., Roland T., Stein A., Le
620 Roux G. 2014. Emissions from pre-Hispanic metallurgy in the South American atmosphere. *PloS*
621 *one* 9:e111315.
622 Droppelmann CA., Keller BA., Campos-Melo D., Volkening K., Strong MJ. 2013. Rho guanine
623 nucleotide exchange factor is an NFL mRNA destabilizing factor that forms cytoplasmic inclusions
624 in amyotrophic lateral sclerosis. *Neurobiology of aging* 34:248–262.
625 Eichenbaum H. 2017. The role of the hippocampus in navigation is memory. *Journal of neurophysiology*
626 117:1785–1796.
627 Fernández-Cadena JC., Andrade S., Silva-Coello CL., De la Iglesia R. 2014. Heavy metal concentration
628 in mangrove surface sediments from the north-west coast of South America. *Marine pollution*
629 *bulletin* 82:221–226.
630 Foot NJ., Gembus KM., Mackenzie K., Kumar S. 2016. Ndfip2 is a potential regulator of the iron
631 transporter DMT1 in the liver. *Scientific reports* 6:24045.
632 Fu X., Yucer N., Liu S., Li M., Yi P., Mu J-J., Yang T., Chu J., Jung SY., O'Malley BW., Gu W., Qin J.,
633 Wang Y. 2010. RFWD3-Mdm2 ubiquitin ligase complex positively regulates p53 stability in
634 response to DNA damage. *Proceedings of the National Academy of Sciences of the United States of*
635 *America* 107:4579–4584.
636 Gallon M., Cullen PJ. 2015. Retromer and sorting nexins in endosomal sorting. *Biochemical Society*
637 *transactions* 43:33–47.
638 Gamper AM., Roeder RG. 2008. Multivalent binding of p53 to the STAGA complex mediates coactivator

639 recruitment after UV damage. *Molecular and cellular biology* 28:2517–2527.

640 Gardiner A., Barker D., Butlin RK., Jordan WC., Ritchie MG. 2008. Drosophila chemoreceptor gene

641 evolution: selection, specialization and genome size. *Molecular ecology* 17:1648–1657.

642 Gardner AL., Wilson DE., Reeder DM. 2005. Mammal species of the world: a taxonomic and geographic

643 reference. *Mammal species of the world: a taxonomic and geographic reference* 12.

644 Gehlenborg N. 2016.UpSetR: A More Scalable Alternative to Venn and Euler Diagrams for Visualizing

645 Intersecting Sets

646 Gong Z., Chen J. 2011. E3 ligase RFWD3 participates in replication checkpoint control. *The Journal of*

647 *biological chemistry* 286:22308–22313.

648 Griffiths-Jones S., Moxon S., Marshall M., Khanna A., Eddy SR., Bateman A. 2005. Rfam: annotating

649 non-coding RNAs in complete genomes. *Nucleic acids research* 33:D121–4.

650 Haas BJ., Papanicolaou A., Yassour M., Grabherr M., Blood PD., Bowden J., Couger MB., Eccles D., Li

651 B., Lieber M., MacManes MD., Ott M., Orvis J., Pochet N., Strozzi F., Weeks N., Westerman R.,

652 William T., Dewey CN., Henschel R., Leduc RD., Friedman N., Regev A. 2013. De novo transcript

653 sequence reconstruction from RNA-seq using the Trinity platform for reference generation and

654 analysis. *Nature protocols* 8:1494–1512.

655 Hall JMF., Cruser D., Podawiltz A., Mummert DI., Jones H., Mummert ME. 2012. Psychological Stress

656 and the Cutaneous Immune Response: Roles of the HPA Axis and the Sympathetic Nervous System

657 in Atopic Dermatitis and Psoriasis. *Dermatology research and practice* 2012:403908.

658 Hambuch TM., Lacey EA. 2002. Enhanced selection for MHC diversity in social tuco-tucos. *Evolution;*

659 *international journal of organic evolution* 56:841–845.

660 Harrall KK., Kechris KJ., Tabakoff B., Hoffman PL., Hines LM., Tsukamoto H., Pravenec M., Printz M.,

661 Saba LM. 2016. Uncovering the liver's role in immunity through RNA co-expression networks.

662 *Mammalian genome: official journal of the International Mammalian Genome Society* 27:469–484.

663 Havens LA., MacManes MD. 2016. Characterizing the adult and larval transcriptome of the multicolored

664 Asian lady beetle, Harmonia axyridis. *PeerJ* 4:e2098.

665 Horiuchi K., Kawamura T., Iwanari H., Ohashi R., Naito M., Kodama T., Hamakubo T. 2013.
666 Identification of Wilms' tumor 1-associating protein complex and its role in alternative splicing and
667 the cell cycle. *The Journal of biological chemistry* 288:33292–33302.
668 Ito S., Ishii A., Kakusho N., Taniyama C., Yamazaki S., Fukatsu R., Sakaue-Sawano A., Miyawaki A.,
669 Masai H. 2012. Mechanism of cancer cell death induced by depletion of an essential replication
670 regulator. *PloS one* 7:e36372.
671 Juhasz S., Balogh D., Hajdu I., Burkovics P., Villamil MA., Zhuang Z., Haracska L. 2012.
672 Characterization of human Spartan/C1orf124, an ubiquitin-PCNA interacting regulator of DNA
673 damage tolerance. *Nucleic acids research* 40:10795–10808.
674 Karn RC, Clark NL, Nguyen ED, Swanson WJ 2008. Adaptive evolution in rodent seminal vesicle
675 secretion proteins. *Molecular biology and evolution* 25:2301–2310.
676 Kim B-H., Shenoy AR., Kumar P., Das R., Tiwari 2. Sangeeta., Mac Micking JD. 2011. A Family of
677 IFN-g–Inducible 65-kD GTPases Protects Against Bacterial Infection. *Science* 332.
678 Kimura F., Bonomi LM., Schneyer AL. 2011. Follistatin regulates germ cell nest breakdown and
679 primordial follicle formation. *Endocrinology* 152:697–706.
680 Kong F., Su Z., Zhou C., Sun C., Liu Y., Zheng D., Yuan H., Yin J., Fang J., Wang S., Xu H. 2011. Role
681 of positive selection in functional divergence of mammalian neuronal apoptosis inhibitor proteins
682 during evolution. *Journal of biomedicine & biotechnology* 2011:809765–809768.
683 Kordonowy LL., MacManes MD. 2016. Characterization of a male reproductive transcriptome for
684 Peromyscus eremicus (Cactus mouse). *PeerJ* 4:e2617.
685 Kosiol C, Vinar T, da Fonseca RR, Hubisz MJ, Bustamante CD, Nielsen R, Siepel A 2008. Patterns of
686 positive selection in six Mammalian genomes. *PLoS genetics* 4:e1000144.
687 Kudryashova E., Lu W., Kudryashov DS. 2015. Defensins versus pathogens: an unfolding story.
688 *Oncotarget* 6:28533–28534.
689 Kuhrt D., Wojchowski DM. 2015. Emerging EPO and EPO receptor regulators and signal transducers.
690 *Blood* 125:3536–3541.

691 Kumar D., Thakur MK. 2015. Age-related expression of Neurexin1 and Neuroligin3 is correlated with
692 presynaptic density in the cerebral cortex and hippocampus of male mice. *Age* 37:17–10.
693 Lacey EA. 2001. Microsatellite variation in solitary and social tuco-tucos: molecular properties and
694 population dynamics. *Heredity* 86:628–637.
695 Lacey EA., Braude SH., Wieczorek JR. 1997. Burrow Sharing by Colonial Tuco-Tucos (Ctenomys
696 sociabilis). *Journal of mammalogy* 78:556–562.
697 Lacey EA., Patton JL. 2000. *Life Underground: The Biology of Subterranean Rodents*. University of
698 Chicago Press.
699 Lacey EA., Wieczorek JR. 2004. Kinship in colonial tuco-tucos: evidence from group composition and
700 population structure. *Behavioral ecology: official journal of the International Society for Behavioral*
701 *Ecology* 15:988–996.
702 Lee S., Chang J., Blackstone C. 2016. FAM21 directs SNX27-retromer cargoes to the plasma membrane
703 by preventing transport to the Golgi apparatus. *Nature communications* 7:10939.
704 Ling F., Fang W., Chen Y., Li J., Liu X., Wang L., Zhang H., Chen S., Mei Y., Du H., Wang C. 2010.
705 Identification of novel transcripts from the porcine MYL1 gene and initial characterization of its
706 promoters. *Molecular and cellular biochemistry* 343:239–247.
707 Lin G-H., Wang K., Deng X-G., Nevo E., Zhao F., Su J-P., Guo S-C., Zhang T-Z., Zhao H. 2014.
708 Transcriptome sequencing and phylogenomic resolution within Spalacidae (Rodentia). *BMC*
709 *genomics* 15:32.
710 Liu S., Chu J., Yucer N., Leng M., Wang S-Y., Chen BPC., Hittelman WN., Wang Y. 2011. RING finger
711 and WD repeat domain 3 (RFWD3) associates with replication protein A (RPA) and facilitates RPA-
712 mediated DNA damage response. *The Journal of biological chemistry* 286:22314–22322.
713 Liu J., Li G., Chang Z., Yu T., Liu B., McMullen R., Chen P., Huang X. 2016. BinPacker: Packing-Based
714 De Novo Transcriptome Assembly from RNA-seq Data. *PLoS computational biology* 12:e1004772.
715 Lovegrove BG. 1986.The Metabolism of Social Subterranean Rodents: Adaptation to Aridity
716 Luna F., Antinuchi CD. 2006. Cost of foraging in the subterranean rodent Ctenomys talarum: effect of

717 soil hardness. *Canadian journal of zoology* 84:661–667.

718 Lutz S., Mohl M., Rauch J., Weber P., Wieland T. 2013. RhoGEF17, a Rho-specific guanine nucleotide
719 exchange factor activated by phosphorylation via cyclic GMP-dependent kinase Iα. *Cellular*
720 *signalling* 25:630–638.

721 Machida YJ., Machida Y., Vashisht AA., Wohlschlegel JA., Dutta A. 2009. The deubiquitinating enzyme
722 BAP1 regulates cell growth via interaction with HCF-1. *The Journal of biological chemistry*
723 284:34179–34188.

724 MacManes MD 2017. Severe acute dehydration in a desert rodent elicits a transcriptional response that
725 effectively prevents kidney injury. *American journal of physiology. Renal physiology* 313:F262–
726 F272.

727 MacManes MD, Austin SH, Lang AS, Booth A, Farrar V, Calisi RM 2017. Widespread patterns of
728 sexually dimorphic gene expression in an avian hypothalamic-pituitary-gonadal (HPG) axis.
729 *Scientific reports* 7:45125.

730 MacManes MD. 2015. *Establishing evidenced-based best practice for the de novo assembly and*
731 *evaluation of transcriptomes from non-model organisms*.

732 MacManes MD., Eisen MB. 2014. Characterization of the transcriptome, nucleotide sequence
733 polymorphism, and natural selection in the desert adapted mouse Peromyscus eremicus. *PeerJ*
734 2:e642.

735 MacManes MD 2014. On the optimal trimming of high-throughput mRNA sequence data. Frontiers in
736 genetics 5:13.

737 MacManes MD., Lacey EA. 2012. The Social Brain: Transcriptome Assembly and Characterization of
738 the Hippocampus from a Social Subterranean Rodent, the Colonial Tuco-Tuco (Ctenomys
739 sociabilis). *PloS one* 7:e45524–8.

740 Malik A., Korol A., Hübner S., Hernandez AG., Thimmapuram J., Ali S., Glaser F., Paz A., Avivi A.,
741 Band M. 2011. Transcriptome sequencing of the blind subterranean mole rat, Spalax galili: utility
742 and potential for the discovery of novel evolutionary patterns. *PloS one* 6:e21227.

743 Mao Z., Sun J., Feng B., Ma J., Zang L., Dong F., Zhang D., Zheng M. 2013. The metastasis suppressor,
744 N-myc downregulated gene 1 (NDRG1), is a prognostic biomarker for human colorectal cancer.
745 *PloS one* 8:e68206.
746 Mares MA. 1975. South American mammal zoogeography: evidence from convergent evolution in desert
747 rodents. *Proceedings of the National Academy of Sciences* 72:1702–1706.
748 Martinez E., Palhan VB., Tjernberg A., Lymar ES., Gamper AM., Kundu TK., Chait BT., Roeder RG.
749 2001. Human STAGA complex is a chromatin-acetylating transcription coactivator that interacts
750 with pre-mRNA splicing and DNA damage-binding factors in vivo. *Molecular and cellular biology*
751 21:6782–6795.
752 Martovetsky G., Bush KT., Nigam SK. 2016. Kidney versus Liver Specification of SLC and ABC Drug
753 Transporters, Tight Junction Molecules, and Biomarkers. *Drug metabolism and disposition: the*
754 *biological fate of chemicals* 44:1050–1060.
755 McCarthy DJ., Chen Y., Smyth GK. 2012. Differential expression analysis of multifactor RNA-Seq
756 experiments with respect to biological variation. *Nucleic acids research* 40:4288–4297.
757 Merlo JL., Cutrera AP., Zenuto RR. 2016. Parasite infection negatively affects PHA-triggered
758 inflammation in the subterranean rodent Ctenomys talarum. *Journal of experimental zoology. Part A,*
759 *Ecological genetics and physiology* 325:132–141.
760 Moreton J., Izquierdo A., Emes RD. 2016. Assembly, Assessment, and Availability of De novo
761 Generated Eukaryotic Transcriptomes. *Frontiers in genetics* 6:3389–3389.
762 Mosesson MW. 2005. Fibrinogen and fibrin structure and functions. *Journal of thrombosis and*
763 *haemostasis: JTH* 3:1894–1904.
764 Muschick M., Indermaur A., Salzburger W. 2012. Convergent evolution within an adaptive radiation of
765 cichlid fishes. *Current biology: CB* 22:2362–2368.
766 Nevo E. 1979. Adaptive Convergence and Divergence of Subterranean Mammals. *Annual review of*
767 *ecology and systematics* 10:269–308.
768 Novikov E., Petrovski D., Mak V., Kondratuk E., Krivopalov A., Moshkin M. 2016. Variability of

769 whipworm infection and humoral immune response in a wild population of mole voles (Ellobius
770 talpinus Pall.). *Parasitology research* 115:2925–2932.
771 Numan S., Russell DS. 1999. Discrete expression of insulin receptor substrate-4 mRNA in adult rat brain.
772 *Brain research. Molecular brain research* 72:97–102.
773 Ohana E., Sekler I., Kaisman T., Kahn N., Cove J., Silverman WF., Amsterdam A., Hershfinkel M. 2006.
774 Silencing of ZnT-1 expression enhances heavy metal influx and toxicity. *Journal of molecular*
775 *medicine* 84:753–763.
776 O'Shaughnessy PJ. 2014. Hormonal control of germ cell development and spermatogenesis. *Seminars in*
777 *cell & developmental biology* 29:55–65.
778 Palevitch O., Abraham E., Borodovsky N., Levkowitz G., Zohar Y., Gothilf Y. 2009. Nasal embryonic
779 LHRH factor plays a role in the developmental migration and projection of gonadotropin-releasing
780 hormone 3 neurons in zebrafish. *Developmental dynamics: an official publication of the American*
781 *Association of Anatomists* 238:66–75.
782 Parker J., Tsagkogeorga G., Cotton JA., Liu Y., Provero P., Stupka E., Rossiter SJ. 2013. Genome-wide
783 signatures of convergent evolution in echolocating mammals. *Nature* 502:228–231.
784 Partha R, Chauhan BK, Ferreira Z, Robinson JD, Lathrop K, Nischal KK, Chikina M, Clark NL 2017.
785 Subterranean mammals show convergent regression in ocular genes and enhancers, along with
786 adaptation to tunneling. *eLife* 6:e25884.
787 Periasamy M., Strehler EE., Garfinkel LI., Gubits RM., Ruiz-Opazo N., Nadal-Ginard B. 1984. Fast
788 skeletal muscle myosin light chains 1 and 3 are produced from a single gene by a combined process
789 of differential RNA transcription and splicing. *The Journal of biological chemistry* 259:13595–
790 13604.
791 Pivkin IV., Peng Z., Karniadakis GE., Buffet PA., Dao M., Suresh S. 2016. Biomechanics of red blood
792 cells in human spleen and consequences for physiology and disease. *Proceedings of the National*
793 *Academy of Sciences of the United States of America* 113:7804–7809.
794 Polo SE., Blackford AN., Chapman JR., Baskcomb L., Gravel S., Rusch A., Thomas A., Blundred R.,

795 Smith P., Kzhyshkowska J., Dobner T., Taylor AMR., Turnell AS., Stewart GS., Grand RJ., Jackson
796 SP. 2012. Regulation of DNA-end resection by hnRNPU-like proteins promotes DNA double-strand
797 break signaling and repair. *Molecular cell* 45:505–516.
798 Ponceau A., Albigès-Rizo C., Colin-Aronovicz Y., Destaing O., Lecomte MC. 2015. αII-spectrin
799 regulates invadosome stability and extracellular matrix degradation. *PloS one* 10:e0120781.
800 van Praag H., Schinder AF., Christie BR., Toni N., Palmer TD., Gage FH. 2002. Functional neurogenesis
801 in the adult hippocampus. *Nature* 415:1030–1034.
802 Qin J., Zhou Z., Chen W., Wang C., Zhang H., Ge G., Shao M., You D., Fan Z., Xia H., Liu R., Chen C.
803 2015. BAP1 promotes breast cancer cell proliferation and metastasis by deubiquitinating KLF5.
804 *Nature communications* 6:8471.
805 Robinson MD., McCarthy DJ., Smyth GK. 2010. edgeR: a Bioconductor package for differential
806 expression analysis of digital gene expression data. *Bioinformatics* 26:139–140.
807 Rodriguez KA., Wywial E., Perez VI., Lambert AJ., Edrey YH., Lewis KN., Grimes K., Lindsey ML.,
808 Brand MD., Buffenstein R. 2011. Walking the oxidative stress tightrope: a perspective from the
809 naked mole-rat, the longest-living rodent. *Current pharmaceutical design* 17:2290–2307.
810 Ruusuvuori E., Li H., Huttu K., Palva JM., Smirnov S., Rivera C., Kaila K., Voipio J. 2004. Carbonic
811 anhydrase isoform VII acts as a molecular switch in the development of synchronous gamma-
812 frequency firing of hippocampal CA1 pyramidal cells. *The Journal of neuroscience: the official*
813 *journal of the Society for Neuroscience* 24:2699–2707.
814 Salnikow K., Kluz T., Costa M., Piquemal D., Demidenko ZN., Xie K., Blagosklonny MV. 2002. The
815 regulation of hypoxic genes by calcium involves c-Jun/AP-1, which cooperates with hypoxia-
816 inducible factor 1 in response to hypoxia. *Molecular and cellular biology* 22:1734–1741.
817 Sang Y., Ortega MT., Blecha F., Prakash O., Melgarejo T. 2005. Molecular cloning and characterization
818 of three beta-defensins from canine testes. *Infection and immunity* 73:2611–2620.
819 Schiöth HB., Boström A., Murphy SK., Erhart W., Hampe J., Moylan C., Mwinyi J. 2016. A targeted
820 analysis reveals relevant shifts in the methylation and transcription of genes responsible for bile acid

homeostasis and drug metabolism in non-alcoholic fatty liver disease. *BMC genomics* 17:462.

Schlange T., Matsuda Y., Lienhard S., Huber A., Hynes NE. 2007. Autocrine WNT signaling contributes to breast cancer cell proliferation via the canonical WNT pathway and EGFR transactivation. *Breast cancer research: BCR* 9:R63.

Scott RW., Olson MF. 2007. LIM kinases: function, regulation and association with human disease. *Journal of molecular medicine* 85:555–568.

Seluanov A., Hine C., Azpurua J., Feigenson M., Bozzella M., Mao Z., Catania KC., Gorbunova V. 2009. Hypersensitivity to contact inhibition provides a clue to cancer resistance of naked mole-rat. *Proceedings of the National Academy of Sciences of the United States of America* 106:19352–19357.

Seri B., García-Verdugo JM., McEwen BS., Alvarez-Buylla A. 2001. Astrocytes give rise to new neurons in the adult mammalian hippocampus. *The Journal of neuroscience: the official journal of the Society for Neuroscience* 21:7153–7160.

Shatz CJ. 2009. MHC class I: an unexpected role in neuronal plasticity. *Neuron* 64:40–45.

Shayakul C., Clémençon B., Hediger MA. 2013. The urea transporter family (SLC14): physiological, pathological and structural aspects. *Molecular aspects of medicine* 34:313–322.

Shimizu T., Hibi M. 2009. Formation and patterning of the forebrain and olfactory system by zinc-finger genes Fezf1 and Fezf2. *Development, growth & differentiation* 51:221–231.

Shusterman E., Beharier O., Shiri L., Zarivach R., Etzion Y., Campbell CR., Lee I-H., Okabayashi K., Dinudom A., Cook DI., Katz A., Moran A. 2014. ZnT-1 extrudes zinc from mammalian cells functioning as a Zn(2+)/H(+) exchanger. *Metallomics: integrated biometal science* 6:1656–1663.

Sikes RS. 2016. 2016 Guidelines of the American Society of Mammalogists for the use of wild mammals in research and education. *Journal of mammalogy* 97:663–688.

Simão FA., Waterhouse RM., Ioannidis P., Kriventseva EV., Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212.

Smith-Unna R., Boursnell C., Patro R., Hibberd JM., Kelly S. 2016. TransRate: reference-free quality

847 assessment of de novo transcriptome assemblies. *Genome research*.

848 Song L., Florea L. 2015. Rcorrector: efficient and accurate error correction for Illumina RNA-seq reads.

849 *GigaScience* 4:48.

850 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large

851 phylogenies. *Bioinformatics* 30:1312–1313.

852 Suyama M., Torrents D., Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments

853 into the corresponding codon alignments. *Nucleic acids research* 34:W609–12.

854 Swanson WJ, Yang Z, Wolfner MF, Aquadro CF 2001. Positive Darwinian selection drives the evolution

855 of several female reproductive proteins in mammals. *Proceedings of the National Academy of*

856 *Sciences* 98:2509–2514.

857 Tammone MN., Lacey EA., Relva MA. 2012. Habitat use by colonial tuco-tucos (Ctenomys sociabilis):

858 specialization, variation, and sociality. *Journal of mammalogy* 93:1409–1419.

859 Team, R C. 2013. R: A language and environment for statistical computing.

860 Tian X., Azpurua J., Hine C., Vaidya A., Myakishev-Rempel M., Ablaeva J., Mao Z., Nevo E.,

861 Gorbunova V., Seluanov A. 2013. High-molecular-mass hyaluronan mediates the cancer resistance

862 of the naked mole rat. *Nature* 499:346–349.

863 Tomotani BM., Flores DEFL., Tachinardi P., Paliza JD., Oda GA., Valentinuzzi VS. 2012. Field and

864 laboratory studies provide insights into the meaning of day-time activity in a subterranean rodent

865 (Ctenomys aff. knighti), the tuco-tuco. *PloS one* 7:e37918.

866 Urrejola D., Lacey EA., Wieczorek JR., Ebensperger LA. 2005. Daily Activity Patterns of Free-Living

867 Cururos (Spalacopus cyanus). *Journal of mammalogy* 86:302–308.

868 Vasicek CA., Oosthuizen MK., Cooper HM., Bennett NC. 2005. Circadian rhythms of locomotor activity

869 in the subterranean Mashona mole rat, Cryptomys darlingi. *Physiology & behavior* 84:181–191.

870 Verga Falzacappa C., Mangialardo C., Patriarca V., Bucci B., Amendola D., Raffa S., Torrisi MR.,

871 Silvestrini G., Ballanti P., Moriggi G., Stigliano A., Brunetti E., Toscano V., Misiti S. 2009. Thyroid

872 hormones induce cell proliferation and survival in ovarian granulosa cells COV434. *Journal of*

873 *cellular physiology* 221:242–253.

874 Vianna MR., Alonso M., Viola H., Quevedo J., de Paris F., Furman M., de Stein ML., Medina JH.,
875 Izquierdo I. 2000. Role of hippocampal signaling pathways in long-term memory formation of a
876 nonassociative learning task in the rat. *Learning & memory* 7:333–340.

877 Wada H., Usui M., Sakuragawa N. 2008. Hemostatic abnormalities and liver diseases. *Seminars in*
878 *thrombosis and hemostasis* 34:772–778.

879 Wang W-H., Giebisch G. 2009. Regulation of potassium (K) handling in the renal collecting duct.
880 *Pflugers Archiv: European journal of physiology* 458:157–168.

881 Waterhouse RM., Tegenfeldt F., Li J., Zdobnov EM., Kriventseva EV. 2013. OrthoDB: a hierarchical
882 catalog of animal, fungal and bacterial orthologs. *Nucleic acids research* 41:D358–65.

883 Wei B., Jin J-P. 2016. TNNT1, TNNT2, and TNNT3: Isoform genes, regulation, and structure-function
884 relationships. *Gene* 582:1–13.

885 Welsh JS., Traum TL. 2016. Regarding Mole Rats and Cancer. *Veterinary pathology* 53:1264–1265.

886 Wu-Baer F., Ludwig T., Baer R. 2010. The UBXN1 protein associates with autoubiquitinated forms of
887 the BRCA1 tumor suppressor and inhibits its enzymatic function. *Molecular and cellular biology*
888 30:2787–2798.

889 Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular biology and evolution*
890 24:1586–1591.

891 Yang J., He MM., Niu W., Wrighton SA., Li L., Liu Y., Li C. 2012. Metabolic capabilities of cytochrome
892 P450 enzymes in Chinese liver microsomes compared with those in Caucasian liver microsomes.
893 *British journal of clinical pharmacology* 73:268–284.

894 Zelová J., Šumbera R., Okrouhlík J., Sklíba J., Lövy M., Burda H. 2011. A seasonal difference of daily
895 energy expenditure in a free-living subterranean rodent, the silvery mole-rat (Heliophobius
896 argenteocinereus; Bathyergidae). *Comparative biochemistry and physiology. Part A, Molecular &*
897 *integrative physiology* 158:17–21.

898 Zhang J, Dyer KD, Rosenberg HF 2000. Evolution of the rodent eosinophil-associated RNase gene family

by rapid gene sorting and positive selection. *Proceedings of the National Academy of Sciences* 97:4701–4706.