

# mPed: a computer program for converting pedigree data to a format used by the PMx-software for conservation genetic analysis

Mija Jansson · Ingvar Ståhl · Linda Laikre

Received: 15 September 2012 / Accepted: 28 January 2013 / Published online: 19 February 2013  
© The Author(s) 2013. This article is published with open access at Springerlink.com

**Abstract** There is a growing need for conservation genetic management of animal populations when individual relatedness data (pedigrees) are available. Such data can be used to monitor rates of inbreeding and loss of genetic diversity. Traditionally, pedigree analysis for conservation management has focused on zoo populations of threatened wild animals; available software has been developed in that context. Population Management x (PMx) is a free software for estimating genetic parameters including inbreeding, kinship, founder allele contribution and survival. PMx is an accessory program to the zoo studbook platform Single Population Analysis and Records Keeping System (SPARKS) and is not easily applied outside this platform, but such use is of interest for various domestic breeds or wild populations. We developed a converter program (mPed) for making pedigrees of any studbook format fitting the input requirements of PMx. mPed can be downloaded free at [www.popgen.su.se/mped.php](http://www.popgen.su.se/mped.php).

**Keywords** PMx · Pedigree analysis · Conservation genetic management · Inbreeding · Genetic variation · Founder statistics

Conservation genetic management of populations with known individual relatedness (pedigrees) has traditionally focused on populations of wild animals bred in zoos (Ballou et al. 1995). For such populations, maintenance of

genetic variation is typically the main goal, and pedigree data can be used to monitor rates of inbreeding and loss of genetic variation measured in relation to alleles contributed by separate founding animals of the population (Lacy 1989). Computer programs have been developed within the zoo community to aid in genetic management by handling studbook data and performing statistical pedigree analyses.

There is an increasing need for methods, including software, developed for zoo populations to be applied to populations kept outside the zoo community. With current worldwide awareness of the need for conserving biological diversity, genetic monitoring is becoming topical for many populations in addition to those of threatened species in zoos. International policy agreements explicitly state that genetic variation of domestic animal breeds should be conserved and used sustainably (FAO 2007). For instance, within the framework of the United Nations Convention on Biological Diversity conservation targets to be reached during the period 2011–2020 were recently adopted, and one of these target focuses on the genetic diversity of farmed and domesticated animals (Decision UNEP/CBD/COP/DEC/X2; [www.cbd.int](http://www.cbd.int)). This target imply that strategies should be “developed and implemented for minimizing genetic erosion and safeguarding their genetic diversity” ([www.cbd.int/sp](http://www.cbd.int/sp)).

We have recently been working with the pedigree of a wild wolf population in Sweden generated through tracking in combination with DNA analyses (Liberg et al. 2005). This population is small and highly inbred and there is a need for application of pedigree analysis to address issues of individual kinship and rate of loss of variation (Laikre et al. 2012). Also, we are working on domestic dog pedigrees to address similar issues (Jansson and Laikre 2011). During this work the need for using already available software became apparent.

---

M. Jansson (✉) · L. Laikre  
Division of Population Genetics, Department of Zoology,  
Stockholm University, 106 91 Stockholm, Sweden  
e-mail: [mija.jansson@zoologi.su.se](mailto:mija.jansson@zoologi.su.se)

I. Ståhl  
Ingvar Ståhl IT, Jenny Linds gata 26, 129 52 Hägersten, Sweden

**Table 1** Examples of software for pedigree analysis including their limitations with respect to pedigrees that can be analyzed and the output data provided

Computer program	Limitations	Generated quantities	Web-page	Cost
ENDOG	Around 1,000,000 individuals	$F$ , $N_e^a$ , founder-representation and F-statistics	<a href="http://www.ucm.es/info/prodanim/html/JP_Web.htm">http://www.ucm.es/info/prodanim/html/JP_Web.htm</a>	Free
Pedigree Viewer	Around 50,000 individuals	Draws pedigree, $F$	<a href="http://personal-une.edu.au/~bkinghor/pedigree.htm">http://personal-une.edu.au/~bkinghor/pedigree.htm</a>	Free
Lathunden ( <i>created in Swedish for the Swedish kennel club</i> )	Handle the Swedish kennel club's data base in number and format	$F$ for four generations, kinship, health, mentality, competitions etc.	<a href="http://www.genetica.se/">http://www.genetica.se/</a>	(1275 SEK) US\$191
Pedigree Explorer (PedX)	100 generations	Draws pedigree, $F$ , kinship	<a href="http://www.breedmate.com/">http://www.breedmate.com/</a>	US\$80
Population Management x (PMx)	Around 20,000 individuals	$F$ , $MK$ , $N_e^b$ , founder-representation and contribution, founder genome equivalents etc.	<a href="http://www.vortex9.org/pm2000.html">http://www.vortex9.org/pm2000.html</a>	Free
Pedig2007	More than 100,000 individuals	$F$ , gene-contribution, kinship	<a href="http://www-sgqa.jouy.inra.fr/article.php?id_article=110">http://www-sgqa.jouy.inra.fr/article.php?id_article=110</a>	Free
Partinbr	100 founders, 500 living individuals	$F$	<a href="http://www.vortex9.org/partinbr.html">http://www.vortex9.org/partinbr.html</a>	Free
PEDSYS (44 programs)	30,000 individuals. Can reduce/minimize pedigrees.	$F$ , kinship, founder-representation, genotype-maps	<a href="http://txbiomed.org/departments/genetics/genetics-detail?r=42">http://txbiomed.org/departments/genetics/genetics-detail?r=42</a>	Free

$F$  inbreeding coefficient,  $MK$  mean kinship,  $N_e$  effective population size

<sup>a</sup>  $N_e = 1/(2\Delta F)$  for each generation having  $F_t > F_{t-1}$  (Gutiérrez and Goyache 2005)

<sup>b</sup>  $N_e = 4N_mN_f/(N_m + N_f)$ ,  $N_m$  = number of living, reproducing males,  $N_f$  = number of living, reproducing females

We searched for computer programs providing inbreeding and kinship coefficients (Wright 1922) including mean kinship (Lacy 1995). The program should also estimate loss of genetic variation measured by founder alleles and provide statistics including founder contribution, founder allele survival, and founder genome equivalents (Lacy 1989). We needed the program to be able to handle large pedigrees of several thousand individuals. We found only one program that conforms to these requirements: Population Management x (PMx; Ballou et al. 2011; Lacy et al. 2011; see Table 1 for a summary of the programs identified). PMx is an accessory program to the Single Population Analysis and Records Keeping System (SPARKS)—a platform for computerized studbooks provided by the International Species Information System (ISIS; [www.isis.org](http://www.isis.org)) and is used by zoos worldwide.

Single Population Analysis and Records Keeping System generates input files for PMx but we discovered that it is difficult to generate such input files outside the SPARKS platform. SPARKS cannot import already available pedigrees (individuals have to be entered one by one), and to create the input file for PMx (the \*.ped file) by hand is very time consuming, particularly for studbooks covering hundreds or thousands of individuals. To simplify the creation of new \*.ped-files from databases that were not originally constructed for PMx we developed the converter program mPed (make ped-file; [www.popgen.su.se/mped.php](http://www.popgen.su.se/mped.php)). mPed

is written in the programming language C. mPed converts a tab-separated text file \*.txt (generated from e.g. MS Excel, Word, or Access) to a correct PMx \*.ped-file.

PMx requires the following information for each animal: identifier (ID), identifier of father (sire ID), mother (dam ID), sex, whether or not the animal is in the population (true or false), whether the animal is alive or dead (true or false), birth date, death date, location, local and other ID, and additional fields that can be used for further information. If some of the data fields PMx requires are missing (e.g. death dates) mPed can estimate such data.

mPed uses an initialization file (\*.ini) for instructions on how to format data. In the \*.ini file the user indicates a date of analysis, and all the individuals born after that date are excluded from the pedigree. Then the user gives the path to the file (or files) to be analyzed and state in which column each type of data/information can be found. The user chooses if individuals that lack a birth date should be excluded or not and states the format of birth date and sex for the input file. If death dates are not known, mPed can provide such dates from information on expected lifespan of the individuals in the population. If birth dates are missing, mPed provides dates estimated from birth year of first offspring and age of sexual maturity.

mPed can reduce (“strip”) pedigrees that are too large to run in PMx (around 20,000 individuals depending on the complexity of the kinship matrix) when using Windows 7,

64 bit (or later) and 8 GB internal memory. The stripping includes deleting dead individuals that, at the time of the analysis, do not have any living descendants. We performed numerous tests to validate the program and currently use mPed in our ongoing research (Jansson and Laikre 2011; Laikre et al. 2013).

**Acknowledgments** We thank Laurie Bingaman Lackey, Bob Lacy, Mari Edman, Nils Ryman, and an anonymous reviewer for valuable input. Financial support from the Swedish Environmental Protection Agency and the Swedish Research Council Formas to L.L. is acknowledged.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Ballou JD, Gilpin M, Foose TJ (ed) (1995) Population management for survival and recovery. Analytical methods and strategies in small population conservation. Columbia University Press
- Ballou JD, Lacy RC, Pollak JP (2011) PMx: software for demographic and genetic analysis and management of pedigreed populations (Version 1.0). Chicago Zoological Society, Brookfield, Illinois. <http://www.vortex9.org/PMx.html>
- FAO—Food and Agriculture Organization of the United Nations (2007) Report of the international technical conference on animal genetic resources for food and agriculture. Interlaken, Switzerland, September 2007, pp 3–7
- Gutiérrez JP, Goyache F (2005) ENDOG: a program for monitoring genetic variability. *J Anim Breed Genet* 122:172–176
- Jansson M, Laikre L (2011) Is there a correlation between conservation genetic status and health in dog breeds? *J Vet Behav* 6:59
- Lacy RC (1989) Analysis of founder representation in pedigrees: founder equivalents and founder genome equivalents. *Zool Biol* 8:111–123
- Lacy RC (1995) Clarification of genetic terms and their use in the management of captive populations. *Zool Biol* 14:565–578
- Lacy RC, Ballou JD, Pollak JP (2011) PMx: software package for demographic and genetic analysis and management of pedigreed populations. *Methods Ecol Evol* 3:433–437
- Laikre L, Jansson M, Allendorf FW, Jakobsson S, Ryman N (2013) Hunting threatens achieving Favourable Conservation Status for an isolated, highly inbred wolf population. *Conserv Biol*. doi: [10.1111/j.1523-1739.2012.01965.x](https://doi.org/10.1111/j.1523-1739.2012.01965.x)
- Liberg O, Andrén H, Pedersen C-H, Sand H, Sejberg D, Wabakken P, Åkesson M, Bensch S (2005) Severe inbreeding depression in a wild wolf *Canis lupus* population. *Biol Lett* 1:17–20
- Wright S (1922) Coefficients of inbreeding and relationship. *Am Nat* 56:330–333