

Review

# The Value of Reference Genomes in the Conservation of Threatened Species

Parice Brandies, Emma Peel, Carolyn J. Hogg and Katherine Belov \*

School of Life & Environmental Sciences, The University of Sydney, Sydney 2006, Australia;  
parice.brandies@sydney.edu.au (P.B.); emma.peel@sydney.edu.au (E.P.); carolyn.hogg@sydney.edu.au (C.J.H.)

\* Correspondence: kathy.belov@sydney.edu.au

Received: 22 September 2019; Accepted: 23 October 2019; Published: 25 October 2019



**Abstract:** Conservation initiatives are now more crucial than ever—over a million plant and animal species are at risk of extinction over the coming decades. The genetic management of threatened species held in insurance programs is recommended; however, few are taking advantage of the full range of genomic technologies available today. Less than 1% of the 13505 species currently listed as threatened by the International Union for Conservation of Nature (IUCN) have a published genome. While there has been much discussion in the literature about the importance of genomics for conservation, there are limited examples of how having a reference genome has changed conservation management practice. The Tasmanian devil (*Sarcophilus harrisii*), is an endangered Australian marsupial, threatened by an infectious clonal cancer devil facial tumor disease (DFTD). Populations have declined by 80% since the disease was first recorded in 1996. A reference genome for this species was published in 2012 and has been crucial for understanding DFTD and the management of the species in the wild. Here we use the Tasmanian devil as an example of how a reference genome has influenced management actions in the conservation of a species.

**Keywords:** conservation; genomes; Tasmanian devil

## 1. Introduction

We are currently in the midst of a global sixth mass extinction event, with biodiversity rapidly declining around the world [1], and extinction rates are accelerating [2]. Australia has the worst mammal extinction rate of any country, with 25 mammals declared extinct since European settlement and almost 20% of current mammalian species listed as vulnerable [2–5]. This significant decline is concerning as Australia is one of seventeen “megadiverse” countries that comprises a large proportion of the Earth’s biological diversity [6]. Megadiverse countries have at least 5,000 endemic plant species and have marine ecosystems within its borders [6]. In addition to this, 87% of Australian mammals, 93% of Australian reptiles, and 94% of Australian frogs are endemic to Australia [7]. Therefore, conservation initiatives that protect and maintain Australia’s biodiversity are now more crucial than ever.

Only 39% of the 1,890 Australian species (517 animals; 1373 plants), listed as threatened under the Environment Protection and Biodiversity Conservation Act (EPBC Act), have a recovery plan in place to improve their threat status [8]. These recovery plans set out management and research actions to slow population decline and promote recovery of threatened species and communities. This is achieved by providing a framework for key interest groups and government agencies to coordinate their efforts to improve the plight of threatened species [8]. Management actions range from mitigating threatening processes such as predation, habitat loss, or change, in addition to research into basic species biology, ecosystem integration, and genetics. The main goal of recovery plans is to maintain the long-term viability of a chosen population/community. Maintaining genetic diversity is an important component of population viability as it assists with mitigating negative effects

associated with inbreeding and arms populations with the **potential to adapt to future environmental change** [9–11]. As such, understanding a populations' inherent genetic diversity, in addition to their historical diversity and future potential, is of utmost importance in species conservation. For this reason, more than 80% of the current 200 Australian national vertebrate recovery plans have some form of genetic action listed in the species' recovery plan. Yet, less than 15% of these recovery plans have any form of genetic or genomic data available, either in existence or currently in development. Here we refer to genetic data as information based on specific, limited regions of the genome (e.g., targeted gene sequencing, microsatellite analysis, etc.), whilst genomic data is information based on the whole genome (e.g., whole genome sequencing/resequencing, whole-genome single nucleotide polymorphism (SNP) analysis/reduced representation sequencing, etc.).

Advances in sequencing technologies and the reduction in sequencing costs have given rise to the era of genomics, whereby **holistic genome-wide approaches are rapidly replacing traditional genetic marker approaches in many non-model species** [12–14]. Although recent reviews have highlighted the importance of implementing genomic data into conservation initiatives [13,15,16], the application of such powerful advances in sequencing technologies is lacking in the current literature. This limited use in conservation may be due to a number of reasons including: **costs, a lack of understanding of the potential of new genomics approaches, lack of expertise in developing and utilizing the data, and the absence of a reference genome for the species of interest (or a closely-related species)** [13,15,17]. The latter is an important concern as the generation of a reference genome requires considerable expertise, funds, computational resources, and time that are not often accessible by wildlife managers and conservation teams [15,18].

Of the 13505 animal species that are listed as threatened (Lower Risk/Conservation Dependent or worse) on the International Union for Conservation of Nature (IUCN) Red List [2], 108 (< 1%) have published genomes on NCBI [19]. This equates to only 6% of the 1842 animal genomes currently available on NCBI [19]. Creating high-quality reference genomes that can provide insights into species evolution and biology is a costly task (~\$30,000 for an average eukaryotic genome size of 2.5 Gbp [20]), and also requires large collaborative groups to provide expertise from varying fields (e.g., [21–23]). Fortunately, in recent years a number of national and international consortia and genome projects have been formed with the aim of creating high-quality reference genomes for species spanning the phylogenetic tree of life including: **the Earth Biogenome Project (EBP) [20], the Genome 10K Project (G10K) [24,25], the Vertebrate Genomes Project (VGP) [26], the Bird 10K Project (B10K) [27], the Bat 1K Project (Bat1K) [28], the Global Invertebrate Genomics Alliance (GIGA) [29,30], and the Oz Mammal Genomics initiative (OMG) [31]**, to name a few. The goal of many of these consortia is to bring together the required expertise to generate reference genomes of a sufficient quality, which are publicly available to the science community, thereby providing the vital resources required to implement genomics into conservation management better [13,15,18]. However, just providing the reference genomes or genomic data is not enough to improve conservation outcomes. Geneticists need to continually communicate how genomic techniques can be utilized in a cost-effective manner to assist species conservation better [17,32]. As highlighted by Taylor et al. [33], targeted education and training is also required to teach conservation managers how to interpret and utilize genomic data. To better assist conservation managers, a number of groups and communities have already been established to assist in providing conservation genetics advice for threatened species management. These include the IUCN/SSC (Species Survival Commission) Conservation Genetics Specialist Group (CGSG), the Genetic Composition Working Group of GEO BON (Group on Earth Observations Biodiversity Observation Network), and the pan-European COST (Cooperation in Science and Technology) action ConGRESS (Conservation Genetic Resources for Effective Species Survival) (for further information and examples from these groups, see Holderegger et al. [34]). Conservationists in their respective countries can get in touch with these groups to obtain the contact details of geneticists who work in their region who may be able to assist them with their management needs.

While a number of papers have reviewed current genomic techniques and the way they can, or have been, applied to assist in conservation decisions across species [15,17], questions are still raised as to whether reference genomes are necessary for species conservation. Reference genomes hold the key to investigate a number of paradigms that are essential for species conservation, including: demography, inbreeding, hybridization, disease susceptibility, behavioral ecology, and adaptation [12,13,15,16,18]. Here we demonstrate the value of a reference genome to the conservation effort of an endangered species, the Tasmanian devil (*Sarcophilus harrisii*), and how this information has been applied in real-time management practice [35].

The Tasmanian devil, an endangered Australian marsupial, is often used in the literature as an example of how genetics/genomics approaches could be used in conservation [12,13,36]. However, something that is not often discussed is that having a reference genome for this species is one of the key factors that contributed to using genomics in management practice. Although this species has a unique conservation issue, low genetic diversity coupled with an infectious clonal cancer, the methods described herein apply to many other threatened species. Here we show how the reference genome has allowed a range of conservation questions to be answered in a timely, cost-effective manner and enabled conservation researchers to adapt to the rapid advances in genomic technologies.

## 2. The Tasmanian Devil and Its Genome

The Tasmanian devil is the largest extant carnivorous marsupial, native to mainland Tasmania, Australia. The emergence of transmissible cancer, devil facial tumor disease (DFTD) in the mid-1990s has led to a rapid population decline of up to 80% across their range [37]. In 2003, the Tasmanian and Australian governments responded to the disease threat by establishing the Save the Tasmanian Devil Program (STDP). Since then, researchers, wildlife managers, and the zoo industry have worked closely with the STDP to ensure that Tasmanian devils have a sustainable ecological function in the Tasmanian ecosystem and landscape [35,38]. This work has included a range of activities such as monitoring of wild populations, developing an insurance population, describing and characterizing the disease, and developing new genomic tools to understand the disease and the Tasmanian devil [38].

Prior to the publication of a reference genome for the Tasmanian devil, traditional genetic approaches such as MHC (major histocompatibility complex) typing and microsatellite analysis were used to explore genetic diversity at specific genes as well as general genetic diversity in the species [39–41]. These techniques were able to show that the Tasmanian devil had low genetic diversity [39–42]. However, the low rates of polymorphism for most of these markers did not have high enough resolution to assist in answering crucial conservation questions such as determining founder relatedness within the insurance population [43,44], identifying high-resolution population substructure [45], or to better understand the origin and evolution of DFTD [46]. In instances such as these, further genomic data was required to improve resolution. For other threatened species, where there may be moderate to high genome-wide diversity, microsatellite markers may be highly polymorphic, and so these markers have value as a continuing genetic management tool.

To overcome this knowledge gap, the Tasmanian devil genome was sequenced independently by two different research groups in 2011 [45,46]. Miller et al. [45] sequenced the nuclear genome of two individuals (originating from extreme northwest and southeast Tasmania), as well as the tumor from one individual, using both Roche and Illumina sequencing platforms. The analysis of genome-wide SNPs confirmed low genetic diversity across the Tasmanian devil genome, as well as enabling the construction of genotyping arrays, which revealed a new population substructure and the identification of tumor-specific SNPs. However, the low contiguity of this reference genome assembly (148,891 scaffolds, scaffold N50 147 kb) limited the applicability of the data in downstream research. In 2012, a more contiguous, annotated nuclear genome (35974 scaffolds, scaffold N50 1.85 Mb), and tumor genome was published by Murchison et al. [46], resulting in the primary reference genome used today. This higher quality assembly facilitated an enormous effort in downstream genetic and genomic research. It should be noted that as of August 2019, the 2012 Tasmanian devil reference genome

paper [46] has been cited over 200 times (Google Scholar Citation Search), highlighting the value of this reference genome to the research community. It is not possible to cover all of the research that has stemmed from the sequencing of the 2012 genome here. Rather, here, we present key examples of how having a reference genome has contributed to conservation decisions and outcomes for the Tasmanian devil. We also note that at the time of this publication, an updated Tasmanian devil genome assembly has been released [47]. This assembly utilized an in vitro proximity ligation technique to further improve the scaffolding of the 2012 assembly (10010 scaffolds, N50 7.75 Mb); however, chromosome assignment and annotation have not been performed at this stage.

### 3. Conservation Applications as a Result of a Reference Genome

#### 3.1. Basic Conservation Management

##### 3.1.1. Microsatellite Analysis

Traditionally population genetic measures to answer basic questions regarding population structure, population size, population dynamics (migration, bottlenecks), kinship, inbreeding, etc. [14,48] have used microsatellites, or short tandem repeats [48]. Where microsatellite markers have already been developed for the species of interest, or in a closely related species that may carry similar markers, they provide a cost-effective, quick conservation management tool [48,49]. However, for those species where appropriate microsatellite markers are not currently available, or cross-species microsatellite amplification is not effective, and a reference genome is also not available, considerable time and resources are required to develop species-specific microsatellite markers. For example, prior to sequencing the Tasmanian devil genome, 11 putatively neutral microsatellite markers were developed to assess genetic diversity in Tasmanian devils [39]. The development of these microsatellites involved the creation and screening of a genomic library, sequencing of positive clones, primer design, and PCR optimization [39]. Several years later, MHC-linked microsatellite markers were developed in a similar manner as a cheaper and faster method of investigating MHC diversity when compared to traditional MHC typing techniques, such as cloning and sequencing particular MHC regions [41]. This traditional microsatellite isolation and the marker development approaches require considerable laboratory expertise, time, and funds [49], that today may be better spent developing more powerful molecular approaches (see Reduced Representation Sequencing section below).

Contrarily, the availability of the Tasmanian devil reference genome enabled 22 additional microsatellite markers to be identified and developed in a much faster, cost-effective manner using bioinformatic methods [50]. More importantly, each of these microsatellites were known to be in non-coding regions across all of the autosomes, providing a greater representation of neutral genome-wide diversity in comparison to the original 11 putatively neutral microsatellites. It has previously been estimated that the development of just 10 microsatellite markers without prior genetic data can cost up to \$10000 [51]. The availability of a reference genome mitigates the need for traditional microsatellite isolation procedures, and therefore, significantly reduces costs associated with marker development (< \$1000 for primer optimization and testing). Additionally, the commercial development of microsatellite-based PCR kits resulted in further reductions in the time and cost associated with microsatellite marker development and use [50]. To date 33 microsatellite markers have successfully been applied to Tasmanian devil conservation to investigate inbreeding [50], reconstruct the pedigree of offspring born in group housing and on Maria Island [50,52–54], and investigate mate choice within captivity and the wild [55] (Table 1). These microsatellite markers have also successfully been applied to genotype individuals using non-invasive scat samples [56], which are notoriously known for producing low quantities of low-quality DNA [57]. Globally, microsatellite markers continue to be an effective tool in conservation decision making by answering population questions [58–62]. They are particularly valuable when using non-invasive samples that are often unsuitable for more complex genomic methods that require high-quality input DNA, such as reduced representation sequencing and other whole-genome sequencing methods [15]. A reference genome allows for fast, easy, and inexpensive development of such markers, improving their utility in the conservation management space.

**Table 1.** Examples of Tasmanian devil conservation questions, actions, and outcomes that have been facilitated by the reference genome.

Reference Genome Use	Conservation Questions Addressed	Conservation Actions	Conservation Outcomes
<ul style="list-style-type: none"> <li>• Microsatellite development</li> <li>• <b>Genome-wide SNP analysis</b></li> </ul>	<ul style="list-style-type: none"> <li>• Were the founders related?</li> <li>• <b>Does the metapopulation have equal founder representation to ensure the maintenance of gene diversity?</b></li> <li>• <b>Is inbreeding accumulating in group housing and Maria island insurance populations?</b></li> <li>• How many DFTD strains exist?</li> </ul>	<ul style="list-style-type: none"> <li>• Resolved relatedness of founders [43]</li> <li>• Resolved parentage in group housing within the metapopulation [50,52,54]</li> <li>• Reconstructed pedigree of island population [53]</li> <li>• <b>Informed translocation recommendations [63]</b></li> <li>• <b>Appropriate management of wild populations [46,64,65]</b></li> <li>• Immunization development and deployment [66]. Immune gene diversity analysis for informed translocation recommendations [67–75]</li> </ul>	<ul style="list-style-type: none"> <li>• Tool for selecting individuals for translocations based on genetic complementation</li> <li>• <b>Improved maintenance of genetic diversity across captive populations</b></li> <li>• Increased genetic diversity of hybrid individuals at wild release sites</li> <li>• Assisted in managing the spread of new DFTD strains</li> <li>• Improved immune responses of devils released to the wild</li> <li>• Improved immunogenetic diversity of released Tasmanian devils and their resultant offspring</li> <li>• <b>Mitigation implemented to reduce the impact on highly consumed species</b></li> <li>• <b>Assisted in understanding regions of the genome that are potentially involved in DFTD resistance</b></li> </ul>
<ul style="list-style-type: none"> <li>• The characterization of DFTD strains</li> </ul>	<ul style="list-style-type: none"> <li>• Can we develop a vaccine for DFTD?</li> <li>• Can we improve Tasmanian devil immune diversity?</li> </ul>	<ul style="list-style-type: none"> <li>• Investigating the impact of an introduced carnivore to island wildlife</li> <li>• <b>Ongoing monitoring to ensure releases do not impact the evolution of potential resistance alleles [76–79]</b></li> </ul>	
<ul style="list-style-type: none"> <li>• The characterization of immune genes</li> <li>• Primer design and SNP panel development</li> <li>• Targeted SNP analysis</li> <li>• Development of blocking primer for metagenomics diet analysis</li> <li>• Alignment of resequenced genomes</li> <li>• SNP Analysis and Annotation</li> <li>• <b>GWAS</b></li> </ul>	<ul style="list-style-type: none"> <li>• What constitutes the complete diet of Tasmanian devils on Maria Island?</li> <li>• Are devils evolving host-parasite resistance to DFTD?</li> </ul>		



### 3.1.2. Reduced Representation Sequencing

While microsatellite analysis is one of the most common population genetics tools, sometimes more statistical power is needed to address specific conservation management questions, particularly in species with low genetic diversity [43,80,81]. For instance, in the Tasmanian devil, microsatellite analysis was unable to accurately estimate the relatedness of founders sourced for the insurance population between 2006 and 2008 [43]. Single nucleotide polymorphisms (SNPs) enable greater resolution for addressing some common conservation issues such as resolving parentage and population structure, understanding genetic diversity, and identifying regions of the genome, which may be linked to important phenotypes [42]. When compared to a microsatellite approach, only 3–8 biallelic SNPs are required to be as informative as one microsatellite marker [82,83]. Reduced representation sequencing (RRS) is a simple, cost-effective approach for generating genome-wide SNP data and is gaining popularity in the conservation sector [15,42,84]. RRS relies on high-throughput sequencing of fragments generated by restriction enzyme digestion of the genome and can, therefore, easily be applied in any species. There are a variety of RRS methods currently available, including traditional RADseq [85], ddRAD [86], DArTseq [87], and others [42].

Both DArTseq and RADseq have been employed to collect RRS data from over 1,000 Tasmanian devils from the insurance population, Maria island and a number of wild sites [76,77,84,88,89]. RRS methods have shown to be superior in accurately estimating diversity and inferring genome-wide heterozygosity compared with microsatellite analysis and other targeted techniques [89]. Although this approach does not require a reference genome for development and use, coupling RRS data with a reference genome is advantageous in that it: i) improves the reliability of genotype calls [90]; ii) reduces the required coverage for accurate genotyping [91]; iii) provides for a greater number of SNPs [92]; iv) improves downstream population genetic inferences [92]; v) allows for SNP annotation with gene information [93]; and vi) provides the ability to compare results from differing RRS methods which are particularly important when different methods are used across time for endangered species.

Using a reference genome guided approach in the Tasmanian devil enabled 2060 SNPs to be identified [84] much more quickly than a de novo approach. Aligning the RRS data to the reference genome provides the ability to identify genes which may be targets of future analysis, and to separate functional vs. non-functional genome diversity which could have conservation implications [94]. For example, the reference genome was able to identify candidate genes within a genomic region that displayed signatures of selection in RRS data [76], and to identify cancer-resistance candidate genes from phenotype association tests of RRS data [77] (Table 1). A number of non-synonymous SNPs have also been identified within particular genes, which have the potential to impact phenotype. Furthermore, reference alignment allows SNPs from alternative RRS datasets to be compared and combined, such as the DArTseq and RADseq data, which are important for reusing previous investments of limited conservation dollars. Recent work investigating New Zealand threatened bird species also showed the benefits of calling SNPs against conordinal, confamilial, cogenetic, and conspecific reference genomes [95]. This highlights that not every threatened species requires a reference genome, although the quality of the SNP data reduces as you move away from the genus and family level.

## 3.2. Further Species-Specific Applications

### 3.2.1. Reference Gene Characterization

A valuable advantage of having access to a reference genome is the ability to characterize particular genes, or gene families, that are relevant to species-specific conservation [23]. Gene characterisation is often undertaken in two main ways: in-depth, manual characterization of a specific set of genes of interest, and automatic, whole-genome annotation. The latter is achieved in two main stages: the computational phase and the annotation phase [96,97]. During the computational phase, initial gene predictions are based on several lines of evidence including transcriptome and protein data from the species of interest and several closely-related, or well-annotated species [96,97]. During the

annotation phase, the most representative gene predictions (defined by the annotation pipeline) are synthesized into the final gene annotations [96,97]. The whole-genome annotation of the Tasmanian devil reference genome was achieved using the Ensemble genome annotation pipeline [46,98,99]. This automatic annotation of 18775 protein-coding genes was critical to the development of targeted SNP panels to explore diversity at important immune genes in the Tasmanian devil [69–71] (see SNP Panel section below), and in the identification of genes that may be linked to DFTD [46,76–78,100] (Table 1).

While modern-day tools, such as trainable automated gene prediction algorithms, have increased the feasibility of genome annotation of newly sequenced species within individual research groups, complete genome annotation still requires considerable bioinformatics expertise [96,97]. Manual annotation of a subset of target genes is often required. This is particularly relevant for genes that have experienced duplications and are, therefore, often unable to be automatically annotated [23,96]. In the Tasmanian devil, this was true for a number of gene families, including the Major Histocompatibility Complex (MHC), toll-like receptors (TLR), natural killer (NK) receptors, cathelicidins, behavior, and reproductive genes which were all manually annotated [69,72,75,101,102]. Annotation of these genes was essential in facilitating species-specific downstream research and informing conservation management decisions in the Tasmanian devil, such as genetic variation analyses [69,70,72,75]; selection of individuals for release to the wild [63], individuals response to the immunotherapy [66]; changes of immune function with the onset of puberty [73]; and the influence of age and DFTD on immune function [74] (Table 1). This highlights the potential of a reference genome for exploratory analysis of gene families involved in key biological processes of threatened species such as immunity, reproduction, and behavior.

### 3.2.2. Targeted SNP Panels

Targeted SNP panels enable diversity at particular genes to be investigated based on current conservation concerns/questions [103]. In the Tasmanian devil, an SNP panel targeting immune, behavioral, and putatively neutral loci was developed and used to genotype over 300 individuals in the insurance population [71]. This involved low-coverage resequencing of a number of individuals (see the Whole-Genome Resequencing section below), alignment of data to the reference genome, identification of target SNPs, primer design, pilot sequencing, and final genotyping. The SNP panel resolved parentage with higher confidence than microsatellite markers and also provided representative measures of genetic diversity at both functional and non-functional loci [71]. Development of another SNP panel, which targeted a range of immune genes, showed considerably low immune diversity in the species [70], which has led to further research into ways of breeding Tasmanian devils to improve genome-wide heterozygosity and functional diversity [67,68]. The Tasmanian devil reference genome was essential for aligning sequencing data and target SNP discovery allowing for management decisions to be based on both genome-wide and functional diversity (Table 1). Although custom SNP panel development can be expensive and is not simple, once developed it provides fast, accurate measures of diversity at particular genes, or genome regions, across a large number of individuals [71,104,105].

### 3.2.3. Whole-Genome Resequencing

Whole-genome resequencing (WGR) involves sequencing the genome of several individuals to a predetermined level of coverage (usually between 2× and 60×) and aligning this data to an available reference genome (for examples in non-model species, see Fuentes-Pardo and Ruzzante [15]). A major application of whole-genome resequencing (WGR) is the identification of variation throughout the genome, enabling the development of more targeted approaches that can be used to explore diversity at key regions in a larger cohort of individuals [70,71]. The Tasmanian devil targeted SNP panels were created using low-coverage WGR (10–15×) data from 7–12 individuals aligned against the annotated reference genome [70,78]. A major limitation of using this low-coverage resequencing strategy is that genome regions with lower coverage can often contain sequencing errors that may not be distinguished from true SNPs [106]. This led to a number of the SNPs identified in the Tasmanian devil resequencing

data not being present in the downstream SNP panel data [70,78]. While the best way to overcome this limitation is to increase the sequencing coverage of individuals, other methods, such as calling SNPs across individuals, can assist in more accurate variant calling in low-coverage WGR datasets [107].

Higher-coverage sequence data enables variants and heterozygosity to be called much more accurately than low-coverage sequence data and hence allows for SNPs to be called more confidently without additional targeted sequencing (e.g., SNP panels) [108]. High-coverage (~45×) WGR of 25 Tasmanian devils has allowed for reliable estimates of genome-wide heterozygosity, which are being used to assess the accuracy of estimates from other techniques including microsatellites, SNP panels and RRS data. The higher cost of high-coverage data causes a trade-off between investigating the whole genome of a relatively small number of individuals versus using a targeted subset of loci across many individuals (as of 2019, WGR routinely costs over \$1000 per individual whereas RRS costs less than \$100 per individual). This trade-off needs to be acknowledged, is dependent on the conservation research questions, and requires careful consideration prior to the commencement of sequencing [13]. Fortunately, a number of alternative cost-effective WGR approaches are available and may be suitable when high-coverage WGR is not possible. For a review of the different types of WGR and their different applications in conservation [15].

Whilst targeted sequencing approaches are useful for the exploration of genes known to be important to species biology, sometimes genetic mechanisms driving particular phenomena that are vital to species adaptation and survival may not be known or detected in other reduced sequencing techniques like RRS [109]. Whole-genome resequencing (WGR) enables conservation researchers to ask and answer a wide range of questions that are not possible using other approaches. For example, WGR also enables the use of genome-wide association studies to determine the genetic basis of particular phenotypic traits that are important to species conservation [13,15]. In the case of the Tasmanian devil, some individuals have been found to display a resistant phenotype to DFTD, enabling spontaneous tumor regression [110]. Identifying the potential genetic basis of this phenotype is important to understanding which individuals may be more resilient to the disease and provide targets for the development of potential treatments [76–78] (Table 1). Low-coverage WGR of individuals showing tumor regression and those that succumbed to the disease enabled a genome-wide association study to be undertaken, which identified two genomic regions that may be associated with resistance to DFTD including PAX3 and TLL1 loci [78]. A follow up study, Wright et al. [78] resequenced 10 individuals to a higher coverage (20–30×) and was able to identify a larger number of genomic regions that may underlie tumor regression in the Tasmanian devil [100]. This work demonstrates the ability of WGR data, along with an annotated reference genome, in exploring the genetic basis of phenotypic traits that could have important conservation implications [13,15,78,100] (Table 1). It is important to note that often larger numbers of individuals are required to identify genes underlying certain phenotypes, particularly in species with higher genetic diversity and reduced selective pressure on the phenotype of interest [111]. This requires careful consideration of trade-offs between the sequencing approach (targeted vs. RRS vs. WGR), number of samples and sequencing coverage, and will often depend upon some prior knowledge (or preliminary testing), budget, and access to samples. Overall, WGR data is better able to separate out and compare functional versus non-functional diversity than RRS methods, which is valuable in understanding the adaptive potential of species [94].

There are many other advantages of using this high-resolution genomic data, including i) more robust insights into the evolutionary and demographic histories of a species; ii) more accurate measures of diversity, inbreeding and population structure; and iii) the ability to identify and investigate signatures of selection and adaptive genetic variation [15,16,18]. WGR data in the Tasmanian devil is currently being employed to assess selection and mutation rates within populations and in identifying runs of homozygosity (ROH) throughout the genome (for examples in other species, see Ceballos, et al. [112] and Hodgkinson, et al. [113]). These analyses are useful in the investigation of well-known issues in conservation, including inbreeding depression [112] and adaptation to captivity [114].



Some of the current limitations for using WGR in conservation contexts are the cost, the required computing power and respective expertise, and the availability of reference genomes [13,15]. Costs vary greatly and depend on the number of individuals or loci you wish to use, and the required depth of sequencing [15]. In addition, this approach requires significant expertise and compute power to execute, which limits its applicability to many conservation contexts [15]. Creating partnerships between academic researchers with the required expertise and computing resources and conservation managers is key to overcoming many limitations of using genomics in conservation, and has been successfully implemented in the conservation of the Tasmanian devil [35]. A reference genome is essential for WGR, so the significant lack of published genomes (<1%) for threatened species (or their closely-related counterparts) prevents many conservation managers from taking full advantage of high-resolution genomic data. However, in the dawn of large genomic consortia such as the Earth Biogenome Project, which aims to sequence the genomes of all of the Earth's eukaryotic biodiversity over the next 10 years [20], lack of a reference genome will soon become a thing of the past.

Overall, WGR paired with an annotated reference genome opens up a realm of possibilities for downstream conservation research by developing more cost-effective approaches when data from a large number of individuals is necessary for making informed conservation management decisions. As costs of sequencing continue to decrease, and the availability of reference genomes continue to rise, the use of this high-resolution genomic data in conservation research will likely become the norm [12] and is already being applied to some bird species [95].

#### 4. Reference Genome Quality

An important factor to consider in the creation of reference genomes is the quality of the assembly. Consortia such as the Vertebrate Genome Project and the Earth Biogenome Project have proposed specific standards that reference genomes should meet [20,26] (Table S1). However, it is important to understand whether such high standards are necessary or achievable for conservation management. A number of statistics are used to evaluate the different aspects of genome quality including accuracy (e.g., average read coverage and quality), continuity (e.g., N50, N90, number of contigs/scaffolds, average length of contigs/scaffolds, gap percentage, etc.), and completeness (e.g., BUSCO (Benchmarking sets of Universal Single-Copy Orthologs)/CEGMA (Core Eukaryotic Genes Mapping Approach) scores, number of genes, etc.) (see Wajid and Serpedin [115] for a more exhaustive list). While the ideal reference genome would consist of a completely annotated, gap-free, chromosome-length assembly, even the some of the best model species genomes, such as the human genome, currently do not reach this standard. Furthermore, the ease and ability to reach chosen standards depends on many factors, including genome size, genome structure (e.g., repetitive content), level of heterozygosity, sample availability/quantity, as well as the cost and expertise of the sequencing types and computing resources available [24] (for reviews on reference genome creation including available sequencing types and their associated advantages/disadvantages see Ekblom and Wolf [96], Wajid and Serpedin [115], and Sedlazeck, et al. [116]). It is important to note that the current Tasmanian devil reference genome was sequenced in 2011 by Murchison et al. [46], so it does not meet the minimum standards set by the EBP (Earth Biogenome Project) or VGP (Vertebrate Genomes Project) (Table S1). Despite this, the Tasmanian devil genome has still been able to facilitate an enormous amount of conservation research. A higher-quality genome which is more complete, correct, and contiguous, has a number of advantages such as improved identification and characterization of genes and other genomic regions; more accurate ROH (runs of homozygosity) analysis and structural variant analysis; and higher resolution of chromosomal organization allowing for improved comparative genomic and evolutionary analyses [117].

Naturally, genome quality is also a factor of input DNA quality. High molecular weight DNA, generally greater than 40 kb in length, is required to generate the multiple sequencing types used to construct a high-quality genome [118]. Extracting high molecular weight DNA often requires additional consideration during the sample collection phase, such as flash-freezing tissues in liquid nitrogen,

storage at  $-80^{\circ}\text{C}$  or below, and avoiding freeze-thaw. However, for species of high conservation concern, or those that inhabit difficult field locations, this could be challenging. In these scenarios, researchers may utilize museum specimens. However, this can introduce additional problems associated with sample preservation and degraded DNA, which may not be suited to long-read sequencing technologies [119]. As such, the ability to collect, store, and extract high-quality DNA should not be underestimated, as this is an essential first step towards generating high-quality genome. However, it is important to weigh up whether the cost, computing resources, expertise, and time of creating an improved or “Gold standard” assembly is necessary to answer the conservation research questions at hand. For example, Patton et al. [47] showed that the improvement of contiguity of the newly released 2019 Tasmanian devil assembly had minimal impacts on inferred patterns of historical effective population size when compared to the current reference assembly. Hence, in many cases, a simple short-read genome assembly is enough to answer many basic conservation management questions and also enable a number of more in-depth species-specific analyses mentioned in the sections above. Nevertheless, as sequencing technologies and computational infrastructure continue to advance and become more affordable, high-quality reference genomes would become easier to create and would overcome many of the limitations of currently fragmented reference assemblies such as incomplete gene characterization, comparative evolutionary limitations, and increased computational requirements [117]. Despite this, without advances in sequencing chemistry and library preparation to reduce input DNA quality and quantity, the availability of high-quality samples and ensuing high molecular weight DNA may continue to limit the creation of high-quality reference genomes in some species.

## 5. Conclusions

The Tasmanian devil reference genome has enhanced our capacity to manage this species in the face of an infectious, clonal cancer. By having the reference genome, we have been able to develop a range of genomic tools that have been used to investigate DFTD (e.g., [46]), investigate the interplay between the Tasmanian devils and the disease (e.g., [76–79]), inform development of immunotherapy and vaccine protocols [66], inform the management of the insurance population [38,65], and provide advice on the translocation of Tasmanian devils to wild populations to improve both genome-wide and functional diversity (e.g., [63,89]). Tasmanian devils are not the only species who are threatened globally by disease; other examples include black-footed ferret and distemper [120], bats and white-nose syndrome [121], and frogs and chytrid [122]. Here we have presented a strong case study of the benefits of using reference genomes for the conservation of threatened species. As the threat to global biodiversity increases, the management of threatened species becomes more pronounced. Reference genomes could be used by conservation managers to develop a range of genetic tools such as designing species-specific microsatellite markers for population data and differentiation; developing targeted SNP panels, or aligning and calling RRS data, for higher resolution population information or data on particular genes of interest; and conducting exploratory analyses (e.g., genome-wide association studies) using variant calling of whole-genome resequencing data.

Despite the challenges in obtaining high-quality samples for genome sequencing and expertise for the creation of reference genomes for threatened species, there is value in them. Reduced costs and lower input DNA requirements, as well as improved bioinformatic assembly and annotation pipelines based on non-model non-eutherian species, mean that these technologies are becoming more attainable by conservation programs and should be used more routinely where budgets allow [96]. Reference genomes enable a wealth of genetic/genomic applications and are an important asset in our ongoing fight to preserve global biodiversity. We would recommend that conservation managers who are seeking to use the types of methods we have described herein collaborate with global genome consortia (like the Earth Biogenome Project) or national/local consortia (like the Oz Mammal Genome Initiative) to utilize the full potential of genomic resources and join the genomics revolution. This

allows conservation managers to focus on conservation and work with geneticists who can help them make adaptive management decisions in real-time [35].

Although here we have presented a unique case study of a species with significantly low levels of genetic diversity and a large threatening disease process, the techniques described for the Tasmanian devil can be applied more broadly to many species of conservation concern. The applications of what we have described herein for devils is not unique to this species as many of the questions we have answered are posed by those managing other threatened species. **These include understanding historical demography and current population structure, minimizing inbreeding, maximizing adaptive potential, and identifying the basis of important phenotypic traits** (whether these be related to disease, behavior, or reproduction). Hence, despite differences in threatening processes and current state of vulnerable species, the nature of their small population sizes will result in a number of common conservation concerns that could be informed using genomic data [15,18]. In the midst of the sixth mass extinction event, we advocate the use of reference genomes and associated genetic tools to arm conservation managers with ways to assist the long-term survival of species.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2073-4425/10/11/846/s1>, Table S1: Comparison of model and non-model mammalian/marsupial reference genomes to the G10K and EBP minimum reference genome quality standards.

**Author Contributions:** This project was conceived by K.B. and C.J.H. All authors contributed to writing and reviewing the manuscript. This work was funded by the Australian Research Council Grants to K.B. and C.J.H.

**Funding:** This research was funded by the Australian Research Council (LP140100508; DP180102465).

**Acknowledgments:** Thanks to all the staff, students, and affiliates of the Australasian Wildlife Genomics Group who have contributed to discussions over the years in regards to the benefit of genomics data to conservation, in particular, R. Johnson, C. Grueber, B. Wright and Y. Cheng. Thanks also to the staff of the Save the Tasmanian Devil Program and the members of the Zoo and Aquarium Association Australasia, without whom our devil work would not be possible.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ceballos, G.; Ehrlich, P.R.; Barnosky, A.D.; García, A.; Pringle, R.M.; Palmer, T.M. Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Sci. Adv.* **2015**, *1*, e1400253. [[CrossRef](#)] [[PubMed](#)]
2. IUCN. The IUCN Red List of Threatened Species. Version 2019-1. Available online: <http://www.iucnredlist.org> (accessed on 2 July 2019).
3. Johnson, C.N.; Isaac, J.L. Body mass and extinction risk in Australian marsupials: The ‘Critical Weight Range’ revisited. *Austral. Ecol.* **2009**, *34*, 35–40. [[CrossRef](#)]
4. Johnson, C.N.; Isaac, J.L.; Fisher, D.O. Rarity of a top predator triggers continent-wide collapse of mammal prey: Dingoes and marsupials in Australia. *Proc. R. Soc. Biol. Sci. Ser. B* **2006**, *274*, 341–346. [[CrossRef](#)] [[PubMed](#)]
5. Short, J.; Smith, A. Mammal decline and recovery in Australia. *J. Mammal.* **1994**, *75*, 288–297. [[CrossRef](#)]
6. Mittermeier, R.A. *Megadiversity: Earth’s Biologically Wealthiest Nations*; Agrupacion Sierra Madre: Mexico City, Mexico, 1997.
7. Chapman, A.D. *Numbers of Living Species in Australia and the World*, 2nd ed.; Australian Government, Department of the Environment and Energy: Canberra, Australia, 2009.
8. Department of the Environment and Energy. Recovery Plans. Available online: <https://www.environment.gov.au/biodiversity/threatened/recovery-plans> (accessed on 8 August 2019).
9. Ballou, J.D.; Lees, C.; Faust, L.J.; Long, S.; Lynch, C.; Bingaman Lackey, L.; Foote, T.J. Demographic and genetic management of captive populations. In *Wild Mammals in Captivity: Principles and Techniques for Zoo Management*, 2nd ed.; The University of Chicago Press: Chicago, IL, USA, 2010; pp. 219–252.
10. Lacy, R.C. Importance of Genetic Variation to the Viability of Mammalian Populations. *J. Mammal.* **1997**, *78*, 320–335. [[CrossRef](#)]

11. Frankham, R.; Ballou, J.D.; Briscoe, D.A. *Introduction to Conservation Genetics*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2010. [CrossRef]
12. Johnson, W.E.; Koepfli, K. The role of genomics in conservation and reproductive sciences. In *Reproductive Sciences in Animal Conservation*; Springer: Berlin, Germany, 2014; pp. 71–96.
13. Supple, M.A.; Shapiro, B. Conservation of biodiversity in the genomics era. *Genome Biol.* **2018**, *19*, 131. [CrossRef]
14. Allendorf, F.W. Genetics and the conservation of natural populations: Allozymes to genomes. *Mol. Ecol.* **2017**, *26*, 420–430. [CrossRef]
15. Fuentes-Pardo, A.P.; Ruzzante, D.E. Whole-genome sequencing approaches for conservation biology: Advantages, limitations and practical recommendations. *Mol. Ecol.* **2017**, *26*, 5369–5406. [CrossRef]
16. Larsen, P.A.; Matocq, M.D. Emerging genomic applications in mammalian ecology, evolution, and conservation. *J. Mammal.* **2019**, *100*, 786–801. [CrossRef]
17. McMahon, B.J.; Teeling, E.C.; Höglund, J. How and why should we implement genomics into conservation? *Evol. Appl.* **2014**, *7*, 999–1007. [CrossRef]
18. Khan, S.; Nabi, G.; Ullah, M.W.; Yousaf, M.; Manan, S.; Siddique, R.; Hou, H. Overview on the role of advance genomics in conservation biology of endangered species. *Int. J. Genomics* **2016**, *2016*, 1–8. [CrossRef] [PubMed]
19. Kitts, P.A.; Church, D.M.; Thibaud-Nissen, F.; Choi, J.; Hem, V.; Sapojnikov, V.; Smith, R.G.; Tatusova, T.; Xiang, C.; Zherikov, A. Assembly: A resource for assembled genomes at NCBI. *Nucleic Acids Res.* **2015**, *44*, D73–D80. [CrossRef] [PubMed]
20. Lewin, H.A.; Robinson, G.E.; Kress, W.J.; Baker, W.J.; Coddington, J.; Crandall, K.A.; Durbin, R.; Edwards, S.V.; Forest, F.; Gilbert, M.T.P. Earth BioGenome Project: Sequencing life for the future of life. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 4325–4333. [CrossRef] [PubMed]
21. Li, Q.; Xuan, Z.; Li, Y.; Zheng, H.; Bai, Y.; Li, B.; Hu, Y.; Liu, X.; Zhang, Z.; Li, D.; et al. The sequence and de novo assembly of the giant panda genome. *Nature* **2010**, *463*, 311–317. [CrossRef]
22. Groenen, M.A.; Archibald, A.L.; Uenishi, H.; Tuggle, C.K.; Takeuchi, Y.; Rothschild, M.F.; Rogel-Gaillard, C.; Park, C.; Milan, D.; Megens, H.-J. Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* **2012**, *491*, 393–398. [CrossRef]
23. Johnson, R.N.; O'Meally, D.; Chen, Z.; Etherington, G.J.; Ho, S.Y.W.; Nash, W.J.; Grueber, C.E.; Cheng, Y.; Whittington, C.M.; Dennison, S.; et al. Adaptation and conservation insights from the koala genome. *Nat. Genet.* **2018**, *50*, 1102–1111. [CrossRef]
24. Koepfli, K.-P.; Paten, B. Genome 10K Community of Scientists; O'Brien, S.J. The Genome 10K Project: A way forward. *Annu. Rev. Anim. Biosci.* **2015**, *3*, 57–111. [CrossRef]
25. Genome 10K Community of Scientists. Genome 10K: A proposal to obtain whole-genome sequence for 10 000 vertebrate species. *J. Hered.* **2009**, *100*, 659–674. [CrossRef]
26. Genome 10K Community of Scientists. Vertebrate Genomes Project. Available online: <https://vertebrategenomesproject.org> (accessed on 16 August 2019).
27. China National GeneBank. B10K. Available online: <https://b10k.genomics.cn/> (accessed on 16 August 2019).
28. Teeling, E.C.; Vernes, S.C.; Dávalos, L.M.; Ray, D.A.; Gilbert, M.T.P.; Myers, E. Bat1K Consortium. Bat biology, genomes, and the Bat1K project: To generate chromosome-level genomes for all living bat species. *Annu. Rev. Anim. Biosci.* **2018**, *6*, 23–46. [CrossRef]
29. GIGA Community of Scientists. The Global Invertebrate Genomics Alliance (GIGA): Developing community resources to study diverse invertebrate genomes. *J. Hered.* **2013**, *105*, 1–18.
30. Voolstra, C.R.; Wörheide, G.; Lopez, J.V. Corrigendum to: Advancing genomics through the Global Invertebrate Genomics Alliance (GIGA). *Invertebr. Syst.* **2017**, *31*, 231. [CrossRef]
31. Potter, S.; Eldridge, M. Oz Mammal Genomics. *Australas. Sci.* **2017**, *38*, 19–21.
32. Ralls, K.; Ballou, J.D.; Dudash, M.R.; Eldridge, M.D.; Fenster, C.B.; Lacy, R.C.; Sunnucks, P.; Frankham, R. Call for a paradigm shift in the genetic management of fragmented populations. *Conserv. Lett.* **2018**, *11*, e12412. [CrossRef]
33. Taylor, H.R.; Dussex, N.; van Heezik, Y. Bridging the conservation genetics gap by identifying barriers to implementation for conservation practitioners. *Glob. Ecol. Conserv.* **2017**, *10*, 231–242. [CrossRef]



34. Holderegger, R.; Balkenhol, N.; Bolliger, J.; Engler, J.O.; Gugerli, F.; Hochkirch, A.; Nowak, C.; Segelbacher, G.; Widmer, A.; Zachos, F.E. Conservation genetics: Linking science with practice. *Mol. Ecol.* **2019**, *28*, 3848–3856. [[CrossRef](#)]
35. Hogg, C.J.; Grueber, C.E.; Pemberton, D.; Fox, S.; Lee, A.V.; Ivy, J.A.; Belov, K. “Devil Tools & Tech”: A Synergy of Conservation Research and Management Practice. *Conserv. Lett.* **2017**, *10*, 133–138.
36. Grueber, C.E. Comparative genomics for biodiversity conservation. *Comput. Struct. Biotechnol. J.* **2015**, *13*, 370–375. [[CrossRef](#)]
37. Lazenby, B.T.; Tobler, M.W.; Brown, W.E.; Hawkins, C.E.; Hocking, G.J.; Hume, F.; Huxtable, S.; Iles, P.; Jones, M.E.; Lawrence, C. Density trends and demographic signals uncover the long-term impact of transmissible cancer in Tasmanian devils. *J. Appl. Ecol.* **2018**, *55*, 1368–1379. [[CrossRef](#)]
38. Hogg, C.J.; Fox, S.; Pemberton, D.; Belov, K. *Saving the Tasmanian Devil*; Hogg, C.J., Fox, S., Pemberton, D., Belov, K., Eds.; CSIRO Publishing: Clayton South, VIC, Australia, 2019.
39. Jones, M.E.; Paetkau, D.; Geffen, E.; Moritz, C. Microsatellites for the Tasmanian devil (*Sarcophilus harrisii*). *Mol. Ecol. Notes* **2003**, *3*, 277–279. [[CrossRef](#)]
40. Siddle, H.V.; Marzec, J.; Cheng, Y.; Jones, M.; Belov, K. MHC gene copy number variation in Tasmanian devils: Implications for the spread of a contagious cancer. *Proc. R. Soc. Biol. Sci. Ser. B* **2010**, *277*, 2001–2006. [[CrossRef](#)]
41. Cheng, Y.; Belov, K. Isolation and characterisation of 11 MHC-linked microsatellite loci in the Tasmanian devil (*Sarcophilus harrisii*). *Conserv. Genet. Resour.* **2012**, *4*, 463–465. [[CrossRef](#)]
42. Andrews, K.R.; Good, J.M.; Miller, M.R.; Luikart, G.; Hohenlohe, P.A. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet.* **2016**, *17*, 81. [[CrossRef](#)] [[PubMed](#)]
43. Hogg, C.J.; Ivy, J.A.; Srb, C.; Hockley, J.; Lees, C.; Hibbard, C.; Jones, M. Influence of genetic provenance and birth origin on productivity of the Tasmanian devil insurance population. *Conserv. Genet.* **2015**, *16*, 1465–1473. [[CrossRef](#)]
44. Hogg, C.J.; Wright, B.; Morris, K.M.; Lee, A.V.; Ivy, J.A.; Grueber, C.E.; Belov, K. Founder relationships and conservation management: Empirical kinships reveal the effect on breeding programmes when founders are assumed to be unrelated. *Anim. Conserv.* **2019**, *22*, 348–361. [[CrossRef](#)]
45. Miller, W.; Hayes, V.M.; Ratan, A.; Petersen, D.C.; Wittekindt, N.E.; Miller, J.; Walenz, B.; Knight, J.; Qi, J.; Zhao, F.; et al. Genetic diversity and population structure of the endangered marsupial *Sarcophilus harrisii* (Tasmanian devil). *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 12348–12353. [[CrossRef](#)] [[PubMed](#)]
46. Murchison, E.P.; Schulz-Trieglaff, O.B.; Ning, Z.; Alexandrov, L.B.; Bauer, M.J.; Fu, B.; Hims, M.; Ding, Z.; Ivakhno, S.; Stewart, C. Genome sequencing and analysis of the Tasmanian devil and its transmissible cancer. *Cell* **2012**, *148*, 780–791. [[CrossRef](#)]
47. Patton, A.H.; Margres, M.J.; Stahlke, A.R.; Hendricks, S.; Lewallen, K.; Hamede, R.K.; Ruiz-Aravena, M.; Ryder, O.; McCallum, H.I.; Jones, M.E.; et al. Contemporary demographic reconstruction methods are robust to genome assembly quality: A case study in Tasmanian Devils. *Mol. Biol. Evol.* **2019**, msz191. [[CrossRef](#)]
48. Selkoe, K.A.; Toonen, R.J. Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecol. Lett.* **2006**, *9*, 615–629. [[CrossRef](#)]
49. Abdul-Muneer, P.M. Application of microsatellite markers in conservation genetics and fisheries management: Recent advances in population structure analysis and conservation strategies. *Genet. Res. Int.* **2014**, *2014*, 1–11. [[CrossRef](#)]
50. Gooley, R.; Hogg, C.J.; Belov, K.; Grueber, C.E. No evidence of inbreeding depression in a Tasmanian devil insurance population despite significant variation in inbreeding. *Sci. Rep.* **2017**, *7*, 1830. [[CrossRef](#)]
51. Abdelkrim, J.; Robertson, B.C.; Stanton, J.-A.L.; Gemmell, N.J. Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *BioTechniques* **2009**, *46*, 185–192. [[CrossRef](#)] [[PubMed](#)]
52. Gooley, R.M.; Hogg, C.J.; Belov, K.; Grueber, C.E. The effects of group versus intensive housing on the retention of genetic diversity in insurance populations. *BMC Zool.* **2018**, *3*, 2. [[CrossRef](#)]
53. McLennan, E.A.; Gooley, R.M.; Wise, P.; Belov, K.; Hogg, C.J.; Grueber, C.E. Pedigree reconstruction using molecular data reveals an early warning sign of gene diversity loss in an island population of Tasmanian devils (*Sarcophilus harrisii*). *Conserv. Genet.* **2018**, *19*, 439–450. [[CrossRef](#)]
54. Farquharson, K.A.; Hogg, C.J.; Grueber, C.E. A case for genetic parentage assignment in captive group housing. *Conserv. Genet.* **2019**, *20*, 1–7. [[CrossRef](#)]



55. Day, J.; Gooley, R.M.; Hogg, C.J.; Belov, K.; Whittington, C.M.; Grueber, C.E. MHC-associated mate choice under competitive conditions in captive versus wild Tasmanian devils. *Behav. Ecol.* **2019**, *30*, 1196–1204. [\[CrossRef\]](#)
56. Grueber, C.E.; Chong, R.; Gooley, R.M.; McLennan, E.A.; Barrs, V.R.; Belov, K.; Hogg, C.J. Genetic analysis of scat samples to inform conservation of Tasmanian devil. *Aust. Zool.* (In press).
57. Taberlet, P.; Waits, L.P.; Luikart, G. Noninvasive genetic sampling: Look before you leap. *Trends Ecol. Evol.* **1999**, *14*, 323–327. [\[CrossRef\]](#)
58. Armstrong, A.J.; Dudgeon, C.L.; Bustamante, C.; Bennett, M.B.; Ovenden, J.R. Development and characterization of 17 polymorphic microsatellite markers for the reef manta ray (*Mobula alfredi*). *BMC Res. Notes* **2019**, *12*, 233. [\[CrossRef\]](#)
59. Faria, J.; Pita, A.; Rivas, M.; Martins, G.M.; Hawkins, S.J.; Ribeiro, P.; Neto, A.I.; Presa, P. A multiplex microsatellite tool for conservation genetics of the endemic limpet *Patella candei* in the Macaronesian archipelagos. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **2016**, *26*, 775–781. [\[CrossRef\]](#)
60. Shaney, K.J.; Adams, R.; Kurniawan, N.; Hamidy, A.; Smith, E.N.; Castoe, T.A. A suite of potentially amplifiable microsatellite loci for ten reptiles of conservation concern from Africa and Asia. *Conserv. Genet. Resour.* **2016**, *8*, 307–311. [\[CrossRef\]](#)
61. Storfer, A.; Epstein, B.; Jones, M.; Micheletti, S.; Spear, S.F.; Lachish, S.; Fox, S. Landscape genetics of the Tasmanian devil: Implications for spread of an infectious cancer. *Conserv. Genet.* **2017**, *18*, 1287–1297. [\[CrossRef\]](#)
62. Grueber, C.E.; Fox, S.; McLennan, E.A.; Gooley, R.M.; Weiser, E.L.; Pemberton, D.; Hogg, C.J.; Belov, K. Complex problems need detailed solutions: Harnessing multiple data types to inform genetic rescue in the wild. *Evol. Appl.* **2019**, *12*, 280–291. [\[CrossRef\]](#) [\[PubMed\]](#)
63. Hogg, C.J.; McLennan, E.A.; Wise, P.; Lee, A.; Pemberton, D.; Fox, S.; Belov, K.; Grueber, C.E. Preserving the integrity of a single source population during multiple translocations. *Biol. Conserv.* (In press).
64. Pye, R.J.; Pemberton, D.; Tovar, C.; Tubio, J.M.; Dun, K.A.; Fox, S.; Darby, J.; Hayes, D.; Knowles, G.W.; Kreiss, A. A second transmissible cancer in Tasmanian devils. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 374–379. [\[CrossRef\]](#) [\[PubMed\]](#)
65. Hogg, C.; Lee, A.; Srb, C.; Hibbard, C. Metapopulation management of an endangered species with limited genetic diversity in the presence of disease: The Tasmanian devil *Sarcophilus harrisii*. *Int. Zoo Yearb.* **2017**, *51*, 137–153. [\[CrossRef\]](#)
66. Pye, R.; Patchett, A.; McLennan, E.; Thomson, R.; Carver, S.; Fox, S.; Pemberton, D.; Kreiss, A.; Baz Morelli, A.; Silva, A. Immunization strategies producing a humoral IgG immune response against devil facial tumor disease in the majority of Tasmanian devils destined for wild release. *Front. Immunol.* **2018**, *9*, 259. [\[CrossRef\]](#)
67. Grueber, C.; Peel, E.; Wright, B.; Hogg, C.; Belov, K. A Tasmanian devil breeding program to support wild recovery. *Reprod. Fertil. Dev.* **2019**, *31*, 1296–1304. [\[CrossRef\]](#)
68. McLennan, E.A.; Grueber, C.E.; Wise, P.; Belov, K.; Hogg, C.J. Mixing genetic lineages successfully boosts diversity of an endangered carnivore. *Anim. Conserv.* (under review).
69. Morris, K.M.; Cheng, Y.; Warren, W.; Papenfuss, A.T.; Belov, K. Identification and analysis of divergent immune gene families within the Tasmanian devil genome. *BMC Genomics* **2015**, *16*, 1017. [\[CrossRef\]](#)
70. Morris, K.M.; Wright, B.; Grueber, C.E.; Hogg, C.; Belov, K. Lack of genetic diversity across diverse immune genes in an endangered mammal, the Tasmanian devil (*Sarcophilus harrisii*). *Mol. Ecol.* **2015**, *24*, 3860–3872. [\[CrossRef\]](#)
71. Wright, B.; Morris, K.; Grueber, C.E.; Willet, C.E.; Gooley, R.; Hogg, C.J.; O'Meally, D.; Hamede, R.; Jones, M.; Wade, C. Development of a SNP-based assay for measuring genetic diversity in the Tasmanian devil insurance population. *BMC Genomics* **2015**, *16*, 791. [\[CrossRef\]](#) [\[PubMed\]](#)
72. Cheng, Y.; Belov, K. Characterisation of non-classical MHC class I genes in the Tasmanian devil (*Sarcophilus harrisii*). *Immunogenetics* **2014**, *66*, 727–735. [\[CrossRef\]](#) [\[PubMed\]](#)
73. Cheng, Y.; Heasman, K.; Peck, S.; Peel, E.; Gooley, R.M.; Papenfuss, A.T.; Hogg, C.J.; Belov, K. Significant decline in anticancer immune capacity during puberty in the Tasmanian devil. *Sci. Rep.* **2017**, *7*, 44716. [\[CrossRef\]](#) [\[PubMed\]](#)
74. Cheng, Y.; Makara, M.; Peel, E.; Fox, S.; Papenfuss, A.T.; Belov, K. Tasmanian devils with contagious cancer exhibit a constricted T-cell repertoire diversity. *Commun. Biol.* **2019**, *2*, 99. [\[CrossRef\]](#) [\[PubMed\]](#)

75. Cui, J.; Cheng, Y.; Belov, K. Diversity in the Toll-like receptor genes of the Tasmanian devil (*Sarcophilus harrisii*). *Immunogenetics* **2015**, *67*, 195–201. [[CrossRef](#)] [[PubMed](#)]
76. Epstein, B.; Jones, M.; Hamede, R.; Hendricks, S.; McCallum, H.; Murchison, E.P.; Schönfeld, B.; Wiench, C.; Hohenlohe, P.; Storfer, A. Rapid evolutionary response to a transmissible cancer in Tasmanian devils. *Nat. Commun.* **2016**, *7*, 12684. [[CrossRef](#)]
77. Margres, M.J.; Jones, M.E.; Epstein, B.; Kerlin, D.H.; Comte, S.; Fox, S.; Fraik, A.K.; Hendricks, S.A.; Huxtable, S.; Lachish, S. Large-effect loci affect survival in Tasmanian devils (*Sarcophilus harrisii*) infected with a transmissible cancer. *Mol. Ecol.* **2018**, *27*, 4189–4199. [[CrossRef](#)]
78. Wright, B.; Willet, C.E.; Hamede, R.; Jones, M.; Belov, K.; Wade, C.M. Variants in the host genome may inhibit tumour growth in devil facial tumours: Evidence from genome-wide association. *Sci. Rep.* **2017**, *7*, 423. [[CrossRef](#)]
79. Hohenlohe, P.A.; McCallum, H.I.; Jones, M.E.; Lawrance, M.F.; Hamede, R.K.; Storfer, A. Conserving adaptive potential: Lessons from Tasmanian devils and their transmissible cancer. *Conserv. Genet.* **2019**, *20*, 81–87. [[CrossRef](#)]
80. Fernández, M.E.; Goszczynski, D.E.; Lirón, J.P.; Villegas-Castagnasso, E.E.; Carino, M.H.; Ripoli, M.V.; Rogberg-Muñoz, A.; Posik, D.M.; Peral-García, P.; Giovambattista, G. Comparison of the effectiveness of microsatellites and SNP panels for genetic identification, traceability and assessment of parentage in an inbred Angus herd. *Genet. Mol. Biol.* **2013**, *36*, 185–191. [[CrossRef](#)]
81. Tokarska, M.; Marshall, T.; Kowalczyk, R.; Wójcik, J.; Pertoldi, C.; Kristensen, T.; Loeschcke, V.; Gregersen, V.; Bendixen, C. Effectiveness of microsatellite and SNP markers for parentage and identity analysis in species with low genetic diversity: The case of European bison. *Heredity* **2009**, *103*, 326. [[CrossRef](#)] [[PubMed](#)]
82. Rosenberg, N.A.; Li, L.M.; Ward, R.; Pritchard, J.K. Informativeness of genetic markers for inference of ancestry. *Am. J. Hum. Genet.* **2003**, *73*, 1402–1422. [[CrossRef](#)] [[PubMed](#)]
83. Schopen, G.; Bovenhuis, H.; Visker, M.; Van Arendonk, J. Comparison of information content for microsatellites and SNPs in poultry and cattle. *Anim. Genet.* **2008**, *39*, 451–453. [[CrossRef](#)] [[PubMed](#)]
84. Wright, B.; Farquharson, K.A.; McLennan, E.A.; Belov, K.; Hogg, C.J.; Grueber, C.E. From reference genomes to population genomics: Comparing three reference-aligned reduced-representation sequencing pipelines in two wildlife species. *BMC Genomics* **2019**, *20*, 453. [[CrossRef](#)] [[PubMed](#)]
85. Davey, J.W.; Blaxter, M.L. RADSeq: Next-generation population genetics. *Briefings Funct. Genomics* **2010**, *9*, 416–423. [[CrossRef](#)]
86. Peterson, B.K.; Weber, J.N.; Kay, E.H.; Fisher, H.S.; Hoekstra, H.E. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE* **2012**, *7*, e37135. [[CrossRef](#)]
87. Von Mark, V.C.; Kilian, A.; Dierig, D.A. Development of DArT marker platforms and genetic diversity assessment of the US collection of the new oilseed crop lesquerella and related species. *PLoS ONE* **2013**, *8*, e64062.
88. Hendricks, S.; Epstein, B.; Schönfeld, B.; Wiench, C.; Hamede, R.; Jones, M.; Storfer, A.; Hohenlohe, P. Conservation implications of limited genetic diversity and population structure in Tasmanian devils (*Sarcophilus harrisii*). *Conserv. Genet.* **2017**, *18*, 977–982. [[CrossRef](#)]
89. McLennan, E.A.; Wright, B.R.; Belov, K.; Hogg, C.J.; Grueber, C.E. Too much of a good thing? Finding the most informative genetic data set to answer conservation questions. *Mol. Ecol. Resour.* **2019**, *19*, 659–671. [[CrossRef](#)]
90. Torkamaneh, D.; Laroche, J.; Belzile, F. Genome-wide SNP calling from genotyping by sequencing (GBS) data: A comparison of seven pipelines and two sequencing technologies. *PLoS ONE* **2016**, *11*, e0161333. [[CrossRef](#)]
91. Davey, J.W.; Hohenlohe, P.A.; Etter, P.D.; Boone, J.Q.; Catchen, J.M.; Blaxter, M.L. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* **2011**, *12*, 499. [[CrossRef](#)] [[PubMed](#)]
92. Shafer, A.B.; Peart, C.R.; Tusso, S.; Maayan, I.; Brelsford, A.; Wheat, C.W.; Wolf, J.B. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods Ecol. Evol.* **2017**, *8*, 907–917. [[CrossRef](#)]
93. Gurgul, A.; Miksza-Cybulska, A.; Szmatola, T.; Jasielczuk, I.; Piestrzyńska-Kajtoch, A.; Fornal, A.; Semik-Gurgul, E.; Bugno-Poniewierska, M. Genotyping-by-sequencing performance in selected livestock species. *Genomics* **2019**, *111*, 186–195. [[CrossRef](#)] [[PubMed](#)]

94. Hoelzel, A.R.; Bruford, M.W.; Fleischer, R.C. *Conservation of Adaptive Potential and Functional Diversity*; Springer: Berlin, Germany, 2019.
95. Galla, S.J.; Forsdick, N.J.; Brown, L.; Hoepfner, M.; Knapp, M.; Maloney, R.F.; Moraga, R.; Santure, A.W.; Steeves, T.E. Reference genomes from distantly related species can be used for discovery of single nucleotide polymorphisms to inform conservation management. *Genes* **2019**, *10*, 9. [[CrossRef](#)] [[PubMed](#)]
96. Ekblom, R.; Wolf, J.B. A field guide to whole-genome sequencing, assembly and annotation. *Evol. Appl.* **2014**, *7*, 1026–1042. [[CrossRef](#)] [[PubMed](#)]
97. Yandell, M.; Ence, D. A beginner's guide to eukaryotic genome annotation. *Nat. Rev. Genet.* **2012**, *13*, 329. [[CrossRef](#)] [[PubMed](#)]
98. Curwen, V.; Eyra, E.; Andrews, T.D.; Clarke, L.; Mongin, E.; Searle, S.M.; Clamp, M. The Ensembl automatic gene annotation system. *Genome Res.* **2004**, *14*, 942–950. [[CrossRef](#)]
99. Potter, S.C.; Clarke, L.; Curwen, V.; Keenan, S.; Mongin, E.; Searle, S.M.; Stabenau, A.; Storey, R.; Clamp, M. The Ensembl analysis pipeline. *Genome Res.* **2004**, *14*, 934–941. [[CrossRef](#)]
100. Margres, M.J.; Ruiz-Aravena, M.; Hamede, R.; Jones, M.E.; Lawrance, M.F.; Hendricks, S.A.; Patton, A.; Davis, B.W.; Ostrander, E.A.; McCallum, H. The genomic basis of tumor regression in Tasmanian devils (*Sarcophilus harrisii*). *Genome Biol. Evol.* **2018**, *10*, 3012–3025. [[CrossRef](#)]
101. Peel, E.; Cheng, Y.; Djordjevic, J.; Fox, S.; Sorrell, T.; Belov, K. Cathelicidins in the Tasmanian devil (*Sarcophilus harrisii*). *Sci. Rep.* **2016**, *6*, 35019. [[CrossRef](#)]
102. Van der Kraan, L.E.; Wong, E.S.; Lo, N.; Ujvari, B.; Belov, K. Identification of natural killer cell receptor genes in the genome of the marsupial Tasmanian devil (*Sarcophilus harrisii*). *Immunogenetics* **2013**, *65*, 25–35. [[CrossRef](#)]
103. Van Tienderen, P.H.; de Haan, A.A.; van der Linden, C.G.; Vosman, B. Biodiversity assessment using markers for ecologically important traits. *Trends Ecol. Evol.* **2002**, *17*, 577–582. [[CrossRef](#)]
104. Russell, T.; Cullingham, C.; Kommadath, A.; Stothard, P.; Herbst, A.; Coltman, D. Development of a novel mule deer genomic assembly and species-diagnostic SNP panel for assessing introgression in mule deer, white-tailed deer, and their interspecific hybrids. *G3 Genes Genomes Genet.* **2019**, *9*, 911–919. [[CrossRef](#)] [[PubMed](#)]
105. Zhao, H.; Fuller, A.; Thongda, W.; Mohammed, H.; Abernathy, J.; Beck, B.; Peatman, E. SNP panel development for genetic management of wild and domesticated white bass (*Morone chrysops*). *Anim. Genet.* **2019**, *50*, 92–96. [[CrossRef](#)] [[PubMed](#)]
106. Li, R.; Li, Y.; Fang, X.; Yang, H.; Wang, J.; Kristiansen, K.; Wang, J. SNP detection for massively parallel whole-genome resequencing. *Genome Res.* **2009**, *19*, 1124–1132. [[CrossRef](#)]
107. Cheng, A.Y.; Teo, Y.-Y.; Ong, R.T.-H. Assessing single nucleotide variant detection and genotype calling on whole-genome sequenced individuals. *Bioinformatics* **2014**, *30*, 1707–1713. [[CrossRef](#)]
108. Kishikawa, T.; Momozawa, Y.; Ozeki, T.; Mushiroda, T.; Inohara, H.; Kamatani, Y.; Kubo, M.; Okada, Y. Empirical evaluation of variant calling accuracy using ultra-deep whole-genome sequencing data. *Sci. Rep.* **2019**, *9*, 1784. [[CrossRef](#)]
109. Hoban, S.; Kelley, J.L.; Lotterhos, K.E.; Antolin, M.F.; Bradburd, G.; Lowry, D.B.; Poss, M.L.; Reed, L.K.; Storfer, A.; Whitlock, M.C. Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *Am. Nat.* **2016**, *188*, 379–397. [[CrossRef](#)]
110. Pye, R.; Hamede, R.; Siddle, H.V.; Caldwell, A.; Knowles, G.W.; Swift, K.; Kreiss, A.; Jones, M.E.; Lyons, A.B.; Woods, G.M. Demonstration of immune responses against devil facial tumour disease in wild Tasmanian devils. *Biol. Lett.* **2016**, *12*, 20160553. [[CrossRef](#)]
111. Hong, E.P.; Park, J.W. Sample size and statistical power calculation in genetic association studies. *Genomics Inform.* **2012**, *10*, 117. [[CrossRef](#)]
112. Ceballos, F.C.; Hazelhurst, S.; Ramsay, M. Assessing runs of homozygosity: A comparison of SNP array and whole genome sequence low coverage data. *BMC Genomics* **2018**, *19*, 106. [[CrossRef](#)] [[PubMed](#)]
113. Hodgkinson, A.; Casals, F.; Idaghdour, Y.; Grenier, J.-C.; Hernandez, R.D.; Awadalla, P. Selective constraint, background selection, and mutation accumulation variability within and between human populations. *BMC Genomics* **2013**, *14*, 495. [[CrossRef](#)] [[PubMed](#)]
114. Willoughby, J.R.; Ivy, J.A.; Lacy, R.C.; Doyle, J.M.; DeWoody, J.A. Inbreeding and selection shape genomic diversity in captive populations: Implications for the conservation of endangered species. *PLoS ONE* **2017**, *12*, e0175996. [[CrossRef](#)] [[PubMed](#)]

115. Wajid, B.; Serpedin, E. Do it yourself guide to genome assembly. *Brief. Funct. Genomics* **2014**, *15*, 1–9. [[CrossRef](#)] [[PubMed](#)]
116. Sedlazeck, F.J.; Lee, H.; Darby, C.A.; Schatz, M.C. Piercing the dark matter: Bioinformatics of long-range sequencing and mapping. *Nat. Rev. Genet.* **2018**, *19*, 329–346. [[CrossRef](#)] [[PubMed](#)]
117. Lee, H.; Gurtowski, J.; Yoo, S.; Nattestad, M.; Marcus, S.; Goodwin, S.; McCombie, W.R.; Schatz, M. Third-generation sequencing and the future of genomics. *BioRxiv* **2016**, 048603.
118. Rhoads, A.; Au, K.F. PacBio Sequencing and Its Applications. *Genomics Proteomics Bioinform.* **2015**, *13*, 278–289. [[CrossRef](#)]
119. McDonough, M.M.; Parker, L.D.; Rotzel McInerney, N.; Campana, M.G.; Maldonado, J.E. Performance of commonly requested destructive museum samples for mammalian genomic studies. *J. Mammal.* **2018**, *99*, 789–802. [[CrossRef](#)]
120. Thorne, E.T.; Williams, E.S. Disease and endangered species: The black-footed ferret as a recent example. *Conserv. Biol.* **1988**, *2*, 66–74. [[CrossRef](#)]
121. Blehert, D.S.; Hicks, A.C.; Behr, M.; Meteyer, C.U.; Berlowski-Zier, B.M.; Buckles, E.L.; Coleman, J.T.; Darling, S.R.; Gargas, A.; Niver, R. Bat white-nose syndrome: An emerging fungal pathogen? *Science* **2009**, *323*, 227. [[CrossRef](#)]
122. Berger, L.; Speare, R.; Hyatt, A. Chytrid fungi and amphibian declines: Overview, implications and future directions. In *Declines and Disappearances of Australian Frogs*; Campbell, A., Ed.; Environment Australia: Canberra, Australia, 1999; pp. 23–33.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).