REVIEW

SPECIAL ISSUE ON INVASIVE MAMMAL SPECIES

Genetic tools in the management of invasive mammals: recent trends and future perspectives

Samuel S. BROWETT* *Ecosystems and Environment Research Centre, School of Science, Engineering and Environment, University of Salford, Salford, M5 4WT, UK. Email: browett.sam@gmail.com*

Denise B. O'MEARA* *Molecular Ecology Research Group, Eco-Innovation Research Centre, School of Science and Computing, Waterford Institute of Technology, Waterford, Ireland. Email: domeara@wit.ie*

Allan D. MCDEVITT* *Ecosystems and Environment Research Centre, School of Science, Engineering and Environment, University of Salford, Salford, M5 4WT, UK. Email: a.mcdevitt@salford.ac.uk*

Keywords: detection, gene drives, genomics, global, hybridisation, mammals, next-generation sequencing

Running head: Genetic tools in the management of invasive mammals

* Correspondence authors

Received: 23 July 2019

Accepted: 27 January 2020

Editor: DR

Special Issue Guest Editor: Sandro Bertolino

ABSTRACT

1. Invasive non-native species are now considered to be one of the greatest threats to biodiversity worldwide. Therefore, efficient and cost-effective management of species invasions requires robust knowledge of their demography, ecology and impacts, and genetic-based techniques are becoming more widely adopted in acquiring such knowledge.
2. We focus on the use of genetic tools in the applied management of mammalian invasions globally, as well as on their inherent advantages and disadvantages. We cover tools that are used in: (1) detecting and monitoring mammalian invaders; (2) identifying origins and invasive pathways; (3) assessing and quantifying the negative impacts of invaders; and 4) population management and potential eradication of invasive mammals.
3. We highlight changes in sequencing technologies, including how the use of techniques such as Sanger sequencing and microsatellite genotyping, for monitoring and tracing invasive pathways respectively, are now giving way to the use of high-throughput sequencing methods. These include the emergence of environmental DNA (eDNA) metabarcoding for the early detection of invasive mammals, and single nucleotide polymorphisms or whole genomes to trace the sources of invasive populations. We are now moving towards trials of genome-editing techniques and gene drives to control or eradicate invasive rodents.
4. Genetic tools can provide vital information that may not be accessible with non-genetic methods, for the implementation of conservation policies (e.g. early detection using systematic eDNA surveillance, the identification of novel pathogens). However, the lack of clear communication of novel genetic

methods and results (including transparency and reproducibility) to relevant stakeholders can be prohibitive in translating these findings to appropriate management actions. Geneticists should engage early with stakeholders to co-design experiments in relation to management goals for invasive mammals.

INTRODUCTION

The introduction of species outside of their native range has escalated due to increased movement of people (Hulme 2009), and invasive non-native species are now considered to be one of the greatest threats to biodiversity worldwide (Bellard et al. 2016). Invasive species disrupt ecosystem services and lead to the introduction of novel diseases, ultimately impacting native wildlife, domesticated species and humans. In response to invasive non-native species, plans and policies are put into place to prevent their entry and reduce or eliminate their impact. Such measures are extremely costly in economic terms. For example, the European Union alone spends approximately €12 billion annually on the control and management of invasive non-native species and on mitigating their adverse impacts.

Efficient and cost-effective management of species invasions requires robust knowledge of their demography, ecology and impacts, and genetic-based techniques are becoming more widely adopted in acquiring such knowledge (Searle 2008, Darling et al. 2017). The genetic tools that we have to study these processes have developed dramatically over time, particularly over the last decade, and have become more affordable, efficient and available for small to medium-scale laboratories, providing new opportunities to study multiple aspects of invasions (Lee 2002). However, genetic tools are variable in methodology, design, price, complexity and the resolution of results. The scope of this review is to provide an accessible synopsis of the genetic techniques for the non-geneticist in order to enable stakeholders, such as state and conservation managers, policy-makers, field biologists and early-career researchers, to work collaboratively with geneticists to address questions related to the prevention and management of mammalian

invasions. We provide a brief overview of effective genetic techniques that are available for four management stages of a mammalian invasion: (1) detection and monitoring of non-native invasive mammals; (2) identifying invaders' origins and invasive pathways; (3) assessing and quantifying the negative impacts of invaders; and (4) population management and the potential eradication of invasive mammals.

DETECTION AND MONITORING

The early and rapid detection of newly introduced mammals is vital to prevent further spread that could subsequently result in a more costly eradication programme. Given the elusive nature of many mammalian species, detection and monitoring often requires indirect observations such as searching for latrines, faeces, hair, or tracks, or direct observations such as live-trapping or camera-trapping surveys (Sales et al. 2020a). These can require differing levels of expertise and resources, but despite high levels of expertise it is not always possible to assign indirect field signs correctly to a species without further confirmation via DNA analysis (Harrington et al. 2010).

Indirect field signs such as hair and faeces can be subjected to genetic non-invasive sampling (gNIS; Ferreira et al. 2018) to confirm species identification. gNIS has the benefit of collecting genetic information without handling animals, which may cause stress. Routine **PCR** (terms in **bold** are defined in the Glossary) methodologies can be applied as diagnostic tools for identifying species from ambiguous field signs such as hair or faeces. The required species-specific primers are already available to identify, for example, Iberian carnivores from faecal DNA, including invasive mammals such as the genet *Genetta genetta*, Egyptian Mongoose *Herpestes ichneumon* and the North American mink *Neovison vison* (Fernandes et al. 2008).

DNA obtained from gNIS may have degraded into smaller fragments due to prolonged exposure to environmental factors such as temperature fluctuations and ultraviolet light. Therefore, PCR detection or identification methods can be used to target short genetic regions (<1000 base pairs). **qPCR** is marginally more complex but has some benefits over traditional PCR for the identification of species from gNIS. qPCR can amplify shorter DNA regions (<100 base pairs) and is more sensitive to smaller starting amounts of DNA. It has been used to detect invasive mammals such as the greater white-toothed shrew *Crocidura russula* and grey squirrel *Sciurus carolinensis* from native pine marten *Martes martes* faeces (O'Meara et al. 2014). qPCR has the additional benefit of providing quality control to select optimal DNA samples for further analysis, such as sequencing and genotyping, thus allowing researchers to avoid wasting resources on poor-quality samples that are unlikely to yield results. Kierepka et al. (2016) used qPCR to screen feral pig *Sus scrofa* faecal-derived DNA prior to genotyping, to generate a robust capture-mark-recapture protocol in order to facilitate accurate estimates of abundance.

Physical samples such as faeces or hair are not always required for species detection. Organisms leave genetic material behind in the surrounding environment (e.g. in water bodies and soil) via excretions and secretions (Harper et al. 2019); this is referred to as **environmental DNA** (**eDNA**). Single-species detection from eDNA is possible using PCR, qPCR or droplet digital PCR (**ddPCR**). Research on invasive wild boar *Sus scrofa* in North America (Williams et al. 2018) has demonstrated the efficiency of a species-specific qPCR approach on samples from various water bodies in detecting the species, but has also highlighted that a minimum number of

131 individuals is required for detection. This clearly has implications for providing early
132 detection of invasive species, which may initially be present in low numbers.
133
134 Single-species detection methods are relatively cheap, fast and robust, but require
135 prior knowledge of the target species to design appropriate detection methods (e.g.
136 O'Meara et al. 2014). If prior knowledge of the target species is unavailable, species
137 can be identified from gNIS using **Sanger sequencing** to generate a **DNA barcode**
138 (Hebert et al. 2003). In the Scottish Highlands, UK, experienced field surveyors used
139 field signs such as faeces to identify 57 sites out of 147 as positive for the presence
140 of invasive North American mink. Subsequent DNA sequencing of a standardised
141 portion of **mitochondrial DNA** (mtDNA) showed that mink faeces were misidentified
142 at all sites, and that they were commonly confused with native carnivore faeces
143 (Harrington et al. 2010). Had management or eradication programmes been
144 designed based on indirect observation, the result would have been a costly, time-
145 consuming, and unnecessary eradication programme.
146
147 **Next-generation sequencing** can facilitate the simultaneous identification of entire
148 communities (i.e. multiple species). **DNA metabarcoding** from environmental
149 samples has the potential to be used as an early warning system for the detection of
150 invasive non-native species, can be used for continuous monitoring programmes,
151 and has been extensively applied for tracking biological invasions in aquatic
152 ecosystems (Deiner et al. 2017). eDNA metabarcoding studies targeting mammalian
153 communities were relatively rare in comparison to other taxonomic groups (Sales et
154 al. 2020a), but this may change now that there are established metabarcoding

protocols for detecting and monitoring whole communities using vertebrate (Harper et al. 2019) or mammal-specific primer sets (Ushio et al. 2017, Sales et al. 2020a,b).

eDNA metabarcoding is an emerging technique for invasive mammal detection and monitoring, and there are important considerations for its use. For example, mammals with larger home ranges (e.g. invasive carnivores) have lower probabilities of detection than more abundant group-living mammals (Harper et al. 2019, Sales et al. 2020a). Due to the high sensitivity of metabarcoding, contamination is a concern (Sales et al. 2020a). It is therefore essential that specialised eDNA lab facilities (akin to working with ancient DNA) are used (Zinger et al. 2019). Another consideration is the existence of gaps in customised or online **reference databases** for identifying sequences to the appropriate species level in under-studied geographic regions (Sales et al. 2020b). However, with a carefully planned experimental design and the appropriate field and lab controls (Zinger et al. 2019), eDNA metabarcoding has the potential to be applied for early detection and ongoing surveillance of invasive mammals (Harper et al. 2019, Sales et al. 2020a).

ORIGINS AND INVASIVE PATHWAYS

Identifying the origins of invasions is a critical management strategy in controlling the spread of invasive species (Hulme 2009). When there is an absence of direct evidence indicating the routes of invasion (such as records from interception at ports), indirect methods such as the analyses of genetic data from invasive populations and putative sources becomes vital (Searle 2008, Gargan et al. 2016).

Studies initially relied upon sequencing mtDNA to track the transport of invasive species, because of the availability of universal primers for mammals (for mtDNA genes such as cytochrome *b* and the control region) and available sequences (from the native ranges for comparisons) in reference databases such as Genbank. For invasive mammals with a global distribution, such as house mice *Mus* spp. and rats *Rattus* spp., phylogenetic analyses of mtDNA have proven extremely useful in tracing multiple introductions to islands and different continents over recent millennia and centuries (Jones et al. 2013). The use of this type of mtDNA marker can be limited over the spatial and temporal scales required for tracking more recent invasions. Although mtDNA accumulates substitutions more rapidly than nuclear DNA, mtDNA markers are generally useful for investigating intraspecific relationships over tens to hundreds of thousands of years. Unless mtDNA variation is sufficiently high in the native range, it is not ideal for tracing most mammalian invasions (Gray et al. 2014) and may reveal the continent of origin as opposed to the country (Gargan et al. 2016). Raccoons *Procyon lotor* show limited mtDNA diversity within their invasive range in Europe, which originally led researchers to believe that they were descended from a small number of founding individuals (Frantz et al. 2013). However, the analysis of more rapidly evolving **microsatellites** led to the conclusion that there were potentially up to four separate sources for the raccoon's current distribution within its invasive range (Fischer et al. 2015). In the same vein, studies of house mice have revealed the importance of using a multiple marker approach (such as microsatellites) when inferring the origins of island populations, as many display admixed origins (e.g. Gray et al. 2014).

Given that we are now firmly entrenched within the genomics era of molecular

ecology research, it is unsurprising that studies inferring the origins of invasive mammals are now switching to **Single Nucleotide Polymorphism (SNP)** marker-based approaches. Compared to microsatellite markers, SNPs usually span across a greater proportion of the genome, can determine population demographics to a finer scale, and do not require calibration between laboratories (Iacolina et al. 2016). Incorporating SNPs in a study previously required a huge investment of time and resources, usually applied only to economically important species (e.g. cattle, dogs, rodents, pigs). The *de novo* discovery of SNPs in non-model organisms is now achievable and affordable through **reduced representation sequencing** techniques (such as Restriction Site Associated sequencing or RAD-seq; Baird et al. 2008).

Puckett et al. (2016) used ~32000 SNPs (derived from ddRAD genotyping) to examine the population genomic structure of brown rats *Rattus norvegicus* throughout their worldwide geographic range. Brown rats were generally grouped into Asian and non-Asian groups, but fine-scale structuring was identified within regions, reflecting more recent invasion pathways. For example, mtDNA data revealed a European origin for contemporary New Zealand and western USA populations, but SNP data revealed ancestry from admixed Asian and non-Asian genomic clusters. In tracking the invasion of raccoon dogs *Nyctereutes procyonoides* in Denmark, Nørgaard et al. (2017) utilised genotyping-by-sequencing to identify over 4000 SNPs to trace their origins to Danish fur farms and reveal subsequent admixture with neighbouring German populations. Unlike with microsatellites, newly generated data on finer spatial scales can be compared with a global dataset of SNPs if a reference genome is available. For example, this allowed Combs et al.

(2018) to determine that the most likely origins of the New York, USA, population of brown rats were France and the British Isles.

NEGATIVE IMPACTS

Diet and competition

Invasive mammals may affect local flora and fauna through predation or ingestion (e.g. feral cats take terrestrial vertebrates; Doherty et al. 2017), or via increased competition (e.g. invasive American mink compete with native European mammalian carnivores Sidorovich et al. 2010). Mammals are notoriously elusive, making their diet difficult to document through direct observations, so that morphological diagnostics of prey remains from stomach contents and faeces are a popular method (Brzeziński et al. 2018). This methodology produces biased results due to variable degradation rates between species and body parts (i.e. soft body parts degrade faster than hard body parts), and residual body fragments that are found are difficult to identify to species level (Deagle et al. 2009). Stable isotope analysis shows promise, but has difficulties identifying prey species when isotopic signatures naturally vary between geographic locations (Chibowski et al. 2019).

Genetic tools require DNA to be extracted from faeces or gut contents using appropriate extraction kits capable of removing inhibitors associated with the digestive tract. Species-specific primers and PCR are straight-forward and cost-effective methods to measure predation rates of a single species of interest (Waraniak et al. 2018). However, invasive mammals can have a variable diet between native and introduced ranges (Ballari & Barrios-García 2014), making it difficult to predict what they will consume in their introduced range. DNA

metabarcoding is a promising method: it allows the identification of multiple dietary components of hundreds of individuals, and increases prey detection from 2% using morphological diagnostics to 70% using metabarcoding (Pompanon et al. 2012, Egeter et al. 2015a).

Not only can DNA metabarcoding accurately document an animal's impact on local resources, but it can also reduce ambiguity. Previous assessments of the impact of invasive rats *Rattus rattus* on endemic amphibians in New Zealand relied on abundance estimates of native frog species in comparison to arrival patterns of the invasive rat (Egeter et al. 2015b). Inconsistencies between observers caused doubt, but DNA metabarcoding clarified the rat's consumption of New Zealand's native frog species and its contribution to the population declines (Egeter et al. 2019). The sensitivity achieved from next-generation sequencing methods allows multiple prey items to be identified to the species level and generates a comprehensive account of multiple animals' resource use and overlap. Telfair's skink *Leiolopisma telfairii* was introduced to Ile aux Aigrettes, Mauritius, Indian Ocean, for conservation purposes, but unexpectedly met potential threats from the invasive Asian musk shrew *Suncus murinus*. Species-specific primers showed the two species did not predate one another (once adulthood was attained), but DNA metabarcoding identified significant prey overlap and resulted in the suggestion that controlling shrew populations would benefit the skink population (Brown et al. 2014).

Metabarcoding projects for dietary studies require some important considerations before they are started (also relevant to eDNA metabarcoding studies, see above). The first is targeting the appropriate genetic region for the target taxa in the diet,

such as vertebrates, invertebrates or plants (Kress et al. 2015). To know the full diet of an omnivorous invader (e.g. wild boar), multiple regions are required for the full taxonomic range within their diet (De Barba et al. 2014). Alternatively, highly degenerative (non-specific) primers can be used to capture a wider range of prey taxa, but this can result in over-representation of higher-quality host DNA (Zeale et al. 2011). The broader the primers' taxonomic range, the more likely the chance of amplifying non-target taxa and reducing the amount of information on a species' diet. Blocking primers can mitigate host DNA amplification, but require more time to design and test, as they may also block the amplification of some target prey taxa (Su et al. 2018). The high sensitivity of PCR and high-throughput sequencing can also result in the detection of taxa through secondary predation (i.e. detecting the food of the food; Sheppard et al. 2005). Another difficulty is the inference of biomass or the number of prey individuals from molecular diet analysis (Deagle et al. 2019). Estimates of prey proportion are biased towards harder-bodied organisms due to differential degradation rates. There are multiple ways to determine the importance of certain taxa within a predator's diet, such as frequency of occurrence or relative abundance (reviewed by Deagle et al. 2019).

Disease

The introduction of mammals into novel environments comes with the risk of co-introducing pathogens or parasites that local fauna have not yet developed resistance to (Paziewska et al. 2011). Mammalian invasions in Europe are likely to have been responsible for the transport of pathogens responsible for salmonellosis, toxoplasmosis and leptospirosis (Hulme 2014), and for the dissemination of the plague across continents via rodent introductions (Gage & Kosoy 2005). Genetic

303 tools are becoming pivotal in disease management in wildlife (DeCandia et al. 2018):
304 PCR is currently used to verify morphological identification of pathogens and
305 parasites (Bagrade et al. 2016), and genetic tools can be used as detection methods
306 when there are difficulties in recreating optimal cell growing conditions to test for
307 prevalence levels (Heuser et al. 2017).
308
309 Different pathogen genotypes or strains can have different infection capabilities
310 (Nally et al. 2016). Sequencing actin genes of pathogenic *Cryptosporidium* revealed
311 that invasive raccoons harboured genotypes capable of infecting humans
312 (Leśniańska et al. 2016). To gain a higher resolution of bacterial population structure
313 and evolution, and to help understand the distribution of pathogenic species and
314 genotypes in novel areas invaded by mammalian hosts, multiple loci or genes can be
315 sequenced in multi-locus sequence typing (Margos et al. 2008). This method was
316 applied to *Borrelia* spp., an important pathogen in zoonotic ecology due to its
317 responsibility for Lyme disease. Sanger sequencing of the housekeeping gene (*clpA*)
318 and the infection-related gene (*ospC*) of *Borrelia burgdorferi* showed that invasive
319 grey squirrels in the UK are reservoirs for multiple *Borrelia burgdorferi* strains that
320 can affect multiple vertebrate clades (Millins et al. 2015). For larger-scale projects
321 and maximum efficiency, next-generation sequencing can be adapted for multi-locus
322 sequence typing from 100–200 samples in a cost-effective manner (Jacquot et al.
323 2014); this method was used to identify different *Borrelia* spp. lineages associated
324 with different small mammal host species (Jacquot et al. 2014).
325
326 Standardisation of sequence data is encouraged, and uploading data to online
327 databases allows combinations of multiple datasets to be incorporated into new and

broad meta-analyses (Maiden 2006). Phylogenetic analysis of openly available sequence data from online reference databases allowed Hayman et al. (2013) to decipher the origins, dissemination and diversification of the zoonotic pathogen *Bartonella* spp. in mammalian clades and introductions.

Hybridisation

Hybridisation among species which are naturally separated is undoubtedly increasing due to anthropogenic impacts, including species' invasions (McFarlane & Pemberton 2019). Extensive introgression from invading populations can put already endangered native populations at risk (Senn et al. 2019). Identifying hybrids based on phenotypic characteristics is problematic due to intermediate phenotypes and observer biases (McDevitt et al. 2009). To increase the efficiency of hybrid identification, molecular markers have long been deployed; microsatellites have been used since the 1990s. The increased use of assignment-based analysis in the early 2000s (e.g. Randi et al. 2001) allowed researchers to identify the proportion of the genome (usually inferring from ≥10 microsatellites) assigned to each species in each individual, which individuals exhibited an admixed genotype and could therefore be labelled as hybrids, and the percentage of the population consisting of hybrids. This type of analyses has been important in providing initial indications of the level of hybridisation between invasive sika deer *Cervus nippon* and red deer *Cervus elaphus* in Europe (e.g. McDevitt et al. 2009), domestic/feral cats and wildcats *Felis silvestris* in Europe (e.g. Randi et al. 2001) and domestic/feral pigs and wild boar in multiple geographic regions (e.g. Scandura et al. 2011). While microsatellite markers can be informative in detecting first or second-generation hybridisation events, their low coverage means that they cannot detect extensive

353 backcrossing over several generations between parental species (McFarlane &
354 Pemberton 2019).
355
356 In order to improve resolution in detecting hybrids and their backcrosses, there is
357 clearly a need to use higher-density and diagnostic SNPs (Mattucci et al. 2019).
358 Several recent studies have highlighted improvements in hybrid detection by using
359 thousands of SNPs rather than 10 - 25 microsatellites. For example, a study on wolf-
360 dog hybridisation showed that only 1-5% of individuals were identified as hybrids
361 when 16 or 18 microsatellites were used (Randi 2008). A later study used 61000
362 SNPs to infer that 62% of Eurasian wolves *Canis lupus* had some level of admixture
363 with domestic dogs *Canis familiaris* (Pilot et al. 2018). In a well-studied hybrid zone
364 between sika deer and red deer in Kintyre, Scotland, an increased panel of 45000
365 SNPs reclassified 26% of individuals as hybrids that had originally been assigned to
366 one of the parental species from a previous study based on 22 microsatellites
367 (McFarlane et al. 2019). In attempting to preserve Scottish wildcats from extensive
368 introgression with feral/domestic cats, only wildcat individuals with both high genetic
369 scores (using a SNP panel) and high phenotype scores of wildcat ‘purity’ are
370 selected for the captive breeding and reintroduction programmes (Senn et al. 2019).
371
372 **MANAGEMENT AND ERADICATION**
373 Given the financial and social commitments required from stakeholders to undertake
374 long-term eradication programmes of species such as grey squirrels and American
375 mink, it is important to be able to gauge the success and impact of these efforts.
376 Microsatellites are very effective in determining recent changes in invasive mammal

377 population demographics in order to assess the progress of management and
378 control schemes (Velando et al. 2017).
379
380 Culling programs are well-established for the control of American mink in several
381 European countries. Fraser et al. (2013) used microsatellites to divide the Scottish
382 mink population into genetic clusters (sub-populations) which were classified as
383 management units. These units were formed through a combination of historical fur-
384 farm escapes and subsequent natural movement through a mosaic landscape
385 throughout Scotland. The genetic analysis of Scottish mink populations
386 corresponded to the habitat characteristics, and allowed Fraser et al. (2013) to
387 create an informed proposal on how to reduce the spread of the species and decide
388 where to direct eradication efforts. However, Oliver et al. (2016) used similar data to
389 identify a possible mechanism for populations in mainland Scotland remaining
390 relatively stable despite culling. They identified an increase in long-distance
391 immigration and an almost three-fold increase in male immigration into culled areas,
392 providing evidence of compensatory immigration during these culling efforts.
393
394 As with identifying invasive pathways with genetic markers, SNPs can provide higher
395 resolution of population demographics, and have been implemented instead of
396 traditional capture-mark-recapture methods to show connectivity and dispersal in
397 brown rat populations in an urban area (Combs et al. 2018). Piertney et al. (2016)
398 used 299 SNPs to identify genomic clusters of brown rats on the island of South
399 Georgia in order to identify the appropriate number of target areas for baiting
400 operations. Although these types of data (microsatellites and SNPs) and analyses
401 (population structure and gene flow) are useful for planning and assessing the

success of management and eradication programmes, an important consideration is the likely response of the invader to control or eradication measures, whether these be chemical or biological. For example, using a genome-wide SNPs, Morgan et al. (2018) demonstrated that invasive house mice on islands off North and South America did not possess rodenticide resistance alleles that are present in parts of Europe (even though the study also found that these house mice were of European ancestry). This has important implications for subsequent eradication and control measures.

Even when genetic tools are used to identify key populations to target for culling programs, culls require a lot of effort and usually only have the power to manage a population rather than eradicate it completely. Island populations of invasive mice have been proposed as targets for trialling more elaborate eradication programmes involving genome-editing techniques using **CRISPR-Cas9** (Breed et al. 2019). The concept of gene drives, whereby the use of genetic engineering alters the probability of how specific alleles are inherited in future generations of offspring, is being tested for eradication programmes in multiple invasive species, particularly invertebrates (see Breed et al. 2019 for examples). Mammals are now being considered, and well-studied model organisms such as mice are an obvious starting point. Transgenic delivery of the male sex-determining factor (Sry) has been proposed to skew the sex ratio heavily towards male mice and thereby control population size (Backus & Gross 2016). This would require repeated releases of engineered males, which could be feasible on small islands (Campbell et al. 2015). Prowse et al. (2019) demonstrated that the Y chromosome can be ‘shredded’ using CRISPR technology in mouse embryonic stem cells, and individual-based simulations show that this targeted

deletion of a sex chromosome has the potential for eradicating an island population of rodents. However, it would require >90% efficiency to produce high probabilities of eradication success, and would be highly susceptible to changes in mating systems and population size (Prowse et al. 2019).

There are additional concerns if genetically altered individuals are ever accidentally released or spread beyond their target areas of control (a hallmark of effective invaders), as gene drives are self-sustaining (Noble et al. 2018). Despite these justifiable concerns, plans are already underway to bring this technique to the field and to select an appropriate island for trials (Scudellari 2019). The use of the technique is clearly complex in terms of scientific, social, regulatory and ethical issues (Breed et al. 2019), and it remains to be determined how effective gene drives will be over large geographic areas. However, gene drives offer a potentially more targeted approach than the use of chemicals, which could impact non-target, native species.

CONCLUSION

The impact and challenges of surveillance programmes related to invasive species are recognised internationally and have led to the creation of policies aimed at preventing and managing invasions. For example, the European Union has created Regulation 1143/2014 on Invasive Alien Species (IAS Regulation) that contains three measures to combat invasive species which include: (1) prevention, (2) early detection and rapid eradication and (3) management. This review has highlighted that genetic tools have multiple applications for the active management of invasive mammalian species. Not only this, but they are reliable, robust, and provide vital

information, that may not be accessible with non-genetic methods, for the implementation of conservation policies (e.g. early detection using systematic eDNA surveillance and the identification of novel pathogens).

However, there are technical challenges associated with the standardisation of genetic methodologies and bioinformatic pipelines used between laboratories, even when researchers are attempting to address similar questions (Zinger et al. 2019). For example, how the samples have been collected (e.g. gNIS or tissue) and stored (e.g. in ethanol or frozen) has implications for what techniques can be performed downstream in the laboratory (e.g. single gene or whole-genome approaches). Another significant challenge is the availability of appropriate funding and expertise. These factors can all limit what questions can be addressed that will translate into management actions and decisions.

In addition, the lack of clear communication of novel genetic methods and results (including transparency and reproducibility) to relevant stakeholders can be prohibitive in translating these findings to appropriate management and eradication action on the ground (Mosher et al. 2019, Ward et al. 2020). These communication challenges have been well documented in relation to the marine sector (e.g. Darling et al. 2017), but little coordination has taken place in relation to invasive mammals, despite the environmental and economic consequences that invasive non-native species pose to native species, habitats and the agricultural industry. Geneticists should engage early with stakeholders in relation to project costs, duration and management goals for invasive mammals. This will allow for robust experimental

476 design using existing genetic tools, and the development of new technologies that
477 can be tailored towards specific management issues (Mosher et al. 2019).
478

REFERENCES

Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA (2016) Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics* 17: 81-92.

Backus GA, Gross K (2016) Genetic engineering to eradicate invasive mice on islands: modeling the efficiency and ecological impacts. *Ecosphere* 7: e01589.

Bagrade G, Deksne G, Ozoliņa Z, Howlett SJ, Interisano M, Casulli A, Pozio E (2016) *Echinococcus multilocularis* in foxes and raccoon dogs: an increasing concern for Baltic countries. *Parasites and Vectors* 9: 615.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA et al. (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3: e3376.

Ballari SA, Barrios-García MN (2014) A review of wild boar *Sus scrofa* diet and factors affecting food selection in native and introduced ranges. *Mammal Review* 44: 124–134.

Bellard C, Cassey P, Blackburn TM (2016) Alien species as a driver of recent extinctions. *Biology Letters* 12: 20150623.

Breed MF, Harrison PA, Blyth C, Byrne M, Gaget V, Gellie NJC et al. (2019) The potential of genomics for restoring ecosystems and biodiversity. *Nature Reviews Genetics* 20: 615–628.

Brown DS, Burger R, Cole N, Vencatasamy D, Clare EL, Montazam A, Symondson WOC (2014). Dietary competition between the alien Asian musk shrew (*Suncus murinus*) and a re-introduced population of Telfair's skink (*Leiolopisma telfairii*). *Molecular Ecology* 23: 3695–3705.

Brzeziński M, Chibowska P, Zalewski A, Borowik T, Komar E (2018) Water vole

(*Arvicola amphibius*) population under the impact of the American mink (*Neovison vison*): are small midfield ponds safe refuges against this invasive predator? *Mammalian Biology* 93: 182–188.

Campbell KJ, Beek J, Eason CT, Glen AS, Godwin J, Gould F et al. (2015) The next generation of rodent eradications: innovative technologies and tools to improve species specificity and increase their feasibility on islands. *Biological Conservation* 185: 47–58.

Chibowski P, Zalewski A, Suska-Malawska M, Brzeziński M (2019) Study on geographical differences in American mink diets reveals variations in isotopic composition of potential mink prey. *Mammal Research* 64: 343-351.

Combs M, Puckett EE, Richardson J, Mims D, Munshi-South J (2018) Spatial population genomics of the brown rat (*Rattus norvegicus*) in New York City. *Molecular Ecology* 27: 83–98.

Darling JA, Galil BS, Carvalho GR, Rius M, Viard F, Piraino S (2017) Recommendations for developing and applying genetic tools to assess and manage biological invasions in marine ecosystems. *Marine Policy* 85: 54–64.

De Barba M, Miquel C, Boyer F, Mercier C, Rioux D, Coissac E, Taberlet P (2014) DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. *Molecular Ecology Resources* 14: 306–323.

Deagle BE, Clarke LJ, Thomas AC, Mcinnes JC, Vesterinen EJ, Clare EL et al. (2019) Counting with DNA in metabarcoding studies: how should we convert sequence reads to dietary data? *Molecular Ecology* 28: 391-406.

Deagle BE, Kirkwood R, Jarman SN (2009) Analysis of Australian fur seal diet by pyrosequencing prey DNA in faeces. *Molecular Ecology* 18: 2022–2038.

529 DeCandia AL, Dobson AP, vonHoldt BM (2018) Toward an integrative molecular
530 approach to wildlife disease. *Conservation Biology* 32: 798–807.
531 Deiner K, Bik HM, Elvira M, Seymour M, Lacoursiére-Roussel A, Altermatt F et al.
532 (2017) Environmental DNA metabarcoding: transforming how we survey animal
533 and plant communities. *Molecular Ecology* 26: 5872–5895.
534 Doherty TS, Dickman CR, Johnson CN, Legge SM, Ritchie EG, Woinarski JCZ
535 (2017) Impacts and management of feral cats (*Felis catus*) in Australia. *Mammal*
536 *Review* 47: 83–97.
537 Egeter B, Bishop PJ, Robertson BC (2015a) Detecting frogs as prey in the diets of
538 introduced mammals: a comparison between morphological and DNA-based
539 diet analyses. *Molecular Ecology Resources* 15: 306–316.
540 Egeter B, Robertson BC, Bishop PJ (2015b) A synthesis of direct evidence of
541 predation on amphibians in New Zealand, in the context of global invasion
542 biology. *Herpetological Review* 46: 512–519.
543 Egeter B, Roe C, Peixoto S, Puppo P, Easton LJ, Pinto J et al. (2019) Using
544 molecular diet analysis to inform invasive species management: a case study of
545 introduced rats consuming endemic New Zealand frogs. *Ecology and Evolution*
546 9: 5032-5048.
547 Fernandes CA, Ginja C, Pereira I, Tenreiro R, Bruford MW, Santos-Reis M (2008)
548 Species-specific mitochondrial DNA markers for identification of non-invasive
549 samples from sympatric carnivores in the Iberian Peninsula. *Conservation*
550 *Genetics* 9: 681–690.
551 Ferreira CM, Sabino-Marques H, Paupério J, Barbosa S, Costa P, Encarnação C et
552 al. (2018) Genetic non-invasive sampling (gNIS) as a cost-effective tool for
553 monitoring elusive small mammals. *European Journal of Wildlife Research* 64:

554 46.
555 Fischer ML, Hochkirch A, Heddergott M, Schulze C, Anheyer-Behmenburg HE, Lang
556 J et al. (2015) Historical invasion records can be misleading: genetic evidence
557 for multiple introductions of invasive raccoons (*Procyon lotor*) in Germany. *PLoS*
558 *ONE* 10: e0125441.
559 Frantz AC, Heddergott M, Lang J, Schulze C, Ansorge H, Runge M et al. (2013)
560 Limited mitochondrial DNA diversity is indicative of a small number of founders
561 of the German raccoon (*Procyon lotor*) population. *European Journal of Wildlife*
562 *Research* 59: 665–674.
563 Fraser EJ, Macdonald DW, Oliver MK, Piertney S, Lambin X (2013) Using population
564 genetic structure of an invasive mammal to target control efforts – an example of
565 the American mink in Scotland. *Biological Conservation* 167: 35–42.
566 Gage KL, Kosoy MY (2005) Natural history of the plague: perspectives from more
567 than a century of research. *Annual Review of Entomology* 50: 505–528.
568 Gargan LM, Cornette R, Yearsley JM, Montgomery WI, Paupério J, Alves PC et al.
569 (2016) Molecular and morphological insights into the origin of the invasive
570 greater white-toothed shrew (*Crocidura russula*) in Ireland. *Biological Invasions*
571 18: 857–871.
572 Gray MM, Wegmann D, Haasl RJ, White MA, Gabriel SI, Searle JB et al. (2014)
573 Demographic history of a recent invasion of house mice on the isolated island of
574 Gough. *Molecular Ecology* 23: 1923–1939.
575 Harper LR, Handley LL, Carpenter AI, Ghazali M, Di Muri C, Macgregor CJ et al.
576 (2019) Environmental DNA (eDNA) metabarcoding of pond water as a tool to
577 survey conservation and management priority mammals. *Biological*
578 *Conservation* 238: 108225.

579 Harrington LA, Harrington AL, Hughes J, Stirling D, Macdonald DW (2010) The
580 accuracy of scat identification in distribution surveys: American mink, (*Neovison*
581 *vison*), in the northern highlands of Scotland. *European Journal of Wildlife*
582 *Research* 56: 377–384.
583 Hayman DTS, McDonald KD, Kosoy MY (2013) Evolutionary history of rat-borne
584 *Bartonella*: the importance of commensal rats in the dissemination of bacterial
585 infections globally. *Ecology and Evolution* 3: 3195–3203.
586 Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003) Biological identifications
587 through DNA barcodes. *Proceedings of the Royal Society of London. Series B:*
588 *Biological Sciences* 270: 313–321.
589 Heuser E, Fischer S, Ryll R, Mayer-Scholl A, Hoffmann D, Spahr C et al. (2017)
590 Survey for zoonotic pathogens in Norway rat populations from Europe. *Pest*
591 *Management Science* 73: 341–348.
592 Hulme PE (2009) Trade, transport and trouble: managing invasive species pathways
593 in an era of globalization. *Journal of Applied Ecology* 46: 10–18.
594 Hulme PE (2014) Invasive species challenge the global response to emerging
595 diseases. *Trends in Parasitology* 30: 267–270.
596 Iacolina L, Scandura M, Goedbloed DJ, Alexandri P, Crooijmans RPMA, Larson G et
597 al. (2016) Genomic diversity and differentiation of a managed island wild boar
598 population. *Heredity* 116: 60–67.
599 Jacquot M, Bisseux M, Abrial D, Marsot M, Ferquel E, Chapuis JL et al. (2014) High-
600 throughput sequence typing reveals genetic differentiation and host
601 specialization among populations of the *Borrelia burgdorferi* species complex
602 that infect rodents. *PLoS ONE* 9: e88581
603 Jones EP, Eager HM, Gabriel SI, Jóhannesdóttir F, Searle JB (2013) Genetic

tracking of mice and other bioproxies to infer human history. *Trends in Genetics* 29: 298–308.

Kierepka EM, Unger SD, Keiter DA, Beasley JC, Rhodes Jr OE, Cunningham FL, Piaggio AJ (2016) Identification of robust microsatellite markers for wild pig fecal DNA. *Journal of Wildlife Management* 80: 1120–1128.

Kress WJ, García-Robledo C, Uriarte M, Erickson DL (2015) DNA barcodes for ecology, evolution, and conservation. *Trends in Ecology and Evolution* 30: 25–35.

Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology and Evolution* 17: 386–391.

Leśniańska K, Perec-Matysiak A, Hildebrand J, Buńkowska-Gawlik K, Piróg A, Popiołek M (2016). *Cryptosporidium spp*. and *Enterocytozoon bieneusi* in introduced raccoons (*Procyon lotor*)—first evidence from Poland and Germany. *Parasitology Research* 115: 4535–4541.

Maiden MCJ (2006) Multilocus sequence typing of bacteria. *Annual Review of Microbiology* 60: 561–588.

Margos G, Gatewood AG, Aanensen DM, Hanincova K, Terekhova D, Vollmer SA et al. (2008) MLST of housekeeping genes captures geographic population structure and suggests a European origin of *Borrelia burgdorferi*. *Proceedings of the National Academy of Sciences* 105: 8730–8735.

Mattucci F, Galaverni M, Lyons LA, Alves PC, Randi E (2019) Genomic approaches to identify hybrids and estimate admixture times in European wildcat populations. *Scientific Reports* 9: 11612.

McDevitt AD, Edwards CJ, O'Toole P, O'Sullivan P, O'Reilly C, Carden RF (2009) Genetic structure of, and hybridisation between, red (*Cervus elaphus*) and sika

(*Cervus nippon*) deer in Ireland. *Mammalian Biology* 74: 263–273.

McFarlane SE, Hunter DC, Senn HV, Smith SL, Holland R, Huisman J, Pemberton JM (2019) Increased genetic marker density reveals high levels of admixture between red deer and introduced Japanese sika in Kintyre, Scotland. *Evolutionary Applications* doi: 10.1111/eva.12880.

McFarlane SE, Pemberton JM (2019) Detecting the true extent of introgression during anthropogenic hybridization. *Trends in Ecology and Evolution* 34: 315-326.

Millins C, Magierecka A, Gilbert L, Edoff A, Brereton A, Kilbride E et al. (2015) An invasive mammal (the gray squirrel, *Sciurus carolinensis*) commonly hosts diverse and atypical genotypes of the zoonotic pathogen *Borrelia burgdorferi* sensu lato. *Applied and Environmental Microbiology* 81: 4236–4245.

Morgan AP, Didion JP, Hughes JJ, Searle JB, Jolley WJ, Campbell KJ et al. (2018) Genetic characterization of invasive house mouse populations on small islands. *BioRxiv*, 332064.

Mosher BA, Bernard RF, Lorch JM, Miller DAW, Richgels KLD, White CL, Grant EHC (2019) Successful molecular detection studies require clear communication among diverse research partners. *Frontiers in Ecology and the Environment* doi: 10.1002/fee.2141

Nally JE, Arent Z, Bayles DO, Hornsby RL, GilmoreC, Regan S et al. (2016) Emerging infectious disease implications of invasive mammalian species: the greater white-toothed shrew (*Crocidura russula*) is associated with a novel serovar of pathogenic leptospira in Ireland. *PLoS Neglected Tropical Diseases* 10: e0005174.

Noble C, Adlam B, Church GM, Esvelt KM, Nowak MA (2018) Current CRISPR gene

654 drive systems are likely to be highly invasive in wild populations. *eLife* 7:
655 e33423.
656 Nørgaard LS, Mikkelsen DMG, Elmeros M, Chriél M, Madsen AB, Nielsen JL et al.
657 (2017) Population genomics of the raccoon dog (*Nyctereutes procyonoides*) in
658 Denmark: insights into invasion history and population development. *Biological*
659 *Invasions* 19: 1637–1652.
660 O'Meara DB, Sheehy E, Turner PD, O'Mahony D, Harrington AP, Denman H et al.
661 (2014) Non-invasive multi-species monitoring: real-time PCR detection of small
662 mammal and squirrel prey DNA in pine marten (*Martes martes*) scats. *Acta*
663 *Theriologica* 59: 111–117.
664 Oliver MK, Piertney SB, Zalewski A, Melero Y, Lambin X (2016) The compensatory
665 potential of increased immigration following intensive American mink population
666 control is diluted by male-biased dispersal. *Biological Invasions* 18: 3047–3061.
667 Paziewska A, Harris PD, Zwolińska L, Bajer A, Siński E (2011) Recombination within
668 and between species of the alpha proteobacterium *Bartonella* infecting rodents.
669 *Microbial Ecology* 61: 134–145.
670 Piertney SB, Black A, Watt L, Christie D, Poncet S, Collins MA (2016) Resolving
671 patterns of population genetic and phylogeographic structure to inform control
672 and eradication initiatives for brown rats *Rattus norvegicus* on South Georgia.
673 *Journal of Applied Ecology* 53: 332–339.
674 Pilot M, Greco C, VonHoldt BM, Randi E, Jędrzejewski W, Sidorovich VE et al.
675 (2018) Widespread, long-term admixture between grey wolves and domestic
676 dogs across Eurasia and its implications for the conservation status of hybrids.
677 *Evolutionary Applications* 11: 662–680.
678 Pompanon F, Deagle BE, Symondson WOC, Brown DS, Jarman SN, Taberlet P

(2012) Who is eating what: diet assessment using next generation sequencing. *Molecular Ecology* 21: 1931–1950.

Prowse TAA, Adikusuma F, Cassey P, Thomas P, Ross JV (2019) A Y-chromosome shredding gene drive for controlling pest vertebrate populations. *eLife* 8: e41873.

Puckett EE, Park J, Combs M, Blum MJ, Bryant JE, Caccone A et al. (2016) Global population divergence and admixture of the brown rat (*Rattus norvegicus*). *Proceedings of the Royal Society B: Biological Sciences* 283: 20161762.

Randi E (2008) Detecting hybridization between wild species and their domesticated relatives. *Molecular Ecology* 17: 285–293.

Randi E, Pierpaoli M, Beaumont M, Ragni B, Sforzi A (2001) Genetic identification of wild and domestic cats (*Felis silvestris*) and their hybrids using Bayesian clustering methods. *Molecular Biology and Evolution* 18: 1679–1693.

Sales NG, Mckenzie MB, Drake J, Harper LR, Browett SS, Coscia I et al. (2020a) Fishing for mammals: landscape-level monitoring of terrestrial and semi-aquatic communities using eDNA from lotic ecosystems. *Journal of Applied Ecology* in press.

Sales NG, da Cruz Kaizer M, Coscia I, Perkins J, Highlands A, Boubli JP et al. (2020b) Assessing the potential of environmental DNA metabarcoding for monitoring Neotropical mammals: a case study in the Amazon and Atlantic Forest, Brazil. *Mammal Review* doi: 10.1111/mam.12183.

Scandura M, Iacolina L, Apollonio M (2011) Genetic diversity in the European wild boar *Sus scrofa*: phylogeography, population structure and wild x domestic hybridization. *Mammal Review* 41: 125–137.

Scudellari M (2019) Self-destructing mosquitoes and sterilized rodents: the promise

of gene drives. *Nature* 571: 160–162.

Searle JB (2008) The genetics of mammalian invasions: a review. *Wildlife Research* 35: 185–192.

Senn HV, Ghazali M, Kaden J, Barclay D, Harrower B, Campbell RD et al. (2019) Distinguishing the victim from the threat: SNP-based methods reveal the extent of introgressive hybridization between wildcats and domestic cats in Scotland and inform future in situ and ex situ management options for species restoration. *Evolutionary Applications* 12: 399–414.

Sheppard SK, Bell J, Sunderland KD, Fenlon J, Skervin D, Symondson WOC (2005) Detection of secondary predation by PCR analyses of the gut contents of invertebrate generalist predators. *Molecular Ecology* 14: 4461–4468.

Sidorovich VE, Polozov AG, Zalewski A (2010) Food niche variation of European and American mink during the American mink invasion in north-eastern Belarus. *Biological Invasions* 12: 2207–2217.

Su M, Liu H, Liang X, Gui L, Zhang J (2018) Dietary analysis of marine fish species: enhancing the detection of prey-specific DNA sequences via high-throughput sequencing using blocking primers. *Estuaries and Coasts* 41: 560–571.

Ushio M, Fukuda H, Inoue T, Makoto K, Kishida O, Sato K et al. (2017) Environmental DNA enables detection of terrestrial mammals from forest pond water. *Molecular Ecology Resources* 17: e63–e75.

Velando A, Morán P, Romero R, Fernández J, Piorno V (2017) Invasion and eradication of the American mink in the Atlantic Islands National Park (NW Spain): a retrospective analysis. *Biological Invasions* 19: 1227–1241.

Waraniak JM, Blumstein DM, Scribner KT (2018) Barcoding PCR primers detect larval lake sturgeon (*Acipenser fulvescens*) in diets of piscine predators.

729 *Conservation Genetics Resources* 10: 259–268.
730 Ward AI, Richardson S, Macarthur R, Mill AC (2020) Using and communicating
731 uncertainty for the effective control of invasive non-native species. *Mammal*
732 *Review* 50: in press.
733 Williams KE, Huyvaert KP, Vercauteren KC, Davis AJ, Piaggio AJ (2018) Detection
734 and persistence of environmental DNA from an invasive, terrestrial mammal.
735 *Ecology and Evolution* 8: 688–695.
736 Zeale MRK, Butlin RK, Barker GLA, Lees DC, Jones G (2011) Taxon-specific PCR
737 for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resources*
738 11: 236–244.
739 Zinger L, Bonin A, Alsos IG, Bálint M, Bik H, Boyer F et al. (2019) DNA
740 metabarcoding—need for robust experimental designs to draw sound ecological
741 conclusions. *Molecular Ecology* 28: 1857–1862.
742
743

Glossary

CRISPR-Cas9
A targeted genome-editing tool comprised of the programmable Cas9 endonuclease, which introduces double-strand breaks into DNA; and a guide RNA, which targets the Cas9 nuclease to a specific DNA sequence. This allows for a portion of a target organism's genome to be modified by adding, removing or altering a DNA sequence.

DNA barcode
A DNA barcode is a standardised fragment of the genome that can be used to identify a species. Cytochrome c oxidase I was traditionally the mtDNA marker of choice in barcoding studies. The region is highly conserved throughout the animal kingdom but is variable enough to differentiate between species (Hebert et al. 2003).

DNA metabarcoding
The use of universal primers to amplify multiple DNA barcodes from bulk samples containing multiple species, such as stomach contents or environmental samples.

Environmental DNA (eDNA)
Extra-organismal DNA molecules that are shed in the environment. In animals, eDNA can originate from skin, mucous, saliva, sperm, secretions, eggs, faeces, urine and blood. eDNA can be used to detect the presence of species from samples of soil, water, or other substances from the environment.

Microsatellite
Microsatellites are regions of nuclear DNA which have tandemly repeated regions. These tandem repeats are generally 2–6 base pairs in length and have a very high mutation rate. The variation of microsatellites between individuals and populations can be used to determine population demographics such as gene flow, relatedness and genetic diversity.

Mitochondrial DNA
Mitochondrial DNA (mtDNA), found in the mitochondria as opposed to in the nucleus, has a number of favourable properties for phylogeographic and phylogenetic studies, such as the absence of recombination (which results in an effectively clonal inheritance from the maternal side) and a lack of both pseudogenes and repetitive DNA. mtDNA tends to accumulate base pair substitutions at a higher rate than nuclear DNA.

Next-generation sequencing
Next-generation sequencing, also known as high-throughput sequencing, is a broad term used to describe a number of different modern sequencing technologies. A large number of sequences (millions to billions of sequence reads) are generated on a single sequencing run.

PCR
The **Polymerase Chain Reaction (PCR)** is the exponential amplification (i.e. makes thousands of copies) of a specifically targeted region of DNA through repeated

heating and cooling cycles. It is an essential component in most genetic methodologies as more copies of the region provides a stronger signal for downstream analysis such as sequencing. Primers are required to target the region of interest and can be designed to be species-specific or to work on a broad range of species.

qPCR

qPCR is a process by which the DNA fragment is amplified like in normal PCR, but the amplification rate of the DNA fragment is continuously monitored using fluorescent light. The starting amount of DNA can then be quantified against a set of known standards. Droplet digital PCR (**ddPCR**) does not monitor the amplification process, but it can accurately quantify the starting amount of DNA without the necessity for standards.

Reduced representation sequencing

In reduced representation sequencing, restriction enzymes are used to cut (digest) the genome at specific cut sites, defined by a specific sequence of nucleotides. Sequencing and clustering of these DNA fragments allows the *de novo* discovery of SNPs. Variations of this method of sequencing include Restriction-site Associated DNA sequencing (RAD-seq), double digest RAD sequencing (ddRAD) and genotyping-by-sequencing (GBS). See Andrews et al. (2016) for a detailed review.

Reference databases

Generated DNA sequences and barcodes need to be compared to existing sequences that have been identified as belonging to a species (or at least as a genus, depending on the taxonomic group) by an expert. Reference databases provide public access to such sequences. Examples include Genbank, the Barcode of Life Database and the CDC Bartonella Laboratory database. Sequences in reference databases should have been subjected to quality control for taxonomic accuracy, but this is not always the case (particularly for older records).

Sanger sequencing

A region of DNA is copied using a fluorescent dye unique to each nucleotide. The colours read by the machine can determine the sequence of nucleotides in the region. Sanger sequencing is a low throughput method suited to sequencing long strands (~1000 base pairs) of a single region of DNA.

Single Nucleotide Polymorphism (SNP)

These are single base pair changes/variations (polymorphisms) spanning across hundreds to thousands of locations (loci) along the genome. Deciphering patterns of these changes between multiple individuals can be used to determine population demographics such as gene flow and levels of inbreeding. They can provide higher resolution information compared to other genetic markers such as microsatellites. In addition, they can be used to identify signatures of selection/adaptation in populations.