JZAR JOURNAL OF ZOO AND AQUARIUM RESEARCH
www.jzar.org

Review article

The use and integration of molecular DNA information in conservation breeding programmes: a review

Elmar Sean Fienieg[1,*] and Peter Galbusera[2]

[1]*Utrecht University, The Netherlands*
[2]*Centre for Research and Conservation, Antwerp, Belgium*
[*]*Correspondence: Elmar Sean Fienieg, Jagerakker 3, 1541VH Koog aan de Zaan, The Netherlands; elmarfienieg@hotmail.com*

Keywords:
molecular information, population management, captive breeding, relatedness estimators

Article history:
Received: 27 March 2013
Accepted: 12 July 2013
Published online: 31 October 2013

Abstract

Conservation breeding programmes often rely on intensive genetic management of the captive population. However, the relatedness between individuals and individual mean kinship are often estimated based on pedigree records, which are frequently incomplete or unreliable. Depending on the quality of a studbook (e.g. expressed as percentage of pedigree known), molecular information can substantially improve knowledge of a population, and therefore contribute to improving the retention of genetic diversity in each generation. As the use of molecular data has been largely under-utilised, this review aims to provide information on the various types of genetic markers that can be used, the estimation of (DNA based) relatedness and pedigrees, their integration in studbooks, the use of molecular information in breeding pair selection, hybridisation issues and population management in general. We discuss recent developments in methodology (e.g. next generation sequencing), theoretical considerations, and software that can aid conservation breeders in each phase of the programme from the founding phase to the (potential) reintroduction, each clarified by various examples from recent literature. Special attention is given to group-managed populations, for which it is difficult to control mating and reconstruct pedigrees as individuals cannot be isolated for management.

Conservation breeding programmes

An increasing number of species are in danger of extinction because of loss of natural habitat, poaching or other, mainly anthropogenic, impacts (Frankham et al. 2010). Largely through conservation breeding programmes, zoos and aquaria have contributed to conservation by reintroduction to the wild, research, fund raising and raising public awareness (Frankham et al. 2010; Lacy 2012). A successful example is the captive European bison (*Bison bonasus*), at one time extinct in the wild, whose captive population grew from seven to 1800 individuals, of which many were successfully reintroduced (Tokarska et al. 2009).

For reintroduction to be successful, the wild population should become self-sustaining in the long term (Frankham et al. 2010). A sustainable population requires sufficient genetic diversity; less genetically diverse populations may suffer from inbreeding depression and reduced ability to adapt (Allendorf et al. 2010; Witzenberger and Hochkirch 2011). Since a decrease in population fitness can lead to reduced genetic diversity via demographic instability, this interaction can force a population into a downward spiral towards extinction (Frankham et al. 2010).

Unfortunately, captive conservation populations typically suffer from two limitations: they are small, and they are descended from few founders (Leberg and Firmin 2008), meaning there is relatively little genetic diversity to start with. Part of this diversity will be lost each generation because of random genetic drift, which is the dominant genetic process in small isolated populations.

Another problem is that the evolutionary force of selection can act on different traits and in different directions in captivity than it does in a species' natural habitat; e.g. individuals carrying alleles that confer more 'docile' behaviour may have relatively high fitness in captivity but low fitness in the wild. Adaption to captivity has already been observed after one generation within a steelhead (*Oncorhynchus mykiss*) hatchery; the most successful individuals in captivity were the least successful in the wild (Christie et al. 2012).

To be sustainable and meet conservation goals, captive populations require genetic management. An important part of this management is the breeding phase. Generally, this consists of selecting breeding pairs and exchanging animals with other institutions. Traditionally, management is based on studbooks. Pedigrees recorded in studbooks are used to calculate the relatedness between individuals and from these values each

Box 1. Genetic markers

Genetic markers are used to measure allelic variation at a given locus (Selkoe and Toonen 2006). Since funds for conservation programmes are very limited, it is desirable to keep analysis costs low (Bömcke et al. 2011; Witzenberger & Hochkirch, 2011). This can be done by selecting the appropriate number of markers and individuals to answer specific research questions; software such as SPOTG can aid in this process (Hoban et al. 2013a). More statistical power to estimate, for example, relatedness, is obtained when markers are used that measure more variable loci, that are evenly divided over the genome and not in linkage disequilibrium (Bomcke and Gengler 2009). Linkage disequilibrium is the process whereby two loci are more likely to be transmitted to offspring as a pair than other loci; for instance, with complete linkage they are just as informative as one locus (Selkoe and Toonen 2006).

The most commonly used markers measure microsatellites: tandem repeats of short DNA sequences found on the non-coding region of the genome. Alleles differ in size; repeat sequences are easily inserted or deleted due to 'slippage' in DNA replication causing microsatellite loci to be highly variable in length. Benefits are that the development of specific primers is relatively easy (Schoebel et al. 2013) and they provide high statistical power per locus (Witzenberger and Hochkirch 2011).

A single nucleotide polymorphism (SNP) is a single base-pair mutation (e.g. C→T) that can be detected with techniques including sequencing, allele-specific PCR or SNP chips (Frankham et al. 2010). The advantages of SNPs over microsatellites are that they can more easily amplify degraded DNA (e.g. museum specimens) due to their shorter target sequence (Allendorf et al. 2010) and they are more useful in parentage analysis when dealing with highly bottlenecked/inbred species, which often have low microsatellite heterozygosity, a common situation in conservation populations (Tokarska et al. 2009). A downside is that many more SNPs are required to acquire the same power as microsatellites.

Analysis of SNPs through chips is relatively inexpensive compared to microsatellites (Tokarska et al. 2009). However, equipment costs are high and development is more costly (Allendorf et al. 2010). Therefore, SNP chips are currently more attractive in species for which genotyping systems already exist but will otherwise take more effort than microsatellites (Frankham et al. 2010).

Next generation sequencing (NGS) allows sequencing of large sections of DNA, up to the entire genome. It may be possible to sequence the entire genome of a population within a reasonable time and budget in the near future (Allendorf et al. 2011). A weakness of NGS is that it is computationally very demanding due to the huge amount of data that is collected (Allendorf et al. 2010). For directly estimating relatedness, the added value of NGS is questionable, since the total sequence of DNA is no more informative than the number of non-linked SNPs it contains; a large number of markers can achieve the same accuracy (Jones and Wang 2010a). For the development of markers, though, it is very useful (Schoebel et al. 2013).

individual's mean kinship (MK) is calculated. MK is defined as the average coefficient of kinship in the population (Ballou & Foose 1995). Individuals are prioritized for breeding based on low MK; in this way founder representation is equalised and the risk of losing unique alleles due to genetic drift is minimised. Simultaneously, this strategy minimises the average increase in inbreeding.

A second type of information is becoming increasingly more accessible: molecular (DNA) information (Abdelkrim et al. 2009). Essentially, molecular markers (Box 1) can measure the specific alleles carried by an entity. This entity can be a chromosome, individual, population, species, etc. This allows for more subtle management through genetic comparisons. For instance, it can assess the relatedness of two individuals compared to the rest of the population. This review will focus on the use of molecular information within the framework of a conservation breeding programme, with emphasis on producing breeding recommendations. Insights gained from modern methods in commercial animal breeding are also discussed.

Estimating relatedness and improving/reconstructing pedigrees

Molecular data can be used to determine the identities of animals with unknown ancestry and correct errors in the studbook (Witzenberger and Hochkirch 2011). First of all, molecular data should be checked for genotype-errors (see supplementary material; www.jzar.org) and then compared with the known pedigree to determine inconsistencies. The program GENOTYPECHECKER assigns probabilities to these inconsistencies, allowing the user to decide whether this is caused by genotyping errors or incorrectly recorded pedigrees (Paterson and Law 2011). For example, analysis of 13 captive Parma wallabies (*Macropus parma*) shows that two dams had been wrongly recorded (Ivy et al. 2009).

Second, molecular data can be used to determine the relationships of individuals with unknown ancestry. Note that there is a difference between the relatedness of a pair and their relationship; a relationship is categorical (e.g. full siblings), while a relatedness value is continuous (e.g. 0.25). Of the multiple definitions of relatedness used in the literature, MKs are normally based on the coefficient of kinship: the probability that an allele chosen randomly in one individual is identical to the allele chosen randomly in another individual, and that these alleles are identical by descent (Malécot 1948). It is important to know which definition is used to prevent comparing the coefficient of kinship with, for instance, Wright's (1922) coefficient, which is twice as high.

Which relatedness estimator will perform best will differ according to the situation because estimator quality depends on many factors. Examples are the quality of genetic and other information, the actual relatedness between individuals and population structure (Blouin 2003; Wang 2011). The best performing estimator is said to be the one with the lowest sampling variance for the greatest number of relationship categories. For instance, all sibling relationships should be assigned similar relatedness values. The program COANCESTRY is able to select the best performing estimator for each data set based on simulations (Wang 2011). When estimating relatedness from molecular data, a challenge is to discriminate between alleles that are identical by descent (IBD) and those that are identical by state (IBS). IBS alleles will be identical because of homologous mutations or because there is no allelic variation on that locus. When IBS alleles are not corrected for, a positive value of relatedness will be produced whenever a pair contains one identical allele (Jones et al. 2002).

Relatedness estimators attempt to adjust for the alleles that are IBS and weigh alleles for the information they provide by using the available information on allele frequencies in the population (Oliehoek et al. 2006). On the extreme side, a locus is not used to estimate relatedness when no variation has been observed on it.

When relatedness is based exclusively on the allelic resemblance of two individuals, corrected for the known allele frequencies in the population, it is called a moment estimator. From this information the most likely relationship category for a pair (e.g. half-siblings) can be determined and with sufficient data the pedigree can be reconstructed. Note that estimates of relatedness/relationship and inbreeding are made through comparisons relative to others. Ideally these are guaranteed unrelated individuals (Witzenberger and Hochkirch 2011). Otherwise, these estimates are too optimistic in bottlenecked populations or inaccurate when there is little variation in the population (Henkel et al. 2012; Santure et al. 2010; see also limitations below under 'Breeding based on molecular information alone').

Often there is more information available than molecular data alone, such as putative relationships. Tests are developed to compare and combine these, and other available data such as social and geographical distance, with molecular data to determine the likelihood of a certain relationship category. These are known as maximum likelihood (ML) tests (Bink et al. 2008). By limiting the possible relations with, for example, incompatible ages, the statistical power of these tests increases (Ford et al. 2011). The statistical power of ML tests may also be increased by fitting the most logical pedigree for multiple individuals at once, preventing conflicting relationships. For multiple parenthoods this may be done using CERVUS (Kalinowski et al. 2007). The likelihood of an entire pedigree can also be determined as implemented in the software COLONY (Jones and Wang 2010b; Wang and Santure 2009). The power of likelihood tests increases in populations that are more related and contain more known or excluded relationships (Wang and Santure 2009).

In the Parma wallaby population discussed previously, the unknown pedigree of seven individuals was resolved using markers. Even for two individuals whose unknown ancestry remained unresolved, relatedness analysis provided useful information; they appeared to be related at the full sibling level, indicating that these lines will produce inbred pairings and are not genetically unique (Ivy et al. 2009).

In most pedigrees founder relationships will be missing. This often results in the assumption that founders are unrelated. Instead, molecular markers can be used to determine founder relatedness in retrospect by molecular analysis of their offspring (Ivy et al. 2009). Of course, better results are obtained if high quality DNA is still available from living founders or from preserved specimens in museums or zoo archives (recommended). The assumption of zero founder relatedness can be replaced by a relatedness estimate of one estimator or by developing a hypothetical pedigree. For the latter, the software MOL COANC is useful (Fernández and Toro 2006).

An interesting, robust method has been proposed for the breeding management of a highly inbred population of Mississippi sandhill cranes (*Grus canadensis pulla*) that uses multiple estimators of founder relatedness (Henkel et al. 2011). In this method, the studbook is first corrected for gaps and errors using DNA analysis. Second, three different studbooks are made in which only the values of founder relatedness differ: (1) founders are assumed unrelated; (2) founder relatedness is based on allelic similarity, which is the uncalibrated proportion of alleles shared; (3) founder relatedness is estimated with G&Q's moment-estimator (Goodnight & Queller 2002). From these studbooks, three different values of MK can also be produced for each individual. Breeding priority is then given to individuals for which all three MK values are lower than the average MK calculated for its respective studbook because all three values provide different information; zero founder relatedness MK gives a qualitative estimate of the part of the genome not measured by molecular markers, while allelic similarity MK indicates how rare the alleles carried by an individual are and breeding pairs are made based on similar Q&G's MK.

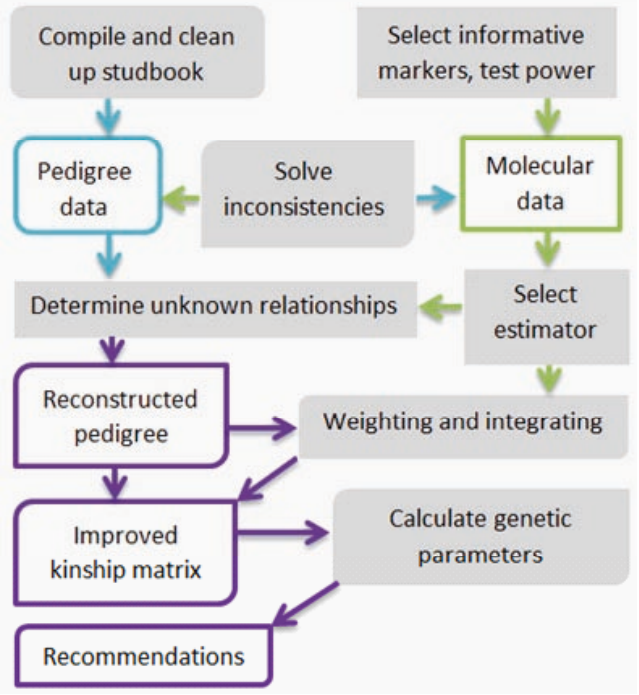

Figure 1. Pedigree reconstruction and integration with molecular data. Flowchart summarising each step from preparatory phase (top) to breeding recommendations.

Because the default assumption in breeding programmes is zero founder relatedness, the effect of knowing this value increases with increasing founder relatedness. When relatedness is very low, it will not affect management at all (Rudnick and Lacy 2008). A simulation of the captive Parma wallaby population shows that implementing the (low) founder relatedness in management hardly influences the genetic diversity maintained over 100 years (Ivy et al. 2009).

Breeding based on molecular information alone

A relatively complete pedigree may give better estimates of relatedness than molecular analysis using up to 100 microsatellite markers, even without taking into account biases caused by genotyping errors, mutations and sampling errors for allele frequencies (Baumung and Solkner 2003; Fernández et al. 2005). However, with the development of SNP-chips and NGS (see Box 1), the accuracy of molecular data has greatly improved (Allendorf et al. 2010). Furthermore, while pedigree-based relatedness assumes no variation in inheritance, in reality there is such variation due to random inheritance and linkage (Engelsma et al. 2011). For example, the relatedness between full siblings in a zebra-finch (*Taeniopygia guttata*) population, analysed with a genome-wide array of markers, had a standard error of 20% (Santure et al. 2010). In other words, for determining the relationship (e.g. full or half sibling) pairwise molecular data will often be inaccurate, but for determining relatedness (e.g. 27% of alleles identical by descent) molecular data can give more accuracy than pedigrees can achieve, on condition that the molecular data is sufficiently comprehensive (de Cara et al. 2011). Based on more accurate relatedness values, a greater amount of genetic diversity can theoretically be maintained (de Cara et al. 2011; Henkel et al. 2011). Whether marker-based management can also achieve lower

levels of inbreeding (depression), though, is still debated (de Cara et al. 2011; Santure et al. 2010; Townsend and Jamieson 2013). Molecular analysis allows for tracking specific alleles in the population and so can be used to prevent unique alleles from being lost from the population (de Cara et al. 2011; Jones et al. 2002). There is a risk associated with blindly equalising allele frequencies, however. Some alleles will be rare because they are deleterious and for some alleles the population's fitness will be optimal at a non-equal frequency (Charlesworth and Charlesworth 2012; Ivy and Lacy 2010). The chances of success of a reintroduced population will be highest with certain allele frequencies best adapted to that environment, such as that of the historical population (Miller et al. 2010). Molecular analysis and the method of Saura et al. (2008) allows the determination of the offspring each individual should leave to the next generation in order to maintain allele frequency distribution at each locus as close as possible to a certain distribution, while simultaneously maintaining acceptable levels of genetic diversity. However, even if historical frequencies can be retrieved, this method is hazardous: genetics in endangered populations are usually more subjected to genetic drift than selection and the habitat of a species is likely to have changed since their capture (Frankham et al. 2010). The safest strategy may therefore simply be to equalise allele-frequencies and so minimise the chance of losing alleles (Miller et al. 2010).

Integrating pedigree and molecular information

Pedigree and molecular relatedness are different types of information. Pedigrees give theoretical information on relatedness over the entire genome while molecular data give empirical evidence on specific parts of the genome. In other words, they give values of relatedness for different parts of the genome and information is lost when both sources are available but only one is used. When combining these two values it is important that the same definition of relatedness is used (preferably the coefficient of kinship; Malécot 1948). Also, if a combined value of pedigree and molecular relatedness is calculated as the average of the two, information will be lost due to unequal accuracy of the two coefficients. Instead, the two values can be weighted for the information that they provide (Bömcke and Gengler 2009; Bömcke et al. 2011; Fernández et al. 2012).

A relatively simple method is described by Fernández et al. (2012); pedigrees and molecular information are given relative weights of 10 and 1 respectively. The added power of molecular markers in this study is used to discriminate between two equally related mates, based on the pedigree (e.g. full siblings), of which one has a lower degree of molecular relatedness. The weight attributed to markers and pedigrees, however, can be made more fit to the situation; Bömcke and Gengler's (2009) method weights pedigree data by the depth of the pedigree (number of known generations and missing parentage in these generations), and weights molecular data by the number of unlinked markers and their polymorphism information content (PIC; useful software: CERVUS, Kalinowski et al. 2007). The PIC value is based on the number of polymorphisms of a certain marker and their distribution in the population. After weighting, the average of the molecular and pedigree coefficients is then the relatedness between two individuals and breeding pairs can be selected based on these values.

Bömcke et al. (2011) use a different approach: molecular markers only give relatedness information on the parts of the genome in linkage disequilibrium with the markers used, therefore molecular data is given a weight that is relative to the proportion of the genome that the markers measure. This proportion is determined through simulation as the power of a set of molecular markers to predict the pedigree. For the unmeasured part of the genome, (incomplete) pedigree data can be used to determine its theoretical relatedness. Again, simulations can be used to determine the power of an incomplete pedigree to predict the actual pedigree. These measures can then be combined into one value using these two weights. A method to fit in individuals for whom only pedigree data is available is also provided. When using these methods it is important to be aware that estimates will still in practice be based on one parameter if the quality of either pedigree or molecular data is relatively low (Bömcke et al. 2011; Bömcke and Gengler 2009). For new integration methods, it may be useful to weigh the relative value of molecular information using software KININFOR (Wang 2006), which uses four measurements of information value for each molecular marker. In addition, this software enables the statistical power of a relatedness analysis to be tested for. The program PMx may be used to determine population genetic parameters, and so produce breeding recommendations (Lacy et al. 2012). It provides the option to combine the pedigree-based relatedness matrix with a molecular (empirical) based matrix, using a chosen weight. A new relatedness matrix can also be inserted directly. For a summary of the steps described above for the use of molecular data for breeding recommendations, see Figure 1.

The methods proposed by Bömcke and Gengler (2009), Bömcke et al. (2011) and Fernández et al. (2012) for integrating molecular data with studbook data are superior over methods that simply reconstruct the pedigree because they use more of the available information. Nevertheless, further improvement may be possible in the future; management methods minimising the population's MK do not discriminate between adaptive variation and neutral variation (Marsden et al. 2013). Traditional molecular markers (e.g. microsatellites) measure neutral DNA directly, but can also measure coding DNA indirectly due to linkage disequilibrium. Novel methods such as NGS (see Box 1) make it possible to measure coding DNA specifically, which opens doors for research and possibly also for maintaining diversity at specific adaptive parts of the genome (Allendorf et al. 2010; Engelsma et al. 2011).

It is questionable, though, whether it is useful to develop methods focused solely on maintaining adaptive diversity. Assessing which adaptive variation is important for each species is likely to require an enormous amount of general and specific research (Witzenberger and Hochkirch 2011). Furthermore, if adaptive variation is evenly spread over the genome, it is much simpler to manage on MK. Nonetheless, there is evidence that genetic diversity is unevenly divided over the genome, so that it may be useful to use a strategy that conserves diversity at specific important regions where it is disappearing (Engelsma et al. 2012). This can be important for loci where variation has major adaptive effects. For example, individuals heterozygous for the major histocompatibility complex (MHC) genes are less susceptible to diseases than homozygous individuals (Hughes 1991). In addition, pairs with dissimilar MHC are thought to be more attracted to each other (Havlicek and Roberts 2009), which could potentially decrease the number of failed pairings in breeding programmes. Marsden et al. (2013) used molecular information on the MHC in combination with information on neutral diversity and the pedigree to evaluate and improve the breeding programme of the African wild dog (*Lycaon pictus*). Results show no significant difference between the neutral and MHC diversity maintained, indicating that the MHC is not under selection in captivity. The authors conclude that MK-based management performs well in the maintenance of MHC diversity.

Comparable to the identification of adaptive variation is the identification of deleterious alleles causing diseases (Maher et al. 2012). This information can be used for purging deleterious alleles from the population: deliberately selecting related breeding pairs to create inbred offspring. Since offspring with two (recessive)

deleterious alleles are less likely to survive, this will decrease its frequency. However, the effects of purging are unpredictable and will often cause the loss of valuable genetic diversity (de Cara et al. 2013; Leberg and Firmin 2008; Boakes et al. 2007). Even when specific carriers are known, excluding these from a small population is not thought to be worth the loss of genetic diversity (Allendorf et al. 2010). In addition, much is still unclear on how multiple deleterious alleles interact (Charlesworth and Willis 2009). Hence, purging is not recommended, even if individual molecular marker data are available (Witzenberger and Hochkirch 2011).

Hybridisation

Because founders for a population are often scarce it often occurs that individuals of different populations are used as founders (Frankham et al. 2010). This introduces the risk of crossing different species and so creating hybrids. Even though hybrids are undesirable, managers should not be overly cautious about including individuals in their breeding programme. This is discussed further in the section 'Conserving the species' below. In case of doubt, molecular analysis can assist in determining if founders are of the same species. In this way, for example, a Nile crocodile (*Crocodylus niloticus*) was discovered within the Philippine crocodile (*C. mindorensis*) breeding programme (Hauswaldt et al. 2013).

When a number of individuals in a population carries unwanted hybrid (exogenous) alleles, simply removing them from the breeding programme is not recommended because this will usually cause an undesirable amount of native diversity to be lost as well (Grobler et al. 2011). Instead, breeding with hybrids can be continued while artificially selecting against hybrid (exogenous) alleles. Native individuals can be given relatively higher breeding priority, based on either studbook or molecular information (Amador et al. 2011, 2012). In this method, markers outperform pedigrees if ten or more informative alleles are used that are private/diagnostic for the conserved species, or if more than 20 alleles are used with a much higher frequency in the conserved species (Amador et al. 2012).

To identify hybrids, genetic data must be available on guaranteed pure individuals. In populations with widespread hybridisation this can cause problems. Complementing genetic data with morphological data can be a solution, as currently done for the black wildebeest (*Connochaetes gnou*). Its wild population seems to mainly consist of hybrids with the blue wildebeest (*C. taurinus*) due to mismanagement in the past (Grobler et al. 2011). Breeding to obtain certain morphology can also be a conservation tool in some cases e.g. breeding-back of the auroch (*Bos primigenius*, van Vuure 2005). Techniques developed for livestock breeding can then be of use. Based on a combination of molecular and pedigree information, Fernández et al. (2012) achieved a 43% increase in frequency of a specific trait of interest while only losing 4% genetic diversity on the rest of the genome in an Iberian pig (*Sus scrofa domesticus*) Dorado strain. Any selection procedure in a conservation programme, however, needs careful consideration since levels of genetic diversity are usually already very low.

Group management

Optimal genetic management requires the availability of individual pedigrees, the ability to breed all individuals and control over breeding pairs. In reality, however, this is often not the case. A common restriction is that a captive population is held on several continents and as a result is managed as multiple sub-populations with limited migration. The programs PMx and METAPOP are able to manage these sub-populations separately, while aiming to maintain genetic diversity for the entire population (Lacy et al. 2012; Pérez-Figueroa et al. 2008). In the absence of shared pedigrees, molecular data can be used to determine the kinship between sub-populations.

For species held as multi-male, multi-female groups, breeding pairs cannot be controlled and deducing pedigrees is not feasible (e.g. shoals of fish). As a result, these populations are managed at the group level, instead of the individual level. Genetic management for these populations consists of artificially exchanging individuals between groups. The genetics of a group change each generation and this change is not visible to the human eye. Fortunately, molecular analysis makes it possible to determine the actual loss of diversity by genetic drift (expressed as the effective population size), identify selective forces and estimate the reproductive success of immigrants (Leus et al. 2011; Hasler et al. 2011; Wang 2004). In this way, relatedness between groups can be determined and individuals can be exchanged at random or, through the use of detailed tests, can even be selected on their genetic suitability (McGreevy et al. 2010; Miller et al. 2010).

When groups are large and fecundity is high, founder representation can be equalised by selecting a relatively small part of the population based on molecular kinship to produce the next generation. Software that can be used includes GENCONT and EVA (Berg & Nielsen 2006; Meeuwissen 2002).

When continuous molecular analysis in each generation is too costly, genetic research can facilitate the use of low-intensity management through models such as Wang's (2004) migration model, by providing knowledge of breeding behaviour and subpopulation structure (Smith 2010).

Conserving the species

At the time of reintroduction, a conservation breeding population ideally represents 95% of the natural gene-pool and is self-sustaining (Frankham 2009; Miller et al. 2010).

A breeding programme will not reach these goals if the founder population was inadequate in the first place. Miller et al.'s (2010) method uses molecular data to select a group of founders that are representative for a (sub)-species, including all its sub-populations. However, analyses of a large number of wild individuals are then required, which will often be unrealistic. It will also often be unrealistic and unnecessary to set up breeding programmes for a large number of subspecies (Frankham 2009; Hedrick and Fredrickson 2009). Instead, conservation actions may best be aimed at preserving a management unit that is both feasible and has the greatest conservation impact (Frankham 2009). If the choice is made to conserve only one sub-species, a combination of genetic and demographic data can be used to determine the extinction risk of subspecies and prioritise the need for a captive breeding programme (FAO 2010; note: developed for breeds of livestock).

If a choice is made to cross multiple populations, the risk of a depression of fitness (outbreeding depression) can be estimated from molecular and ecological information (see the decision tree of Frankham et al. 2011). It is important to note that significant genetic differentiation does not necessarily cause outbreeding depression and that the risk of outbreeding depression will often be low compared to the risk of inbreeding depression in endangered species (Frankham et al. 2011; FAO 2010).

Programme coordinators may be convinced that their programme has started with a large enough proportion of a species' genetic diversity and that their programme is adequately maintaining it. However, without qualitative evaluation of both the wild and captive population, this conclusion cannot be made (Witzenberger and Hochkirch 2011). While pedigrees are often missing for wild populations, samples for DNA analysis can easily

be extracted from hair and faeces (Oliveira and Gaiotto 2011). This has already led to a large number of captive breeding programme evaluations (Gonçalves da Silva et al. 2010; McGreevy et al. 2010; Shen et al. 2009; Tzika et al. 2008). In some studies this did not lead to any concerns (Gonçalves da Silva et al. 2010), but in others it became apparent that diversity levels were dangerously low and additional founders from the wild were required; three out of five groups of yellow-breasted capuchin monkeys (*Cebus xanthosternos*) had dangerously low diversity levels (Oliveira and Gaiotto 2011).

On the other hand, molecular analysis can also reveal a need for reintroduction due to low genetic diversity of the wild population. A way to improve the genetic health of both the wild and captive population is an 'open' system with continuous exchange between the two (Lacy 2012). Such an open system is used for the giant panda (*Ailuropoda melanoleuca*), where molecular data showed that both the captive and wild population were at risk without genetic exchange (Shen et al. 2009). In this case, molecular analysis may be used to construct a pedigree covering both the wild and captive population and individuals can be selected for exchange that optimises genetic diversity of the species as a whole (Allendorf et al. 2010).

DNA analysis in zoos and aquaria has so far been restricted by financial requirements and lack of expertise. Fortunately, these techniques are becoming increasingly simple and exponentially more affordable (Allendorf et al. 2010). Zoos and aquaria currently underestimate the interest of research institutes in collaboration. Geneticists can benefit from zoo studies through overlapping research questions and publications. Their research is often restricted by the number of DNA samples, while zoos and aquaria can provide these with relative ease, possibly combined with historical quantitative and medical data. Communication between breeding programme managers and geneticists is not only important in discovering opportunities on both sides, but also in preventing unforeseen restrictions after analysis (Hoban et al. 2013b). General guidance and case studies for the use of molecular markers in conservation breeding programs and related topics can be found on www.ConGRESSgenetics.eu and a summary of the methods and software described in this review is available as supplementary material (www.jzar.org).

Molecular information has huge conservation potential, but instead of blindly replacing all other information, it should be cleverly integrated into population management. This will facilitate intelligent and informed decisions, which can save money, resources and, most importantly, species.

Acknowledgements

This research was partly funded by the Flemish Government as part of the structural funding for the Centre for research and Conservation. Kaj van Tienderen is kindly acknowledged for designing the box with marker information and Sean Hoban, Marie José Duchateau, Jill Shephard, Sharron Ogle and two anonymous referees for helpful comments.

References

Abdelkrim J., Robertson B., Stanton J., Gemmell N. (2009) Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *BioTechniques* 46: 185–192.

Allendorf F.W., Hohenlohe P.A., Luikart G. (2010) Genomics and the future of conservation genetics. Nature reviews. *Genetics* 11: 697–709.

Amador C., Toro M.A., Fernández T. (2011) Removing exogenous information using pedigree data. *Conservation Genetics* 12: 1565–1573.

Amador C., Toro M.A., Fernández T. (2012) Molecular markers allow to remove introgressed genetic background: a simulation study. *PLoS ONE* 7(11): e49409. doi:10.1371/journal.pone.0049409.

Ballou J.D., Foose T.J. (1996) Demographic and genetic management of captive populations. In: Kleiman D.G., Allen M.E., Thompson K.V., Lumpkin S. (eds) *Wild Animals in Captivity: Principles and Techniques*. Chicago: University of Chicago Press, 263–283.

Baumung R., Solkner J. (2003) Pedigree and marker information requirements to monitor genetic variability. *Genetics Selection Evolution* 35: 369–383.

Berg P., Nielsen J., Sørensen, M.K. (2006) EVA: realized and predicted optimal genetic contributions. In: *Proceeding of the 8thWorld Congress on Genetics Applied to Livestock Production*, CD communication 27-09.

Bink M.C.A.M., Anderson A.D., van de Weg E., Thompson E.A. (2008) Comparison of marker-based pairwise relatedness estimators on a pedigreed plant population. *Theoretische und Angewandte Genetik* 117: 843–855.

Blouin M.S. (2003) DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends in Ecology and Evolution* 18: 503–511.

Boakes E.H., Wang J., Amos W. (2007) An investigation of inbreeding depression and purging in captive pedigreed populations. *Heredity* 98: 172–182.

Boettcher P., Gandini G., Fernández J., Joshi B.K., Oldenbroek K., Sponenberg P. (2010) *FAO Guidelines for the In Vivo Conservation of Animal Genetic Resources.* Roma: Food & Agriculture Organization of the United Nations.

Bömcke E., Gengler N. (2009) Combining microsatellite and pedigree data to estimate relationships among Skyros ponies. *Journal of Applied Genetics* 50: 133–143.

Bömcke E., Soyeurt H., Szydlowski M., Gengler N. (2011) New method to combine molecular and pedigree relationships. *Journal of Animal Science* 89: 972–978.

de Cara M.A.R., Fernández J., Toro M.A., Villanueva B. (2011) Using genome-wide information to minimize the loss of diversity in conservation programmes. *Journal of Animal Breeding and Genetics [Zeitschrift für Tierzüchtung und Züchtungsbiologie]* 128: 456–464.

de Cara M.A.R., Villanueva B., Toro M.A., Fernández J. (2013) Purging deleterious mutations in conservation programmes: combining optimal contributions with inbred matings. *Heredity* 110: 530–537.

Charlesworth D., Willis J.H. (2009) The genetics of inbreeding depression. *Genetics* 10: 783–96.

Charlesworth B., Charlesworth D. (2012) *Elements of Evolutionary Genetics.* Greenwood Village, Colorado: Roberts and Company.

Christie M.R., Marine M.L., French R.A., Blouin M.S. (2012) Genetic adaptation to captivity can occur in a single generation. *Proceedings of the National Academy of Sciences of the United States of America* 109: 238–242.

Engelsma K.A., Veerkamp R.F., Calus M.P.L., Windig J.J. (2011) Consequences for diversity when prioritizing animals for conservation with pedigree or genomic information. *Journal of Animal Breeding and Genetics [Zeitschrift für Tierzüchtung und Züchtungsbiologie]* 128: 473–481.

Engelsma K.A., Veerkamp R.F., Calus M.P.L., Bijma P., Windig J.J. (2012) Pedigree- and marker-based methods in the estimation of genetic diversity in small groups of Holstein cattle. *Journal of Animal Breeding and Genetics [Zeitschrift für Tierzüchtung und Züchtungsbiologie]* 129: 195–205.

Fernández J., Toro M.A. (2006) A new method to estimate relatedness from molecular markers. *Molecular Ecology* 15: 1657–1667.

Fernández J., Clemente I., Amador C., Membrillo A., Azor P., Molina A. (2012) Use of different sources of information for the recovery and genetic management of endangered populations: example with the extreme case of Iberian pig Dorado strain. *Livestock Science* 149: 282–288.

Fernández J., B Villanueva, Pong-Wong R., Toro M.A. (2005) Efficiency of the use of pedigree and molecular marker information in conservation programs. *Genetics* 170: 1313–1321.

Ford M.J., Hanson M.B., Hempelmann J.A., Ayres K.L., Emmons C.K., Schorr G.S., Baird R.W., Balcomb K.C., Wasser S.K., Parsons K.M., Balcomb-Bartok K. (2011) Inferred paternity and male reproductive success in a killer whale (*Orcinus orca*) population. *Journal of Heredity* 102: 537–553.

Frankham R. (2009) Where are we in conservation genetics and where do we need to go? *Conservation Genetics* 11: 661–663.

Frankham R., Ballou J.D., Briscoe D.A. (2010) *Biological Conservation: Introduction to Conservation Genetics*. Cambridge: Cambridge University Press.

Frankham R., Ballou J.D., Elridge M.D.B., Lacy R.C., Ralls K., Dudash M.R., Fenster C.B. (2011) Predicting the probability of outbreeding depression. *Conservation Biology* 25: 465–475.

Fraser D.J. (2008) How well can captive breeding programs conserve biodiversity? A review of salmonids. *Evolutionary Applications* 1: 535–586.

Gonçalves da Silva A., Lalonde D.R., Quse V., Shoemaker A., Russello M.A. (2010) Genetic approaches refine ex situ lowland tapir (*Tapirus terrestris*) conservation. *Journal of Heredity* 101: 581–590.

Goodnight K.F., Queller D.C. (2002) Computer software for performing likelihood tests of pedigree relationship using genetic markers. *Molecular Ecology* 8: 1231–1234.

Grobler J.P., Rushworth I., Brink J.S., Kotze A., Bloomer P., Reilly B., Vrahimis S. (2011) Management of hybridization in an endemic species : decision making in the face of imperfect information in the case of the black wildebeest – *Connochaetes gnou. European Journal of Wildlife Research* 57: 997–1006.

Hasler H., Flury C., Menet S., Haase B., Leeb T., Simianer H., Poncet P A., Rieder S. (2011) Genetic diversity in an indigenous horse breed: implications for mating strategies and the control of future inbreeding. *Journal of Animal Breeding and Genetics [Zeitschrift für Tierzüchtung und Züchtungsbiologie]* 128: 394–406.

Havlíček J., Roberts C.S. (2009) MHC-correlated mate choice in humans: a review. *Psychoneuroendocrinology* 34: 497–512.

Hauswaldt J.S., Vences M., Louis E., Brennemann R., Ziegler T. (2013) Genetic screening of captive Philippine crocodiles (*Crocodylus mindorensis*) as prerequisite for starting a conservation breeding program in Europe. *Herpetological Conservation and Biology* 8: 75–87.

Hedrick P.W., Fredrickson R. (2009) Genetic rescue guidelines with examples from Mexican wolves and Florida panthers. *Conservation Genetics* 11: 615–626.

Henkel J.R., Jones K.L., Hereford S.G., Savoie M.L., Leibo S.P., Howard J.J. (2011) Integrating microsatellite and pedigree analyses to facilitate the captive management of the endangered Mississippi sandhill crane (*Grus canadensis pulla*). *Zoo Biology* 31: 322–335.

Hoban S., Gaggiotti O., ConGRESS Consortium, Bertorelle G. (2013a) Sample planning optimization tool for conservation and population genetics (SPOTG): a software for choosing the appropriate number of markers and samples. *Methods in Ecology and Evolution* 4: 299–303.

Hoban S.M., Hauffe H.C., Pérez-Espona S., Arntzen J.W., Bertorelle G., Bryja J., Frith K., Gaggiotti O.E., Galbusera P., Godoy J.A., Hoelzel A.R., Nichols R.A., Primmer C.R., Russo I., Segelbacher G., Siegismund H.R., Sihvonen M., Vernesi C., Vilà C., Bruford M.W. (2013b) Bringing genetic diversity to the forefront of conservation policy and management. *Conservation Genetics Resources* 5: 593–598.

Hughes, A.L. (1991) MHC polymorphism and the design of captive breeding programs. *Conservation Biology* 5: 249–251.

Ivy J.A., Lacy R.C. (2010) Using molecular methods to improve the genetic management of captive breeding programs for threatened species. In: DeWoody J.A., Bickham J.W., Michler C.H., Nicols K.M., Rhodes O.E., Woeste K.E. (eds) *Molecular Approaches in Natural Resource Conservation and Management*. Cambridge: Cambridge University Press, 267–295.

Ivy, J.A., Miller A., Lacy R.C., Dewoody J.A. (2009) Methods and prospects for using molecular data in captive breeding programs: an empirical example using parma wallabies (*Macropus parma*). *Journal of Heredity* 100: 441–454.

Jones K.L., Glenn T.C., Lacy R.C., Pierce J.R., Unruh N., Mirande C.M., Chavez-Ramirez F. (2002) Refining the whooping crane studbook by incorporating microsatellite DNA and leg-banding analyses. *Conservation Biology* 16: 789–799.

Jones O.R., Wang J. (2010a) Molecular marker-based pedigrees for animal conservation biologists. *Animal Conservation* 13: 26–34.

Jones O.R., Wang J. (2010b) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* 10: 551–555.

Kalinowski S.T., Taper M.L., Marshall T. (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* 16: 1099–1106.

Lacy R.C. (2012) Achieving true sustainability of zoo populations. *Zoo Biology* 13: 1–13.

Lacy R.C, Ballou J.D., Pollak J.P. (2012) PMx: software package for demographic and genetic analysis and management of pedigreed populations. *Methods in Ecology and Evolution* 3: 433–437.

Leberg P.L., Firmin B.D.. (2008) Role of inbreeding depression and purging in captive breeding and restoration programmes. *Molecular Ecology* 17: 334–343.

Leus K., Traylor-Holzer K., Lacy R.C. (2011) Genetic and demographic population management in zoos and aquariums: recent developments, future challenges and opportunities for scientific research. *International Zoo Yearbook* 45: 213–225.

Maher M.C., Uricchio L.H., Torgerson D.G., Hernandez R.D. (2012) Population genetics of rare variants and complex diseases. *Human Heredity* 74: 118-128.

Malécot G. (1948) *Les Mathématiques de l'Hérédité*. Paris: Masson et Cie.

Marsden C.D., Verberkmoes H., Thomas R., Wayne R.K., Mable B.K. (2013) Pedigrees, MHC and microsatellites : an integrated approach for genetic management of captive African wild dogs (*Lycaon pictus*). *Conservation Genetics* 14: 171–183.

McGreevy T.J., Dabek L., Husband T.P. (2010) Genetic evaluation of the Association of Zoos and Aquariums Matschie's tree kangaroo (*Dendrolagus matschiei*) captive breeding program. *Zoo Biology* 30: 636–646.

Meeuwissen T.H.E. (2002) GENCONT: An operational tool for controlling inbreeding in selection and conservation schemes. In: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, Montpellier, France, 769–770.

Miller W., Wright S.J., Zhang Y., Schuster S.C., Hayes V.M. (2010) Optimization methods for selecting founder individuals for captive breeding or reintroduction of endangered species. *Pacific Symposium on Biocomputing* 15: 43–53.

Oliehoek P.A., Windig J.J., van Arendonk J.A.M., Bijma P. (2006) Estimating relatedness between individuals in general populations with a focus on their use in conservation programs. *Genetics* 173: 483–96.

Oliveira C.G., Gaiotto F.A. (2011) Molecular genetic analysis of the yellow-breasted capuchin monkey: recommendations for ex situ conservation. *Genetics and Molecular Research* 10: 1471–1478.

Paterson T., Law A. (2011) Genotypechecker: an interactive tool for checking the inheritance consistency of genotyped pedigrees. *Animal Genetics* 42: 560–562.

Pérez-Figueroa A., Saura M., Fernández J., Toro M.A., Caballero A. (2008) METAPOP—A software for the management and analysis of subdivided populations in conservation programs. *Conservation Genetics* 10: 1097–1099.

Rudnick J.A., Lacy R.C. (2008) The impact of assumptions about founder relationships on the effectiveness of captive breeding strategies. *Conservation Genetics* 9: 1439–1450.

Santure A.W., Stapley J., Ball A.D., Birkhead T.R., Burke T., Slate J. (2010) On the use of large marker panels to estimate inbreeding and relatedness: empirical and simulation studies of a pedigreed zebra finch population typed at 771 SNPs. *Molecular Ecology* 19: 1439–51.

Saura M., Pérez-Figueroa A., Fernández J., Toro M.A., Caballero A. (2008) Preserving population allele frequencies in ex situ conservation programs. *Conservation Biology* 22: 1277–1287.

Schoebel C.N., Brodbeck S., Buehler D., Cornejo C., Gajurel J., Hartikainen H., Keller D., Leys M., Říčanová S., Segelbacher G., Werth S., Csencsics D. (2013) Lessons learned from microsatellite development for nonmodel organisms using 454 pyrosequencing. *Journal of Evolutionary Biology* 26: 600–611.

Selkoe K.A., Toonen R.J. (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9: 615–629.

Shen F., Zhang Z., He W., Yue B., Zhang A., Zhang L., Hou R., Wang C., Watanabe T. (2009) Microsatellite variability reveals the necessity for genetic input from wild giant pandas (*Ailuropoda melanoleuca*) into the captive population. *Molecular Ecology* 18: 1061–1070.

Smith B.R. (2010) *Genetic Management of Groups.* Unpublished dissertation. College Park: University of Maryland.

Tokarska M.T., Kowalczyk M.R., Wójcik J.M., Pertoldi C., Kristensen T.N., Loeschcke V., Gregersen V.R., Bendixen C. (2009) Effectiveness of microsatellite and SNP markers for parentage and identity analysis in species with low genetic diversity: the case of European bison. *Heredity* 103: 326–332.

Townsend S.M., Jamieson I.G. (2013) Molecular and pedigree measures of relatedness provide similar estimates of inbreeding depression in a bottlenecked population. *Journal of Evolutionary Biology* 26: 889–899.

Tzika A.C., Remy C., Gibson R., Milinkovitch M.C. (2008) Molecular genetic analysis of a captive-breeding program: the vulnerable endemic Jamaican yellow boa. *Conservation Genetics* 10: 69–77.

Van Vuure C. (2005) *Retracing the Aurochs : History, Morphology and Ecology of an Extinct Wild Ox.* Sofia–Moscow: Pensoft Publishers.

Wang J. (2004) Monitoring and managing genetic variation in group breeding populations without individual pedigrees. *Conservation Genetics* 5: 813–825.

Wang J. (2006) Informativeness of genetic markers for pairwise relationship and relatedness inference. *Theoretical Population Biology* 70: 300–321.

Wang J. (2011) COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources* 11: 141–145.

Wang J., Santure A.W. (2009) Parentage and sibship inference from multilocus genotype data under polygamy. *Genetics* 181: 1579–94.

Witzenberger K.A., Hochkirch A. (2011) Ex situ conservation genetics: a review of molecular studies on the genetic consequences of captive breeding programmes for endangered animal species. *Biodiversity and Conservation* 20: 1843–1861.

Wright S. (1922) Coefficients of inbreeding and relationship. *American Naturalist* 56: 330.