bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Crowdfunded whole-genome sequencing of the celebrity cat Lil BUB identifies causal mutations for her osteopetrosis and polydactyly

Mike Bridavsky[1], Heiner Kuhl[2,3], Arthur Woodruff[4], Uwe Kornak[5,6,7], Bernd Timmermann[3], Norbert Mages[3], 99 Lives Consortium, Darío G. Lupiáñez[8,*],Orsolya Symmons[9,*], Daniel M. Ibrahim[5,6,7,*]

*equal contributions, corresponding authors

1. Lil BUB LLC, Bloomington, IN 47401, USA.

2. Department of Ecophysiology and Aquaculture, Leibniz Institute for Freshwater Ecology and Inland Fisheries, Berlin, Germany.

3. Max Planck Institute for Molecular Genetics, Sequencing Core Facility, Berlin, Germany.

4. Pet Housecalls Inc., Clear Creek, IN 47401, USA.

5. Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany.

6. Institute for Medical and Human Genetics, Charité Universitätsmedizin Berlin, 13353 Berlin, Germany.

7. Berlin-Brandenburg Center for Regenerative Therapies, Charité-Universitätsmedizin Berlin, D 13353 Berlin, Germany.

8. Epigenetics and Sex Development Group, Berlin Institute for Medical Systems Biology, Max-Delbrück Center for Molecular Medicine, 13125 Berlin-Buch, Germany.

9. Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104, USA.

Corresponding authors: Dario.Lupianez@mdc-berlin.de (D.G.L.), osymmons@seas.upenn.edu (O.S.), ibrahim@molgen.mpg.de (D.M.I)

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Rare diseases and their underlying molecular causes are often poorly studied, posing challenges for patient diagnosis and prognosis. The development of next-generation sequencing and its decreasing costs promises to alleviate such issues by supplying personal genomic information at a moderate price. Here, we used crowdfunding as an alternative funding source to sequence the genome of Lil BUB, a celebrity cat affected by rare disease phenotypes characterized by supernumerary digits, osteopetrosis and dwarfism, all phenotypic traits that also occur in human patients. We discovered that Lil BUB is affected by two distinct mutations: a heterozygous mutation in the limb enhancer of the *Sonic hedgehog* gene, previously associated with polydactyly in Hemingway cats; and a novel homozygous frameshift deletion affecting the *TNFRSF11A* (*RANK*) gene, which has been linked to osteopetrosis in humans. We communicated the progress of this project to a large online audience, detailing the 'inner workings' of personalized whole genome sequencing with the aim of improving genetic literacy. Our results highlight the importance of genomic analysis in the identification of disease-causing mutations and support crowdfunding as a means to fund low-budget projects and as a platform for scientific communication.

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Rare diseases represent a challenge for the research community and a public health problem. Although each disease individually affects a small number of people (no more than 1 person per 1,000-2,000, depending on the country), more than 7,000 different rare diseases have been reported. Therefore, it is estimated that the collective number of cases can be as high as 30 million in Europe and 25 million in the USA, representing 5-10% of the population[1]. Research on individual rare diseases has been traditionally neglected in favor of other, more common, medical conditions. Consequently, patients affected by these conditions often receive inadequate diagnostic and medical support, leading to so-called "diagnostic Odysseys".

Comprehensive genetic testing holds promise to help such patients by identifying the cause of their disease, as well as enabling more informed, personalized treatments and prognoses. However, despite more than 10 years since the sequence of the human genome was revealed, our capability to interpret the effects of genomic variation remains limited. As a result, testing of candidate loci identifies disease-causing mutations for only a fraction of patients, and leaves a substantial number of unsolved cases. The increasing efficiency and decreasing cost of genome sequencing are expected to increase the number of available genomes, promising to improve the identification of disease-causing variants[2] and further enhance our knowledge of the human genome. However, the next frontier lies on the inevitable challenges of evaluating hitherto unknown variants[3], as well as communicating the potential and challenges of sequencing efforts to the patients and the general public, who often have no formal training in genetics, potentially leading to uncertainty and misunderstandings in their interpretation of the results[4–6].

Animals provide a valuable resource in this framework to better understand genetic diseases, with domesticated animals and pets in particular displaying a rich tapestry of traits that often share the same genetic cause as in human[7–9]. One such example is Lil BUB, a pet cat with several congenital malformations (**Figure 1**), whose special appearance and social media presence has granted her international fame and a fandom across the globe. Lil BUB was diagnosed with infantile malignant osteopetrosis, a very rare condition that, in humans, affects 1 in 250,000-300,000 births. The disorder is caused by an impairment of osteoclasts, leading to an imbalance between the processes of bone resorption, performed by such cells, and

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

bone formation, controlled by osteoblasts[10,11]. Together, these two cell types constantly remodel bone tissue throughout life to ensure its proper growth, mechanical stability and physiological functions. Thus, impairment of osteoclast differentiation or loss of resorptive activity can lead to the accumulation of bone tissue, causing increased bone density, susceptibility to fractures, as well as several traits that are secondary to those bone changes. Genetically, the disease is heterogeneous and causal mutations for the recessive, more severe form has been found in seven genes, which differ somewhat in their clinical manifestations and patient prognosis[10]. Currently, the only curative therapy is transplantation of hematopoietic stem cells, although alternative treatments such as gene therapy or cytokine-replacement therapy have been reported[10]. In cats, reports of osteopetrosis are rare[12,13], and no examples of the congenital form are known, although it has been shown that osteopetrosis can be acquired through infection with retroviruses[14,15] .

In addition, Lil BUB presents preaxial polydactyly, a condition characterized by the presence of an extra digit on the thumb side of the paw. This trait is not uncommon in cats, with a high occurrence in several breeds (*e.g.* Maine Coon, Pixie bob) and multiple outbred cats[16,17]. Three mutations on chromosome A2, in the regulatory (enhancer) region of the *Sonic hedgehog* gene, have been previously identified as dominant causal variants[18]. Of these, one is known to be prevalent in North American cats, including in a colony of polydactyl cats on Ernest Hemingway's estate on Key West[19]. Additional mutations on this enhancer and the associated polydactyly have been reported to occur in chicken, dog, mouse or human[18,20,21], highlighting the evolutionary conservation of this developmental pathway.

The combined manifestation of congenital osteopetrosis and polydactyly, as observed in Lil BUB, is quite remarkable and has never been reported previously. To identify the underlying cause of the observed phenotypes, we performed a comprehensive genetic analysis for Lil BUB, and found that she is affected by two distinct mutations: a heterozygous mutation on the limb enhancer of the *Sonic hedgehog* gene; and a frameshift deletion affecting the *TNFRSF11A* (*RANK*) gene. We considered this a test case for precision medicine, since it would allow a more precise diagnosis and prognosis for Lil BUB and a general improvement of our understanding of the disease's etiology. Due to the small-scale characteristics of the project, we determined to use a crowdfunding approach to finance the research. Moreover, we decided to communicate the planning, execution, and analysis of this study to the broader public using a range on online media platforms, to raise

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

awareness for rare diseases and explain the challenges, promises and steps involved in whole genome sequencing. Furthermore, our results highlight the importance of genomic analyses as a means to improve our understanding of developmental disease.

Results

Phenotypic description

Lil BUB was found as a feral kitten that was abandoned by her mother in rural Indiana, USA and was subsequently hand-reared. She was born with several malformations, mainly related to skeletal growth and development. At the age of 7 months, she stopped growing and remained the size of a kitten (adult weight: 1.8 kg; **Figure 1A-C**). This manifested as an extreme case of dwarfism, causing her limbs to be remarkably small in comparison to her body. From birth she was smaller than other cats and by 10 months her movements became severely restricted. Radiologic studies revealed opaque medullary cavities and widened metaphyses resulting in an "Erlenmeyer Flask Deformity" and bowing of long bones (**Figure 1D**). Multiple radiopaque and radiolucent transverse bands were present in the metaphysis, a typical sign of problems with bone resorption. Ribs and sternae were thickened and the epiphyses of the long bones were small, while adjacent physes were thin. Moreover, her skull and jaw remained underdeveloped, with her tongue and eyes displaying a normal size and thus protruding notably from the cranial cavities. The teeth did not erupt or are not present, except for one or two small crowns on the right maxilla. There is a complete absence of bony palate. Despite her jaws and teeth being underdeveloped, Lil BUB has no difficulty feeding and eats food (both wet and dry) with no problem.

Based on her features and the radiological observations, she was diagnosed with a severe form of osteopetrosis, displaying similar features as described in human patients. Although this condition often co-occurs with additional symptoms that affect hormone and blood homeostasis, thyroid screening showed normal T4 (3.8 µg/ml) and Ca^{2+} (9.5mg/dl) levels and bloodwork showed no abnormalities, although enlarged platelets were noted. Lil BUB displayed low alkaline phosphatase levels (<20 IU/L; normal range 23-107 IU/L), which indicates low bone formation. However, in patients affected with osteopetrosis, these levels are often normal or even elevated because of the increased bone surface and varying total osteoblast numbers, and are thus unreliable biomarkers for osteopetrosis.. Of note, Lil BUB

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

received continuous treatment with an electromagnetic pulse device (Assisi loop) since the age of 22 months. Coinciding with the treatment, her condition improved, and Lil BUB regained her mobility.

In addition to the traits associated with osteopetrosis, Lil BUB displays preaxial polydactyly affecting both forelimbs and hindlimbs (**Figure 1C and D**). A complete duplication of a normal biphalangeal thumb is observed in all paws, resulting in 22 digits (6 on each forepaw and 5 on both hindpaws).

Project funding and outreach aspect

Given Lil BUB's unique combination of osteopetrosis traits and polydactyly, we aimed to investigate whether her condition was an unusual form of (feline) osteopetrosis that included polydactyly, or if she was simultaneously affected by two rare conditions. We opted to perform whole genome sequencing to address this question, because a candidate gene approach did not seem feasible. We estimated that the financial requirements needed for such an endeavor would be moderate, and therefore decided to finance the project via crowdfunding[22], with an initial goal of $6,500. We reasoned that Lil BUB's recurrent engagement with the internet public presented a compelling case for "citizen science" that would benefit from alternative funding approaches, while simultaneously providing a platform to communicate genomic analyses to the broader public.

For this purpose, we used the crowdfunding platform "Experiment.com", a site that focuses exclusively on raising funds for research and science projects. During the fundraising period, our campaign attracted the attention of more than 25,250 people (**Figure 2A**). After the 45-day campaign, the project was successfully funded, raising a total of $8,225, 26% more than the initial goal. The project was supported by a total of 248 "backers", representing a 1% of the page views at the crowdfunding site during this period. Backers donated $33.17 on average. During the campaign, page views and donations peaked when the project was mentioned by the official Lil BUB Facebook page or major media outlets, such as the Washington Post or Der Spiegel (**Figure 2A** and **Supplementary Table S1**).

Beyond crowdfunding, we further used our social media networks to engage with the public and communicate the step-by-step progress of our research project. Our aim

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

was to provide comprehensive explanations to the biology behind the project to a non-scientific audience, which represents most of the funders. One example was a Reddit Ask Me Anything (AMA) session that, over the course of 24 hours, allowed us to provide answers to specific project-related questions from the audience. We also produced YouTube video blogs to disseminate our results in a visual, more accessible format. As a consequence of these efforts, traffic on our sites continued after the crowdfunding period, with almost 11,000 additional visitors to our crowdfunding site, more than 5,000 views of our research updates and 15,000 visitors of our blog as of December 2018. We reached audiences from 109 countries on six continents, with approximately 50% of our websites' traffic from the US and 11 other countries (US, Germany, UK, France, Canada, Hungary, Australia, Spain, Brazil, Italy and India) making up for 90% of our readers (**Figure 2B**).

Genome sequencing and initial analysis of informative variants

To identify Lil BUB's disease-causing mutations, we sequenced her genomic DNA using short-read sequencing technology. We mapped reads against the domestic cat reference genome (Felis_catus 6.2), obtaining ~50X coverage and, through comparison with the reference genome, identified a total of ~6 million SNVs and small insertions/deletions, as well as 424 larger structural variants (**Figure 3A-C**). However, the relatively large number of structural variants and their genomic distribution (**Figure 3B**) suggests that the great majority are likely false-positives due to problems in the genome annotation. We therefore did not follow up on those identified variants. Of note, we observed an excess of homozygosity throughout the entire genome, likely indicating inbreeding events over the course of multiple generations (**Figure 3D**).

As a first step of analyzing Lil BUB's genome sequence, we looked at known variants previously linked to distinct phenotypical traits in cats (**Table 1**). In most cases, the studied variants matched with Lil BUB's phenotypical characteristics. For example, Lil BUB does not harbor any of the mutations known to cause retinal atrophy, feline hereditary myopathy, hypertrophic cardiomyopathy or albinism. We found, however, that Lil BUB is a carrier for the variant c.123delCA in the *ASIP* gene that has been linked with black coat colour[23]. In addition, she carries two variants that affect the pattern of her coat's tabby markings[24]. Such markings consist of light background hair, interspersed by darker hair in specific patterns, of which mackerel and blotched

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

are the most frequent. In mackerel cats, the light and dark hairs form regular stripes. In contrast, the blotched pattern is characterized by disorganized stripes with the dark hair forming whorls. Lil BUB's coat clearly displays the mackerel tabby pattern, although she is heterozygous for two distinct mutations (p.Trp841* and p.Thr139Asn) in the *Taqpep* gene that are associated with the blotched pattern. While our short-read sequencing cannot discern whether both variants are on the same chromosome, the co-occurrence of the two mutations with a mackerel pattern is consistent with published data, which found that 7 of 10 cats with the p.Trp841*/p.Thr139Asn mutations display this pattern[24].

In addition to visible traits, her genetic findings were also consistent with molecular/physiological traits. For example, Lil BUB did not have the mutation associated with hypokalemia (intermittent reduced potassium levels) in Burmese cats. Consistent with this, she did not show any hypokalemic symptoms and her bloodwork displayed normal potassium levels (4.3 mmol/l; clinical signs appear below 3 mm/l)[25]. Similarly, we found Lil BUB to have the haplotype associated with blood type A in cats[26,27]. While her blood type was unknown at the time of our initial sequencing analysis, further tests confirmed that she indeed has blood type A. A major unexpected finding from our initial variant analysis was that, despite her white gloving phenotype (white fur only on her paws), she did not carry the variant known to cause gloving in the Birman breed. Such result suggests the existence of additional uncharacterized variants for gloving, an hypothesis also supported by observations in other breeds, such as the Ragdoll, displaying "mitted" patterns in the absence of the Birman mutation[28].

Identification of disease-causing mutations

Having performed this initial genotype-to-phenotype characterization, we next sought to identify the mutation(s) responsible for Lil BUB's osteopetrosis and polydactyly traits. Initially, we chose a targeted approach and specifically analyzed the ZRS enhancer of the *Sonic hedgehog* gene, a genomic region associated to the appearance of polydactyly across mammals. We found that Lil BUB was heterozygous for an A to G mutation at chrA2:167,313,488, a variant previously described to cause the appearance of additional digits in Hemingway cats. We verified both the nature and the heterozygous state of the mutation using Sanger sequencing (**Figure 4A and B**). The affected enhancer is well-known for its

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

association with polydactyly, with mutations in humans, mouse and cats identified in different parts of the element, which cause ectopic expression of *Sonic hedgehog* during embryonic limb development, subsequently leading to the formation of additional digits (**Figure 4C and D**).

Having identified a likely candidate for the polydactyly phenotype, we next searched for causes underlying her osteopetrosis. Although we initially considered the possibility that retroviral infection may have given rise to Lil BUB's osteopetrosis, she tested negative for both feline leukemia virus (FELV) and feline immunodeficiency virus (FIV). In addition, several healthy siblings were born in the same litter, thus eliminating this as a plausible causal mechanism. We therefore focused on the analysis of her genome sequence. We filtered Lil BUB's variants in order to retain SNVs and inDels that affect genes (~32k variants) and specifically focused on those with a known role in bone mass regulation and/or osteoclast differentiation in humans (**Supplementary Table S2**). Having thus narrowed the list to 12 candidate variants (**Table 2**), we further eliminated seven of these based on read coverage, mapping quality, and sequence complexity. We discarded three additional exonic variants, as manual inspection suggested that they were in fact intronic. This was subsequently confirmed by gene annotations in the genome builds 8.0 and 9.0.

Of the two remaining variants, one involved *ZMPSTE24*, a gene that can cause mandibular defects in both human and mice when mutated[29–31]. While Lil BUB's jaw is underdeveloped, providing superficial resemblance to the human and mouse *ZMPSTE24* loss-of-function phenotypes, diverse evidences suggest that it is not the major causal mutation for her traits. First, there are multiple features, typically associated with human and mouse *ZMPSTE24*-deficiency, that are inconsistent with Lil BUB's phenotype. These include spontaneous bone fractures due to decreased bone density, lipodystrophy (general fat loss, although this feature might be secondary to feeding difficulties), atrophy of the skin and loss of hair or sparse, brittle hair[29,30,32,33]. Second, the variant identified in Lil BUB (a T insertion at chrC1:30,324,507, causing a frameshift in the protein) appears to be in a heterozygous state. This contrasts with other studies showing that affected human patients are homozygous or compound heterozygous, while heterozygous individuals are reported to be healthy[29,33,34]. In addition, heterozygous *Zmpste24* knockout mice are either indistinguishable from their wild-type littermates[32] or develop much milder phenotypes[30]. For these reasons, we consider unlikely that Lil BUB's *ZMPSTE24* mutation is causative for her skeletal malformations.

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In contrast, several lines of evidence support that the mutation affecting the gene *TNFRSF11A* is likely the disease-causing variant for Lil BUB's osteopetrosis. *TNFRSF11A* codes for a transmembrane receptor, RANK, that is expressed on osteoclasts and osteoclast precursors. Activation of this receptor by its ligand RANKL is necessary for osteoclast differentiation, survival, and function, and is therefore essential for the process of bone resorption. The identified mutation is a homozygous deletion of an A in a CAT sequence in exon 8 (c.806delA; **Figure 5A**), causing a frameshift and subsequent premature truncation of the protein (p.His272Leufs*16; **Figure 5B**). This truncation removes the majority of RANK's intracellular domain, including two regions that are essential for oligomerization of the receptor and signal activation, thus rendering the protein dysfunctional[35]. Due to the central role of RANK in osteoclast function, variants that truncate murine or human RANK cause autosomal-recessive osteopetrosis, characterized by a complete absence of osteoclasts (**Figure 5C**). Affected individuals display early onset osteopetrosis, sharing several phenotypic features with Lil BUB, such as growth arrest and small stature, Erlenmeyer flask bones, exophtalamus (protruding eyes), dysmorphologies of the skull and jaw, and – in mice - failure of tooth eruption[36–40]. Two variants in particular, the human p.Gly280* and p.Trp434* mutations, are of interest as they also truncate the intracellular domain of RANK and biochemical assays of the deleted regions confirmed their involvement in RANK function[41,42]. Moreover, the identified variant was not present in 131 cats sequenced by the 99 Lives Consortium - none of which has osteopetrosis (data accessed August 2018). Taken together, these data strongly suggest that the homozygous, truncating *TNFRSF11A/*RANK variant is the causative mutation for the osteopetrotic phenotype observed in Lil BUB.

Discussion

Genetic studies in livestock and pet animals can provide valuable insight into human disease, due to the evolutionary conservation of mammalian body plans and their gene regulatory networks[43,44]. Domesticated animals, in particular, often display a vast array of readily observable phenotypes as a result of inbreeding[45,46], which causes a gradual increase in homozygosity and the accumulation of recessive traits. Additionally, rare spontaneous mutations that mimic human conditions have also been documented[20,47,48]. Recent improvements in sequencing technologies have boosted our capacity to analyze entire genomes, expanding the potential to identify

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

pathogenic mutations and constituting an important step towards disease management. Here, we demonstrate how personalized genomics and the conservation of mammalian disease phenotypes can be effectively linked, by sequencing the genome of a pet cat, Lil BUB, and identifying likely causal mutations of her polydactyly and osteopetrosis.

Lil BUB displays an unusual combination of polydactylous and osteopetrotic traits. However, mapping the genetic mutation(s) responsible for these traits through breeding or pedigree analysis was not possible, as she is spayed and her parents unknown. Therefore, we opted to perform whole genome sequencing to identify potential causal variants. By doing so, we identified approximately 6 million single nucleotide variants in Lil BUB compared to the reference genome, obtained from a female Abyssinian cat. This number of variants is similar to that observed in other breeds sequenced at high coverage[28]. Interestingly, we observed a substantially higher degree of homozygosity throughout Lil BUB's genome than expected, based on prior studies in both inbred and random populations. In such studies, average heterozygosity measurements at the population level ranged between 0.51 and 0.65 (based on microsatellites) or 0.53-0.85 (based on short tandem repeat analysis[49,50]). Hence, our data suggests that multiple rounds of inbreeding occurred in Lil BUB's ancestry, either as a result of a small cat population at her place of birth or because her ancestry is partially of a certain cat breed.

In the case of the osteopetrosis, we excluded the possibility that retroviral infection may have given rise to the disease, and identified a mutation in *TNFRSF11A*, a gene well known to cause such abnormalities. The identified mutation was homozygous, consistent with the autosomal recessive osteopetrosis subtype typically associated with *TNFRSF11A*/RANK (**Figure 5**). Beyond the phenotypical similarities between Lil BUB and mice or humans with RANK mutations, the identified variant was absent in all other tested cats of the 99 Lives Consortium, further supporting a causal relationship. Unfortunately, functional validations, such as an *in vitro* osteoclast differentiation assay, are not possible in this case, given the large blood sample requirements.

However, despite the overall similarity in phenotype with human and mouse *RANK*-associated osteopetrosis, we noted some differences: while the spleen is typically enlarged in humans and mice, no hepatosplenomegaly was observed in Lil BUB's case. Furthermore, human patients affected by *RANK* mutations often display severe

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

visual impairment, although Lil BUB's vision seems to be unaffected. It is possible, however, that the described visual impairment is human-specific, as it has not been reported in mice[38–40]. Other Lil BUB´s features, such as severely hypoplastic jaw bones, have not been reported in human patients with RANKL-related autosomal recessive osteopetrosis, but pronounced mandibular hypoplasia was observed in cattle with *CLCN7*-related osteopetrosis, where it was paralleled by gingival hamartomas[51]. Thus, skull bones seem to be especially prone to species differences in autosomal recessive osteopetrosis. In addition, *RANK*-associated osteopetrosis can be linked with hypogammaglobulinemia and an increased risk of infections in some patients[36,37]. Although such parameters were not measured in Lil BUB´s bloodwork, the absence of remarkable recurrent infectious processes during her life suggests normal immunoglobulin levels. Lil BUB's low alkaline phosphatase levels are notable since the enlarged osteopetrotic bone surface should lead to higher numbers of osteoblasts producing this enzyme. RANK has recently been identified as a coupling factor, which can stimulate osteoblast activity through binding to RANKL, thereby initiating reverse signaling[52]. Thus, a loss of RANK protein might explain this unusually weak bone formation, although cannot discern whether the identified frameshift mutation leads to a truncated, partly functional RANK protein or to a complete loss.ç

Currently, the only known therapy for autosomal recessive osteopetrosis in humans is bone marrow transplantation, which has not been required in the case of Lil BUB due to the favorable evolution of her condition. Such improvement was noticed soon after receiving regular sessions of pulsed electromagnetic field therapy. However, currently a scientific basis for the use of this method to treat the disease is missing and, to date, no other validated cases of feline osteopetrosis exist to verify if the observed effects are reproducible.

In contrast to osteopetrosis, polydactyly is a frequently observed trait in cats and the variant identified in Lil BUB has previously been documented in Hemingway cats, as well as other North American polydactyl breeds[53]. This non-coding mutation is assumed to result in polydactyly due to ectopic expression of the signaling molecule *Sonic hedgehog* in the anterior part of the developing limb buds (where it is normally not expressed). Interestingly, Hemingway cats are polydactylous mainly on the forelimbs[53], while all four paws are affected in Lil BUB´s case. However, polydactyly has highly variable expressivity, even within a colony. At the same time, morphological variation of polydactyl cats is often reduced within breeding lines[16,17,53],

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

suggesting the existence of possible modifier genes. Based on our data we cannot distinguish whether there is a genetic component for the four-pawed polydactyly in Lil BUB or if her phenotype is simply part of a natural spectrum associated with the Hemingway mutation.

We also identified an additional mutation (in the *ZMPSTE24* gene*)* that could potentially have modified the outcome of her skeletal traits, raising the possibility that other exist. We also eliminated non-exonic variants, which can also affect gene regulation and/or splicing and need to be considered if no plausible coding mutation is detected. In addition, given that gene annotation of the cat genome is still a work-in-progress, it is possible that some variants currently annotated as non-coding will later be re-assigned, or *vice versa*, as we noticed during the analysis of Lil BUB´s genome (see **Methods**). This underlines the importance of thorough and comprehensive genome annotation for personalized genomics.

Finally, the efforts presented in this paper were also part of a public outreach project to demonstrate the use of personalized genomics: we raised funds for sequencing through a crowdfunding campaign and reported on various steps of sequencing and analysis on Twitter, through Facebook, blog posts and Reddit, as well as YouTube videos. Given the rise of direct-to-consumer genetic testing both for humans and, more recently, for pets[54,55], we hope that our public engagement activities contributed to a broader understanding of how genetic information is obtained and what insights can (and cannot) be gained for a single individual case with unknown ancestry or family medical history.

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Methods

Ethics statement

Blood samples from Lil BUB were collected by a licensed veterinarian, with the consent of her owner. According to the framework defined by the Directive 2010/63/UE of the European Parliament and of the Council, no animal experimentation, only non-experimental clinical veterinary practice (taking biopsies and imaging for disease diagnosis) was performed.

Sample collection, DNA extraction and sequencing

A sample of EDTA anti-coagulated whole blood was collected during a routine veterinary visit and genomic DNA was isolated using standard procedures. The quality and integrity of the DNA was assessed using the A260/280 ratio and agarose gel electrophoresis. Libraries were then prepared using Truseq sample prep for Next Seq Version2: DNA was sheared to 300bp and 500bp with Covaris according to the protocol provided by Illumina, adaptors were ligated to the two libraries, and correctly sized fragments excised from gel. We performed paired end 2x 150bp sequencing of libraries on an Illumina NextSeq 500.

Variant calling and analysis of variants

We obtained a total of 144,172,364 and 310,362,266 reads for the two libraries (~500bp and ~250bp length, respectively). Reads were mapped against Felis_catus_6.2 using Bowtie2 and we were able to map 98.38% of reads, corresponding to 40x coverage. Non-uniquely mapping reads were placed in a random fashion by Bowtie2. We did not apply repeat masking and all mapped reads were used. Sequence read de-duplication, indel re-assembly and subsequent SNV calling was performed with Platypus[56]. In total, we identified 6.19 million SNVs and InDels, 4.4 million (65.4%) were intergenic and 1.75 million (34.6%) were located in annotated genes. Variants were visualized using IGV browser[57]. We used variant predictor SnpEff[58] to determine the potential pathogenic effects of candidate osteopetrosis genes. To exclude that the identified *TNFRSF11A* variant was not simply a low frequency variant in cats, we interrogated the 99 Lives Consortium data:

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

in August 2018 we downloaded all variants identified in *TNFRSF11A/RANK* (this corresponded to the January 2017 variant analysis). The D3:81,422,862 variant was not present in 131 cats from this dataset.

Evaluation of coding variants

Due to the incomplete annotation of the cat genome, five variants affecting transcripts of bone disease related genes had to be evaluated individually. For this the most likely correct transcription and coding unit was evaluated by integrating all available gene prediction algorithms, sequence conservation, and expressed sequences in the UCSC genome browser for releases felCat5 and felCat8.0 of the *Felis catus* genome. Annotation inconsistencies of the *TNFRSF11A* transcript and coding sequence lead to ambiguities in the nomenclature of the frameshift variant. The most appropriate nomenclature is ENSEMBL c.806delA (ENSFCAT00000032280.1) and p.H270LfsX16 (ENSFCAP00000022567.1) in felCat5 or Augustus c.815delA (g12385.t1) and p.H272LfsX16 (g12385.t1_prot) in felCat8.

Sanger sequencing of the ZRS region

A fragment containing the ZRS region was PCR generated using primers that amplify a 2.8kb region surrounding the 800bp ZRS (fwd_primer: CCTTGAAGTGGAAATCTCTCCTG, rev_primer: ATCAATTGCGTGAAAACTGCAAGGG). The obtained PCR products were subcloned into a pTA-vector and four individual subclones were sequenced using six sequencing primers that together covered 2.7kb of the region: catZRSeq1: TTGATGGGGTTTTCCTCGAAC, catZRSeq2: TCAGCTTTATAGGCCTTCCCAG, catZRSeq3: CAAGACGCAAACCGCGGAG, catZRSeq4: GGGCGGATGCAGAGCTTG, catZRSeq5: TTAGTGAGATATGAGTCCATTTTCTGT, catZRSeq6: GAGCATAGCACACGGTCT.

Data availability

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Sequencing data was deposited in the Sequence Read Archive with accession number PRJNA512113. It is available using the following link: https://www.ncbi.nlm.nih.gov/sra/PRJNA512113

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. Schieppati, A., Henter, J.-I., Daina, E. & Aperia, A. Why rare diseases are an important medical and social issue. *The Lancet* **371**, 2039–2041 (2008).
2. Stavropoulos, D. J. *et al.* Whole Genome Sequencing Expands Diagnostic Utility and Improves Clinical Management in Pediatric Medicine. *NPJ Genomic Med.* **1**, (2016).
3. Dewey, F. E. *et al.* Clinical interpretation and implications of whole-genome sequencing. *JAMA* **311**, 1035–1045 (2014).
4. Biesecker, B. B. *et al.* How do research participants perceive 'uncertainty' in genome sequencing? *Genet. Med. Off. J. Am. Coll. Med. Genet.* **16**, 977–980 (2014).
5. Vos, J. *et al.* The counsellees' view of an unclassified variant in BRCA1/2: recall, interpretation, and impact on life. *Psychooncology.* **17**, 822–830 (2008).
6. Reiff, M. *et al.* 'What does it mean?': uncertainties in understanding results of chromosomal microarray testing. *Genet. Med. Off. J. Am. Coll. Med. Genet.* **14**, 250–258 (2012).
7. Gurda, B. L., Bradbury, A. M. & Vite, C. H. Focus: Comparative Medicine: Canine and Feline Models of Human Genetic Diseases and Their Contributions to Advancing Clinical Therapies. *Yale J. Biol. Med.* **90**, 417 (2017).
8. Hytönen, M. K. & Lohi, H. Canine models of human rare disorders. *Rare Dis.* **4**, (2016).
9. Pinnapureddy, A. R. *et al.* Large animal models of rare genetic disorders: sheep as phenotypically relevant models of human genetic disease. *Orphanet J. Rare Dis.* **10**, 107 (2015).
10. Sobacchi, C., Schulz, A., Coxon, F. P., Villa, A. & Helfrich, M. H. Osteopetrosis: genetics, treatment and new insights into osteoclast function. *Nat. Rev. Endocrinol.* **9**, 522–536 (2013).
11. Villa, A., Guerrini, M. M., Cassani, B., Pangrazio, A. & Sobacchi, C. Infantile malignant, autosomal recessive osteopetrosis: the rich and the poor. *Calcif. Tissue Int.* **84**, 1–12 (2009).
12. Fujita, M. *et al.* Osteopetrosis-like disease in a cat with respiratory distress. *J. Vet. Med. Sci.* **69**, 687–690 (2007).
13. Kramers, P., Flückiger, M. A., Rahn, B. A. & Cordey, J. Osteopetrosis in cats. *J. Small Anim. Pract.* **29**, 153–164 (1988).
14. Hoover, E. A. & Kociba, G. J. Bone Lesions in Cats With Anemia Induced by Feline Leukemia Virus. *JNCI J. Natl. Cancer Inst.* **53**, 1277–1284 (1974).
15. Labat, M. L. Retroviruses, immunosuppression and osteopetrosis. *Biomed. Pharmacother. Biomedecine Pharmacother.* **40**, 85–90 (1986).
16. Danforth, C. H. Heredity of polydactyly in the cat. *J. Hered.* **38**, 107–112 (1947).
17. Hamelin, A., Begon, D., Conchou, F., Fusellier, M. & Abitbol, M. Clinical characterisation of polydactyly in Maine Coon cats. *J. Feline Med. Surg.* **19**, 382–393 (2017).
18. Point mutations in a distant sonic hedgehog cis-regulator generate a variable regulatory output responsible for preaxial polydactyly | Human Molecular Genetics | Oxford Academic. Available at: https://academic.oup.com/hmg/article/17/7/978/869373. (Accessed: 11th August 2018)
19. Hemingway's Cats: An Illustrated Biography: Carlene Brennen, Hilary Hemingway: 9781561644896: Amazon.com: Books. Available at: https://www.amazon.com/Hemingways-Cats-Illustrated-Carlene-Brennen/dp/1561644897. (Accessed: 29th August 2018)
20. Maas, S. A., Suzuki, T. & Fallon, J. F. Identification of spontaneous mutations within the long-range limb-specific Sonic hedgehog enhancer (ZRS) that alter Sonic

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

hedgehog expression in the chicken limb mutants oligozeugodactyly and silkie breed. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* **240**, 1212–1222 (2011).
21. Park, K., Kang, J., Subedi, K. P., Ha, J.-H. & Park, C. Canine polydactyl mutations with heterogeneous origin in the conserved intronic sequence of LMBR1. *Genetics* **179**, 2163–2172 (2008).
22. Sauermann, H., Franzoni, C. & Shafi, K. *Crowdfunding Scientific Research*. (National Bureau of Economic Research, 2018). doi:10.3386/w24402
23. Eizirik, E. *et al.* Molecular genetics and evolution of melanism in the cat family. *Curr. Biol. CB* **13**, 448–453 (2003).
24. Kaelin, C. B. *et al.* Specifying and sustaining pigmentation patterns in domestic and wild cats. *Science* **337**, 1536–1541 (2012).
25. Gandolfi, B. *et al.* First WNK4-hypokalemia animal model identified by genome-wide association in Burmese cats. *PloS One* **7**, e53173 (2012).
26. Bighignoli, B. *et al.* Cytidine monophospho-N-acetylneuraminic acid hydroxylase (CMAH) mutations associated with the domestic cat AB blood group. *BMC Genet.* **8**, 27 (2007).
27. Omi, T. *et al.* Molecular Characterization of the Cytidine Monophosphate-N-Acetylneuraminic Acid Hydroxylase (CMAH) Gene Associated with the Feline AB Blood Group System. *PloS One* **11**, e0165000 (2016).
28. Montague, M. J. *et al.* Comparative analysis of the domestic cat genome reveals genetic signatures underlying feline biology and domestication. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 17230–17235 (2014).
29. Agarwal, A. K., Fryns, J.-P., Auchus, R. J. & Garg, A. Zinc metalloproteinase, ZMPSTE24, is mutated in mandibuloacral dysplasia. *Hum. Mol. Genet.* **12**, 1995–2001 (2003).
30. Bergo, M. O. *et al.* Zmpste24 deficiency in mice causes spontaneous bone fractures, muscle weakness, and a prelamin A processing defect. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 13049–13054 (2002).
31. de Carlos, F. *et al.* Microcephalia with mandibular and dental dysplasia in adult Zmpste24-deficient mice. *J. Anat.* **213**, 509–519 (2008).
32. Pendás, A. M. *et al.* Defective prelamin A processing and muscular and adipocyte alterations in Zmpste24 metalloproteinase-deficient mice. *Nat. Genet.* **31**, 94–99 (2002).
33. Navarro, C. L. *et al.* New ZMPSTE24 (FACE1) mutations in patients affected with restrictive dermopathy or related progeroid syndromes and mutation update. *Eur. J. Hum. Genet. EJHG* **22**, 1002–1011 (2014).
34. Diociaiuti, A. *et al.* Teledermatology diagnosis of the first Italian patient affected with restrictive dermopathy due to ZMPSTE24 homozygous mutation. *J. Eur. Acad. Dermatol. Venereol. JEADV* (2018). doi:10.1111/jdv.15351
35. Crockett, J. C., Mellis, D. J., Scott, D. I. & Helfrich, M. H. New knowledge on critical osteoclast formation and activation pathways from study of rare genetic diseases of osteoclasts: focus on the RANK/RANKL axis. *Osteoporos. Int. J. Establ. Result Coop. Eur. Found. Osteoporos. Natl. Osteoporos. Found. USA* **22**, 1–20 (2011).
36. Guerrini, M. M. *et al.* Human osteoclast-poor osteopetrosis with hypogammaglobulinemia due to TNFRSF11A (RANK) mutations. *Am. J. Hum. Genet.* **83**, 64–76 (2008).
37. Pangrazio, A. *et al.* RANK-dependent autosomal recessive osteopetrosis: characterization of five new cases with novel mutations. *J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res.* **27**, 342–351 (2012).
38. Dougall, W. C. *et al.* RANK is essential for osteoclast and lymph node development. *Genes Dev.* **13**, 2412–2424 (1999).
39. Kapur, R. P. *et al.* Malignant autosomal recessive osteopetrosis caused by spontaneous mutation of murine Rank. *J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res.* **19**, 1689–1697 (2004).

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

40. Li, J. *et al.* RANK is the intrinsic hematopoietic cell surface receptor that controls osteoclastogenesis and regulation of bone mass and calcium metabolism. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 1566–1571 (2000).
41. Xu, D., Wang, S., Liu, W., Liu, J. & Feng, X. A novel receptor activator of NF-kappaB (RANK) cytoplasmic motif plays an essential role in osteoclastogenesis by committing macrophages to the osteoclast lineage. *J. Biol. Chem.* **281**, 4678–4690 (2006).
42. Kanazawa, K. & Kudo, A. Self-assembled RANK induces osteoclastogenesis ligand-independently. *J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res.* **20**, 2053–2060 (2005).
43. Martinez-Morales, J. R. Toward understanding the evolution of vertebrate gene regulatory networks: comparative genomics and epigenomic approaches. *Brief. Funct. Genomics* **15**, 315–321 (2016).
44. Wagner, G. P. The developmental genetics of homology. *Nat. Rev. Genet.* **8**, 473–479 (2007).
45. Lyons, L. A. DNA mutations of the cat: the good, the bad and the ugly. *J. Feline Med. Surg.* **17**, 203–219 (2015).
46. Wiener, P. & Wilkinson, S. Deciphering the genetic basis of animal domestication. *Proc. Biol. Sci.* **278**, 3161–3170 (2011).
47. Bourneuf, E. *et al.* Rapid Discovery of De Novo Deleterious Mutations in Cattle Enhances the Value of Livestock as Model Species. *Sci. Rep.* **7**, 11466 (2017).
48. Peters, M. *et al.* DNA-based diagnosis of rare diseases in veterinary medicine: a 4.4 kb deletion of ITGB4 is associated with epidermolysis bullosa in Charolais cattle. *BMC Vet. Res.* **11**, 48 (2015).
49. Lipinski, M. J. *et al.* The ascent of cat breeds: Genetic evaluations of breeds and worldwide random-bred populations. *Genomics* **91**, 12–21 (2008).
50. Menotti-Raymond, M. *et al.* Patterns of molecular genetic variation among cat breeds. *Genomics* **91**, 1–11 (2008).
51. Sartelet, A. *et al.* A missense mutation accelerating the gating of the lysosomal Cl-/H+-exchanger ClC-7/Ostm1 causes osteopetrosis with gingival hamartomas in cattle. *Dis. Model. Mech.* **7**, 119–128 (2014).
52. Ikebuchi, Y. *et al.* Coupling of bone resorption and formation by RANKL reverse signalling. *Nature* **561**, 195–200 (2018).
53. Lange, A., Nemeschkal, H. L. & Müller, G. B. Biased Polyphenism in Polydactylous Cats Carrying a Single Point Mutation: The Hemingway Model for Digit Novelty. *Evol. Biol.* **41**, 262–275 (2014).
54. Deane-Coe, P. E., Chu, E. T., Boyko, A. R. & Sams, A. J. Direct-To-Consumer DNA testing of 6,000 dogs reveals 98.6-kb duplication causing blue eyes and heterochromia in Siberian Huskies. *bioRxiv* 247973 (2018). doi:10.1101/247973
55. Cat DNA Test Kit | Find about your pet's genetics | Basepaws. Available at: https://www.basepaws.com/. (Accessed: 16th December 2018)
56. Rimmer, A. *et al.* Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat. Genet.* **46**, 912–918 (2014).
57. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
58. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80–92 (2012).

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure

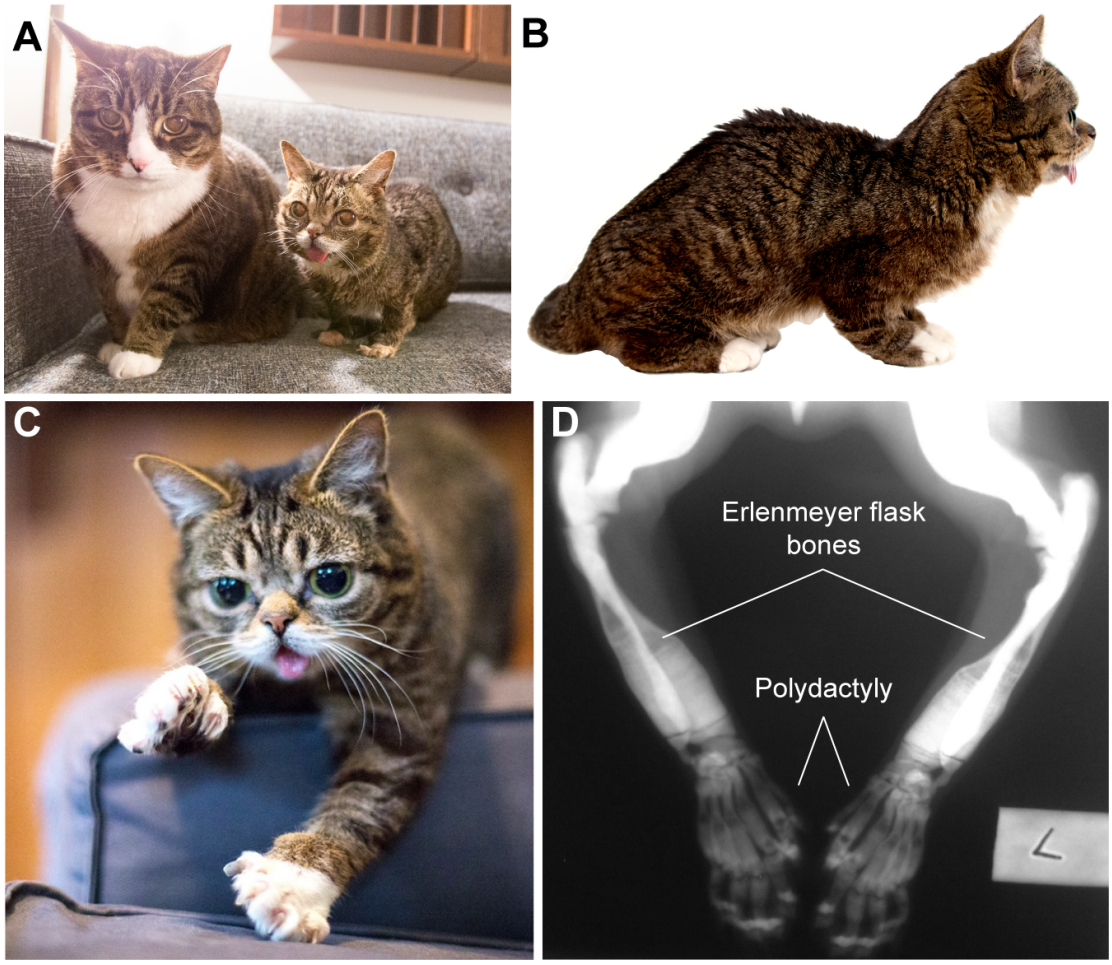

Figure 1. Phenotypic traits of Lil BUB. (A) Appearance of Lil BUB shows small size compared to an average adult domestic short haired cat (B) Short skull and body size, as well as protruding tongue shown in lateral view (C) Polydactyly of the front paws (D) Radiographs of forelimbs show high bone mineral density and Erlenmeyer flask shape of long bones as well as extra digits

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

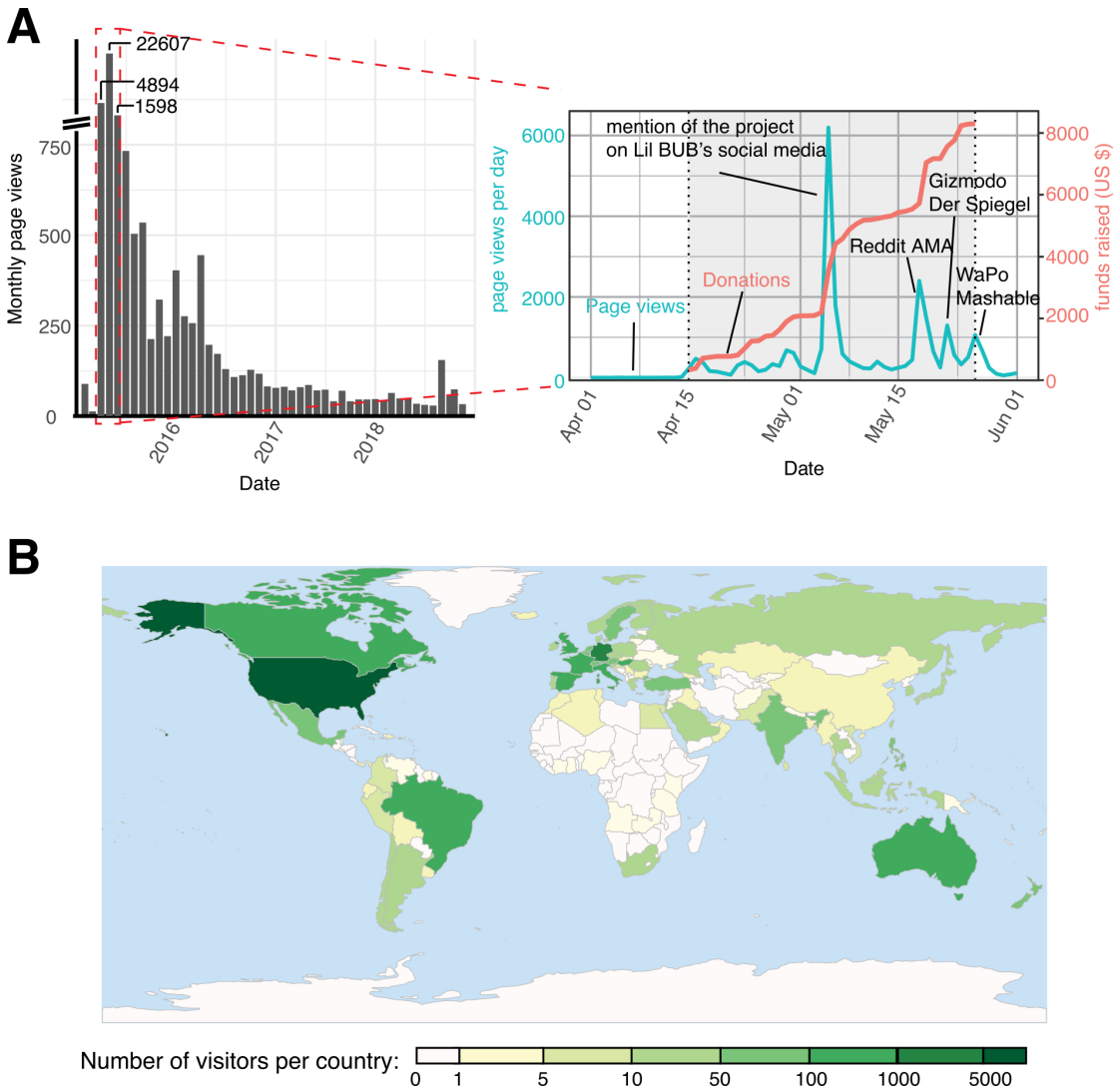

Figure 2. Web traffic on the homepages of the Lil BUBome. (A) Monthly number of visitors on the crowdfunding pages of the Lil BUBome at the experiment.com site between February 2015 and December 10, 2018 (left). It should be noted that for the three months with the highest number of visitors (April-June 2015) the height of the bars is not to scale. For these months we also show a daily breakdown of the site's traffic (right), as well as the progress of fundraising during the crowdfunding period (April 15 to May 25, time highlighted in grey). On the days where we observed a spike in visitors, we also indicate what major media outlets or social media sites reported about the Lil BUBome. (B) Geographic distribution of visitors to the Lil BUBome's Wordpress blog.

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

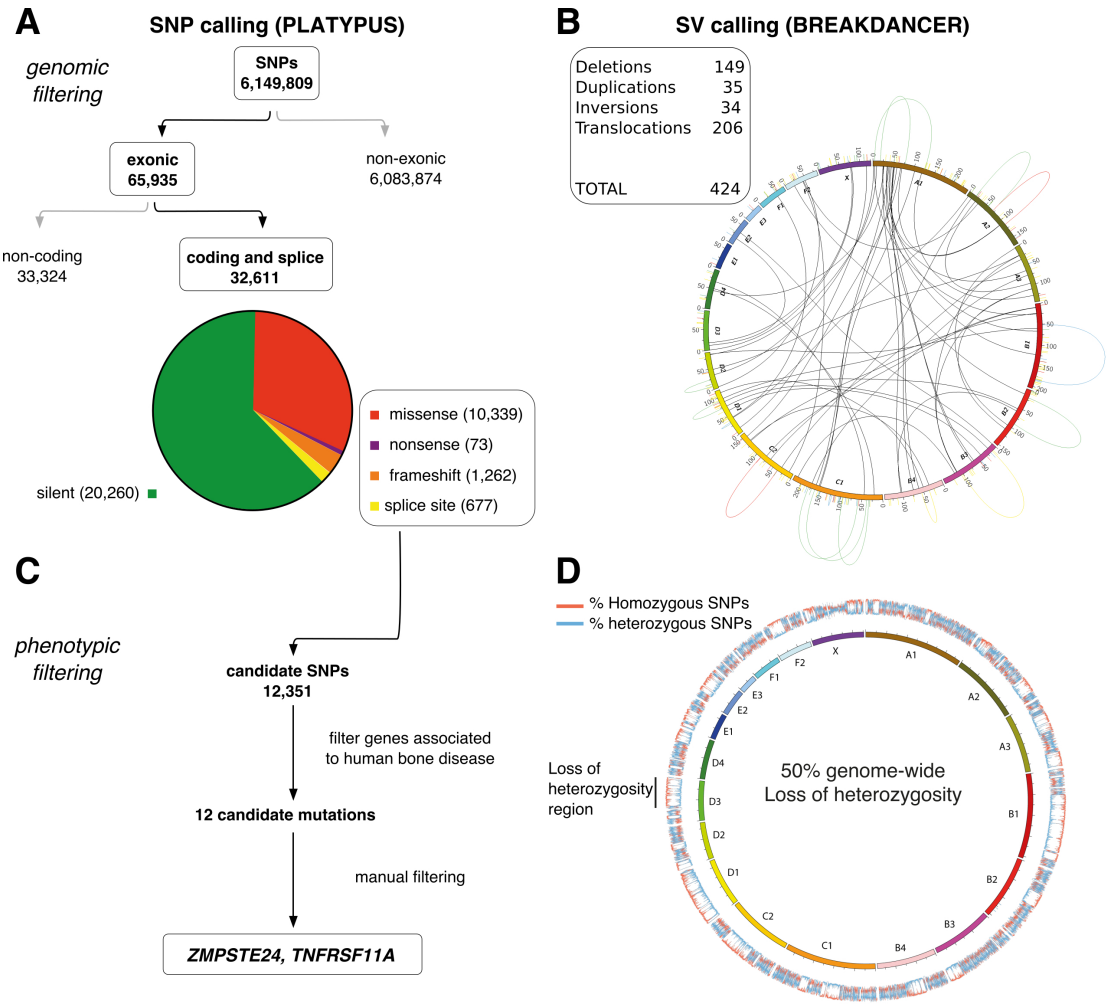

Figure 3. Variant detection in Lil BUB's genome sequence. (A, B) Number and type of single nucleotide variants (A) and structural variants (B) detected in BUB's genome relative to the reference cat genome, *Felis catus* v6.2, as determined using Platypus and Breakdancer tools, respectively. For each category, we indicate the number of observations in that category. The Circos plot for the structural variants shows the position and links of the detected structural variants. **(C)** Filtering of single nucleotide variants to identify candidate mutations for osteopetrosis. **(D)** Percentage of heterozygous and homozygous single nucleotide variants in 50kb windows along each chromosome show the genome-wide pattern of loss of heterozygosity.

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

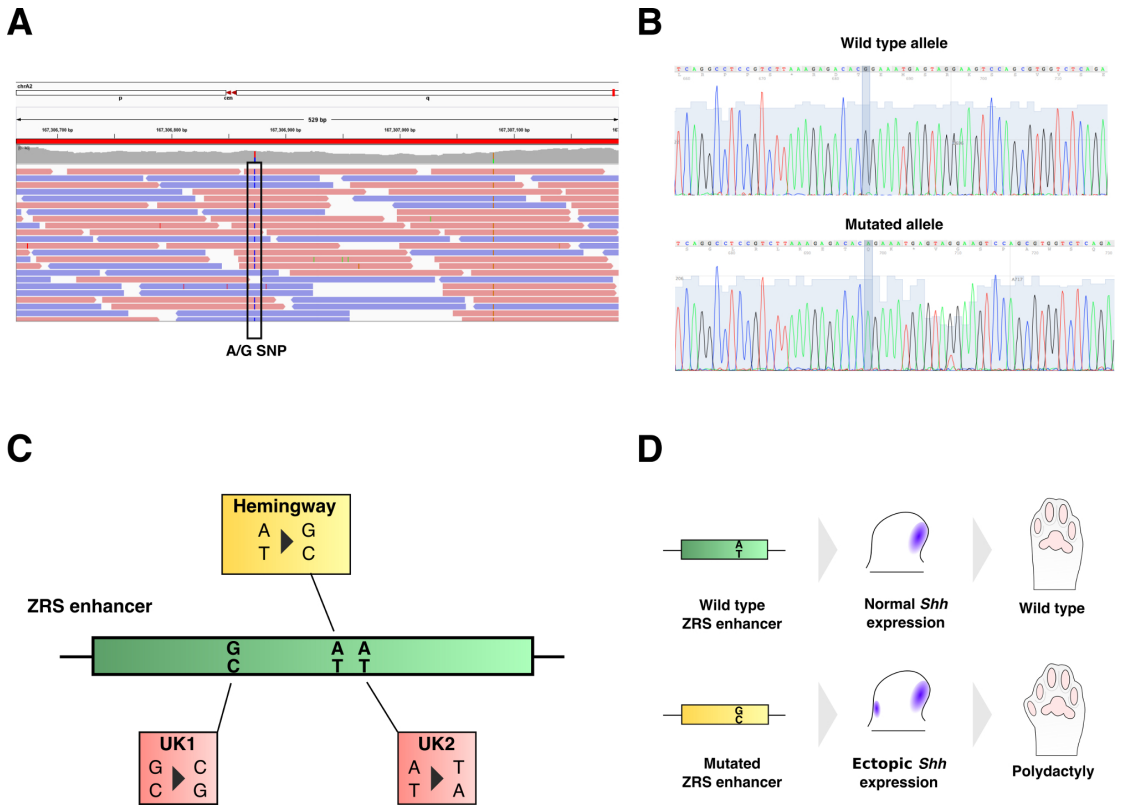

Figure 4. Polydactyly-associated mutations in the *Sonic hedgehog* limb enhancer. (A) A heterozygous A to G mutation at chrA2:167,313,488, in the limb enhancer of the *Sonic hedgehog* gene, as identified by whole genome sequencing in Lil BUB's genome sequence. **(B)** Verification of Lil BUB's mutation by Sanger sequencing, which showed the presence of a wild type allele (top), as well as a mutated allele (bottom). **(C)** Distribution of known polydactyly-associated mutations in the *Sonic hedgehog* limb enhancer in cats. **(D)** Developmental mechanism for polydactyly: The Hemingway-variant of the limb enhancer causes ectopic expression in the anterior part of the developing limb bud, giving rise to additional digits.

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

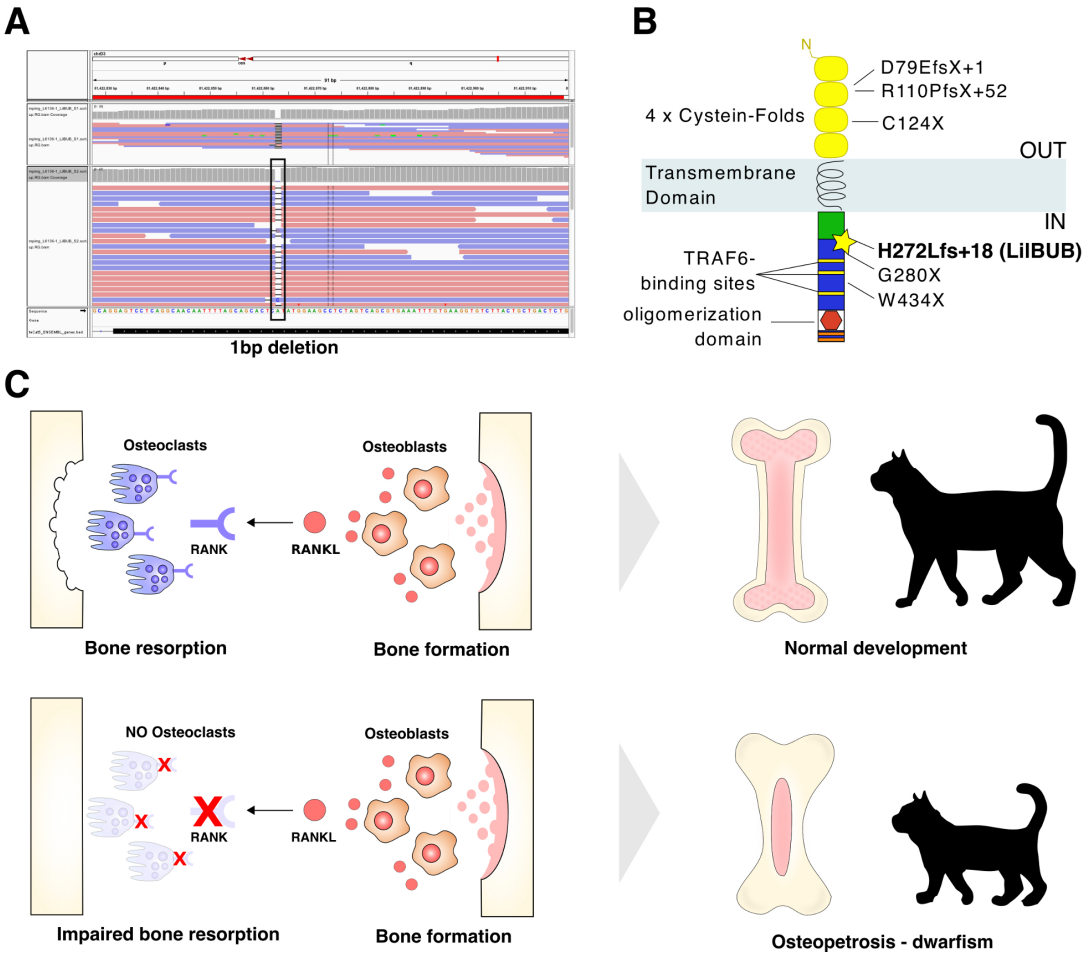

Figure 5. Osteopetrosis-associated variants in *TNFRSF11A*/RANK. (A) A homozygous 1-basepair deletion at chrD3:81,422,853, in exon 8 of *TNFRSF11A* (which codes for the protein RANK), as identified by whole genome sequencing in Lil BUB's genome sequence. **(B)** Schematic structure of the RANK protein and known human and mouse mutations. The position of Lil BUB's mutation is indicated by the star, and the two functional domains that are deleted due to the early truncation of the protein (the STAT-6 binding domain and the oligomerization domain) are also shown. **(C)** Developmental mechanism for osteopetrosis: bone formation and bone resorption are maintained by a balance of osteoblasts and osteoclasts, which express RANK-ligand (RANKL) and its receptor (RANK), respectively. A mutation that abolishes the function of RANK causes impaired bone resorption and thereby osteopetrosis and dwarfism.

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 1. Status of known feline phenotype-associated variants in Lil BUB's genomic sequence.

Variants	Breed	Inheritance	Gene	Reported variant	LilBUB	Publication
Coat color/trait variants						
black coloration	multiple	AR	*ASIP*	c.123delCA	HET for CA deletion	Eizirik et al, 2003
Cinnamon	multiple	AR	*TYRP1*	c.298C>T	C/C	Lyons et al, 2005 and Schmidt-Küntzel et al, 2005
Albino	Siamese	AR	*TYR*	c.901G>A	G/G	Schmidt-Künzel et al, 2005
Albino	Burmese	AR	*TYR*	c.679G>T	G/G	Schmidt-Künzel et al, 2005
Dilution	multiple	AR	*MLPH*	c.83delT	no deletion	Ishida et al, 2006
Dominant White (white fur, blue iris, hearing loss)	multiple	AD	*KIT*	FERV1 LTR in intron1 of KIT	no insertion	David et al, 2014
white Spotting	multiple	AD	*KIT*	full-length FERV1 in intron1 of KIT	no insertion	David et al, 2014
(white) gloving	Burman	AR	*KIT*	c.1035_1036delInsCA	no del/ins	Montague et al, 2014
Blotched tabby	multiple	AR	Taqpep	protein variants reported*	HET for 2 of the variants: W841X and T139N	Kaelin et al, 2012
hairless	Sphynx	AR	KRT71	c.816+1G>A	G/G	Gandolfi et al, 2010
curly	Devon Rex	AR	KRT71	complex**	WT/WT	Gandolfi et al, 2010
curly	Selkirk Rex	AD	KRT71	c.445-1G>C	G/G	Gandolfi et al, 2013
Disease variants and other traits						
progressive retinal atrophy	Abyssinians	AD	*CRX*	c.546delC	no deletion	Menotti-Raymond et al, 2010
progressive retinal atrophy	Abyssinians	AR	*CEP290*	IVS50+9T>G	T/T	Menotti-Raymond et al, 2007
progressive retinal atrophy	Persians	AR	*AIPL1*	c.577C>T	C/C	Lyons et al, 2016
feline hereditary myopathy	Devon Rex and Sphynx	AR	COLQ	c.1190G>A	G/G	Gandolfi et al, 2015
bobbed tail	Japanese Bobtail		HES7	c.5A>G	A/A	Lyons et al, 2016
Niemann-Pick-like disorder	single cat	-	NPC1	c.1322A>C	A/A	Mauler et al, 2017
osteochondrodysplasia	Scottish fold	AD	*TRPV4*	c.1024G>T	G/G	Gandolfi et al, 2016
Gangliosidosis Type 1	not specified	AR	GLB1	c.1448G>C	G/G	Martin et al, 2008
blood type B	multiple	AR	CMAH	complex haplotype***	blood type A haplotype	Bighignoli et al, 2007 and Omi et al, 2016
hypertrophic cardiomyopathy	Maine Coon	AD	MYBPC3	c.91G>C	G/G	Meurs et al, 2005
hypokalemia	Burmese	AR	WNK4	c.2899C>T	C/C	Gandolfi et al, 2012

*the reported variants are: p.S59X , p.W841X, p.T139N

** the reported variant consists of an 81-bp deletion (c.1108-4_1184del), 8-bp insertion (c.1184_1185insAGTTGGAG), base insertion (c.1196insT))

*** the haplotype includes: c.-217G>A, c.-371C>T, c.-70insaacgagcaaccgaagctg, c.139G>A, c.265T>A, c.1600G>A, c.1266G>A + more…

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 2. Shortlisted candidate genes for Lil BUB's osteopetrosis.
Yellow indicate variants that pass quality filters but are discarded due to unclear annotation in cat genome. Green indicate variants that pass quality filters and further retained for examination.

Chr	Start	reference	variant	SnpEff score	Gene symbol	Allele status	Filtering	Exclusion criteria
A2	154754211	C	CTT	24	TBXAS1	HET	SC;badReads;HapScore;MQ;QD	Low quality reads
E1	37622020	A	AC	66	DLX3	HET	SC;badReads;MQ	Low quality reads
A2	15980613	CT	C	197	PTH1R	HOM	PASS	Intronic in newer annotations
A2	154754199	CTTTTTTCTTTCCT	CTTTTTTTCTTTCCT, TTTTTTTTTTTTT	273	TBXAS1	Low complexity	HapScore;MQ	Low quality reads
E2	35128777	CCA	C	503	MMP2	HOM	MQ	Low quality reads
C2	155537049	CG	C	792	CRTAP	HOM	PASS	Unclear annotation
D1	109006572	C	CT	827	TRPT1	HOM	badReads	Low quality reads
D1	116768932	C	T	1076	ATHL1	HOM	badReads	Low quality reads
C1	30324507	A	AT	1225	ZMPSTE24	HET	PASS	No exclusion
D1	116768902	T	TA	1554	ATHL1	HOM	PASS	Intronic
A2	118427509	G	GA	1810	SNX10	Low complexity	SC	Low quality reads
D3	81422862	CA	C	1899	TNFRSF11A	HOM	PASS	No exclusion

quality pass variant
candidate variant

Table S1. A sample of media outlets that reported on about the sequencing of Lil BUB's genome around the world.

Table S2. Genes with known role in bone mass regulation and/or osteoclast differentiation in humans.

Acknowledgements

We thank all the generous donors who contributed to our crowdfunding campaign, and without whom this project would not have been possible. We are also grateful for the guidance we received from experiment.com staff, especially Cindy Wu and Denny Luan, who helped us navigate through the funding process and aided our communication endeavors. We would further like to express our gratitude to our mentors, Stefan Mundlos, Gaël Yvert and Arjun Raj, who supported our work on this project. We would also like to thank Sarah Malek and Karun Kiani for providing critical feedback on the manuscript. Funding for the 99 Lives Cat Genome Sequencing Initiative has been provided by the Winn Feline Foundation and the George Sydney and Phyllis Redman Miller Trust (MT–13–010), the National Geographic Society Education Foundation (2P-14), the University of Missouri Gilbreath-McLorn Endowment (LAL), the Cornell Feline Health Center (ARB, MGC, RJT), contributions from the 99 Lives Consortium participants, donations from Associazione Nazionale Felina Italiana (ANFI), Zoetis, Orivet Genetic Pet Care, Langford Veterinary Services, the World Cat Federation, and public donations. We

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

also wish to thank Illumina for providing a complimentary NextSeq 500 and Library Prep kit.

Author contributions

D.M.I, D.G.L. and O.S. conceived the project, M.B. and A.W. performed disease diagnosis. M.B., U.W., B.T. and A.W. provided reagents. D.M.I., D.G.L and N.M. performed experiments, H.K., D.M.I., D.G.L. and O.S. analyzed the data. The 99 Lives Consortium provided the variant genetic database. D.M.I., D.G.L. and O.S. wrote the manuscript with feedback from all other authors.

99 Lives Consortium participants

Leslie A. Lyons[1], Danielle Aberdein[2], Paulo C. Alves[3,4], Gregory S. Barsh[5,6], Rebecca R. Bellone[7], Tomas F. Bergström[8], Adam R. Boyko[9], Jeffrey A. Brockman[10], Margret L. Casal[11], Marta G. Castelhano[12], Ottmar Distl[13], Nicholas H. Dodman[14], N. Matthew Ellinwood[15], Jonathan E. Fogle[16], Oliver P. Forman[17], Dorian J. Garrick[2,15], Edward I. Ginns[18], Jens Häggström[19], Robert J. Harvey[20], Daisuke Hasegawa[21], Bianca Haase[22], Christopher R. Helps[23], Isabel Hernandez[12], Marjo K. Hytönen[24], Maria Kaukonen[24], Christopher B. Kaelin[5,6], Tomoki Kosho[25], Emilie Leclerc[26], Teri L. Lear[27], Tosso Leeb[28], Hannes Lohi[24], Maria Longeri[29], Ronald H. L. Li[30], Mark A. Magnuson[31], Richard Malik[32], Shrinivas Mane[33], John S. Munday[2], William J. Murphy[34], Niels C. Pedersen[35], Simon M. Peterson-Jones[36], Max F. Rothschild[15], Clare Rusbridge[37], Beth Shapiro[38], Joshua A. Stern[35], William F. Swanson[39], Karen A. Terio[40], Rory J. Todhunter[12], Yu Ueda[35], Wesley C. Warren[41], Elizabeth A. Wilcox[12], Julia H. Wildschutte[42], Yoshihiko Yu[21].

[1]Department of Veterinary Medicine and Surgery, College of Veterinary Medicine, University of Missouri, Columbia, Missouri, 65211 USA

[2]School of Veterinary Biomedical Science, Massey University, Palmerston North 4442 New Zealand

[3]CIBIO/InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos/InBIO Associate Lab & Faculdade de Ciências, Universidade do Porto, Campus e Vairão, 4485–661 Vila do Conde, Portugal.

[4]Wildlife Biology Program, University of Montana, Missoula, Montana, 59812 USA.

[5]HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, 35806 USA

[6]Department of Genetics, Stanford University, Stanford, California, 94305 USA

[7]Veterinary Genetics Laboratory, University of California, Davis, Davis California, 95616 USA

[8]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, 750 07 Uppsala, Sweden

[9]Department of Biomedical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, New York, 14853 USA

[10]Hill's Pet Nutrition Inc., Topeka, Kansas 66601 USA

[11]Reproduction, and Pediatrics, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA 19104 USA

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[12]Department of Clinical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, New York, 14853 USA

[13]Institute for Animal Breeding and Genetics, University of Veterinary Medicine Hannover (Foundation), 30559, Hannover, Germany

[14]Department of Clinical Sciences, Cummings School of Veterinary Medicine, Tufts University, Grafton, MA, 01536 USA

[15]Department of Animal Science, College of Agriculture and Life Sciences, Iowa State University, Ames, Iowa, 50011 USA

[16]College of Veterinary Medicine, North Carolina State University, Raleigh, NC 27607

[17]WALTHAM Centre for Pet Nutrition, Freeby Lane, Waltham on the Wolds, Leicestershire, LE14 4RT UK

[18]Department of Psychiatry, University of Massachusetts Medical School, Worcester, MA, 01655 USA

[19]Department of Clinical Sciences, Faculty of Veterinary Medicine and Animal Science, Swedish University of Agricultural Sciences, Uppsala, SE-750 07 Sweden.

[20]School of Health and Sport Sciences, University of the Sunshine Coast, Sippy Downs, QLD, Australia

[21]Department of Clinical Veterinary Medicine, Nippon Veterinary and Life Science University, Tokyo 180-8602 Japan

[22]Sydney School of Veterinary Science, Faculty of Science, University of Sydney, Sydney, NSW, 2006, Australia

[23]Langford Vets, University of Bristol, Langford, Bristol, BS40 5DU UK

[24]Department of Veterinary Biosciences and Department of Medical Genetics, University of Helsinki and Folkhälsan Research Center, Helsinki 00014 Finland

[25]Department of Medical Genetics, Center for Medical Genetics, Shinshu University Hospital, Matsumoto, Nagano 390-8621, Japan

[26]SPF - Diana Pet food – Symrise Group – 56250 Elven, France

[27]Department of Veterinary Science, University of Kentucky - Lexington, Lexington, KY, 40506 USA (*In memoriam*)

[28]Vetsuisse Faculty, Institute of Genetics, University of Bern, 3001 Bern, Switzerland

[29]Dipartimento di Medicina Veterinaria, University of Milan, 20122 Milan, Italy

[30]Department of Surgical and Radiological Sciences, School of Veterinary Medicine, University of California at Davis, Davis, California, 95616 USA

[31]Departments of Molecular Physiology and Biophysics, Cell and Developmental Biology, and Medicine, Vanderbilt University, School of Medicine, Nashville, Tennessee, 37232 USA

[32]Centre for Veterinary Education, University of Sydney, Sydney, NSW, 2006 Australia

[33]Elanco Animal Health, Greenfield, IN 46140 USA

[34]Department of Veterinary Integrative Biosciences, College of Veterinary Medicine, Texas A&M University, College Station, Texas, 77845 USA

[35]Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California at Davis, Davis, California, 95616 USA

[36]Small Animal Clinical Sciences, College of Veterinary Medicine, Michigan State University, East Lansing, Michigan, 48824 USA

[37]School of Veterinary Medicine, Faculty of Health & Medical Sciences, Vet School Main Building (VSM), Daphne Jackson Road, Guildford, Surrey, GU2 7AL, United Kingdom

bioRxiv preprint doi: https://doi.org/10.1101/556761; this version posted February 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[38]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, Santa Cruz, Calfornia 95064 USA.

[39]Center for Conservation and Research of Endangered Wildlife (CREW), Cincinnati Zoo & Botanical Garden, Cincinnati, Ohio, 45220 USA

[40]Zoological Pathology Program, University of Illinois, Brookfield, Illinois 60513 USA

[41]Division of Animal Sciences, College of Agriculture, Food and Natural Resources; School of Medicine, University of Missouri, Columbia, Missouri 65211, USA;

[42]Bowling Green State University, Department of Biological Sciences, Bowling Green, Ohio 43403 USA