Simulating genetic management strategies & genetic defects in dogs

AUTHOR: ROXY BERGSMA
SUPERVISOR: JACK WINDIG

MSc thesis – Animal Sciences – Chair group Animal Breeding & Genomics

Wageningen University - Registration number 1035497

August 2022

Title: Simulating genetic management strategies and genetic defects in dogs

Author: Roxy Bergsma

Registration number: 1035497

Degree: MSc thesis Animal Sciences – Animal Breeding & Genomics

Institution: Wageningen University

Supervisor: dr. Jack Windig

Date: 16-08-2022

PREFACE

This thesis was a work of passion. When asked as a child what I wanted to be when I grew up, I would answer: “Something with animals”. For a long time I didn’t know whether that dream would come true, but here we are at last. To say that I am happy for completing an MSc thesis in Animal Sciences would be an understatement. I am exhilarated, and proud, but also sad that there are two notable absences in my life that aren’t here to witness this accomplishment. They always supported me in following my dreams, so I hope they are proud, too.

Completing this thesis was quite a struggle, and I could never have done it without the continued support, advice, and feedback from my wonderful and patient supervisor Jack Windig. I feel so honored that he took me on as a student and trusted me to complete this thesis. Thank you.

Special thanks goes towards MJ and Floris, who graciously allowed me to work in their office building; Kor (Made by Korlieke), who made beautiful photos of my beloved white Swiss shepherd Mike, seen on the cover page; Aukje and Mirte, who encouraged me during hard times; my sweet Jelmer and my sisters Daisy and Santana with their unwavering faith in me; and Esther and Mathijs, who opened up their home to me in the final stretch of this thesis’ completion so that nobody would be able to find me and distract me. Much obliged!

SUMMARY

Dog breeds are characterised by closed populations, high inbreeding levels, and high rates of genetic defects. Breed management strategies focus on using selection to reduce defects, or on increasing the population size to reduce inbreeding levels and genetic drift. The aim of this research was to find out how different kinds of genetic management strategies affect the allele frequencies of various genetic defects in a small sized and medium sized effective population, and see which strategy most effectively reduces the number of defects. The program Pointer was used to run simulations. Four breeding strategies were simulated: Excluding carriers of disease alleles, breeding the least related animals, breeding the animals with the lowest mean kinship, and putting restrictions on number of litters per sire. Excluding carriers, which focused on rigorous selection, resulted in swift extinction of almost all populations. The other strategies focused on reducing drift by increasing the effective population size. They all successfully reduced inbreeding levels and rate of inbreeding, lowered the amount of fixed deleterious alleles, and helped prevent the fixation of disease alleles by decreasing the average disease allele frequencies. Focusing on mean kinship was most effective. When the effective population size is increased, some disease alleles may take longer to eliminate, but they linger at low frequencies in a heterozygous state where they do not pose a problem.

SAMENVATTING

Hondenrassen kenmerken zich door gesloten populaties, hoog niveau van inteelt en hoge mate van genetische defecten. Management strategieën richten zich op het reduceren van defecten door middel van selectie, of op het vergroten van de populatiegrootte zodat het inteeltniveau en het effect van genetische drift afnemen. Het doel van dit onderzoek was bekijken hoe allelfrequenties van verschillende genetische defecten zich gedragen in kleine en middelgrote effectieve populaties onder invloed van verschillende genetisch management strategieën en welke strategie het aantal defecten het beste reduceert. Het programma Pointer werd gebruikt voor simulaties. Vier fokstrategieën werden gesimuleerd: Uitsluiten van dragers van defecten, fokken van de minst verwante dieren, fokken van dieren met de laagste mean kinship en als laatste een restrictie op het aantal nesten per dekreu. Het uitsluiten van dragers, wat gebruik maakte van rigoreuze selectie, resulteerde in het snelle uitsterven van bijna alle populaties. De andere strategieën richtten zich op het reduceren van drift door de effectieve populatie te vergroten. Zij waren allemaal succesvol in het verlagen van inteelt, het reduceren van gefixeerde defecten en hielpen fixatie te voorkomen door het verlagen van de gemiddelde allelfrequenties van genetische defecten. Het gebruik van mean kinship was het meest effectief. Wanneer de effectieve populatiegrootte toeneemt, duurt het langer om sommige genetische defecten te elimineren, maar ze blijven heterozygoot op een lage frequentie in de populatie zitten en zorgen daar niet voor problemen.

TABLE OF CONTENTS

1. INTRODUCTION

1.1. CONTEXT

Hundreds of dog breeds exist in the world today. Though the exact date of dog domestication and the manner in which domestication occurred are not exactly known, we know it resulted in a myriad of genetically distinct dog types (Galeta et al., 2021; Parker et al., 2017). Historically, dog breeding focused on function based on human needs, which resulted in dogs suited for specific jobs such as herding, guarding, hunting, and pulling sleds (Jung & Pörtl, 2019). This changed about 150 years ago during the Victorian time period, when dog breeding became fashionable and done for sports. It was during this time that modern dog breeding was invented, which is characterized by written breed standards, dog shows, and closed studbooks (Worboys et al., 2018). Nowadays, dog breeds are distinguished from each other due to breed type, which is the degree to which a purebred dog adheres to their respective breed standard.

Breed type is created by selection. Strong artificial selection was practiced during breed formation to fix behavioral and physical traits (Mooney et al.,2021). One form of selection is deliberate inbreeding, which is the mating of two animals who share one or multiple common ancestors. Another type of inbreeding is forced inbreeding, which is not deliberate. For example: closed studbooks are inherent to most pedigreed dogs, and have been since Victorian times, so most or all dogs within a breed by necessity share common ancestors. This means that mating two dogs within a breed will nearly always result in some degree of inbreeding, though not deliberate on the breeder's part. Inbreeding increases the likelihood that an offspring inherits desirable traits from common ancestors. However, because traits are passed down by chance and every individual carries genetic defects in their DNA, inbreeding also means an increased likelihood that an individual inherits detrimental traits from common ancestors. In other words, inbreeding also has negative effects, like inbreeding depression and an increase in genetic defects (Hedrick & Garcia-Dorado, 2016; Charlesworth & Willis, 2009). Nowadays, dog breeds are known for high levels of inbreeding and diminished genetic diversity, which makes them interesting to study (Leroy, 2011, Yordy et al., 2019).

Each animal carries genetic defects. Because dominant defects with a severe impact on an animal's health tend to be quickly removed by either natural selection or selection practiced by breeders, most genetic defects that exist in a population are fully recessive (Keightley et al., 1998). It's important to realize that each gene has two alleles (one inherited from the sire, one from the dam). If one allele has a defect, while the other healthy allele still contains the information that makes the gene function properly, the animal will be healthy. Individuals with one healthy and one defect allele are therefore called 'carriers'. The deleterious allele is present in a heterozygous state, and while the individual does not express the genetic defect, they may still pass on the deleterious allele to their offspring. This means that genetic defects may not be easily detectable or selected against (Lewis & Windig, 2016; Oldenbroek & Windig, 2012).

When an animal has two defect alleles at a locus, the genetic defect may have different modes of expression. These animals are called 'affected' because the gene does not function properly. If the gene codes for an important trait that is essential for life (a life history trait, or fitness trait), this is called a lethal gene. Lethals cause an animal to die before the age of sexual reproduction, but they are rarely an issue because they are not likely to reach high frequencies in a population (Keightley et al., 1998; Windig & Doekes, 2018). However, many genes code for less visible traits, for instance because the gene has a small effect in a homozygous state. Genes may have such a small effect on a trait such as fertility, that the effects may not be noticeable unless you look at data collected over several generations (Oldenbroek & windig, 2012; Lewis & Windig, 2016).

Frequencies of genetic defects may change in a breed due to two important causes. The first cause is selection, either natural selection (such as selection against individuals with lethal genes, who naturally die before reproduction) or artificial selection (selection by humans, discussed above). Selection is targeted and aimed at decreasing the frequencies of genetic defects (Keightley et al., 1998). The second cause that changes

frequencies in a breed is genetic drift, which is the change in frequency of an allele due to chance. Genetic drift is random, and may either increase or decrease the frequencies of genetic defects. This is especially important in dog breeds with a small effective population, where only few animals contribute their genes to the next generation (Masel, 2011; Odom, 2022; Oldenbroek & Windig, 2012.

Breeding strategies may be employed to reduce genetic defects. Different types of breeding strategies use selection to reduce genetic defects, or reduce drift by increasing the effective population size. An example of a strategy that focuses on selection is the use of genetic tests to exclude individuals that carry a deleterious allele. Examples of strategies that focus on reducing drift are limiting the use of popular sires, mating individuals with a low relatedness, and expanding the population using outcross (Lewis et al., 2015; Lewis & Windig, 2016). A breeding strategy of a different order is genetic purging, which uses deliberate inbreeding to expose genetic defects and then exclude affected individuals from the population (Hedrick & Garcia-Dorado, 2016).

What kind of breeding strategy is better suited to combat genetic defects in dog breeds: strategies aimed at using selection or strategies aimed at reducing drift?

1.2. RESEARCH QUESTIONS

The aim of this research is to find out how different kinds of genetic management strategies affect the allele frequencies of genetic defects, and see which strategy most effectively reduces the number of defects. Therefore, our research question is the following:

RQ: What kinds of genetic management most effectively reduce genetic defects in dog breeding?

There are several things we need to know in order to answer the main research question. For instance, does the effectiveness of a breeding strategy change between alleles with different ages of onset? And how do various breeding strategies cope with reducing genetic defects of varying strengths of expression? Dog breeds vary in the size of their effective population, so does that make a difference in combating genetic defects? And last but not least: each individual carries more than one disease allele. So do disease alleles behave differently as a single allele versus in a group? Once we know the answers to these questions, we may find out if there is perhaps a single best strategy for tackling multiple genetic defects at once. The following sub-questions are formulated to help us answer our main research question:

SQ1. What effect does the age of onset of a disease allele have on the success of a genetic management strategy?

SQ2. What effect does the strength of expression of a disease allele have on the success of a genetic management strategy?

SQ3. What effect does the effective population size have on the success of a genetic management strategy?

SQ4. Do disease alleles behave differently as a single allele versus in a group?

In sub-questions 1, 2, and 3, ‘success’ is defined as the reduction of the number of disease alleles.

2. METHODS

2.1. SIMULATION PROGRAM: POINTER

To answer the research questions, I ran simulations in Pointer, which is a program that simulates genetic management in captive populations. Pointer is a free software written by K. Oldenbroek and J. Windig, made available by the Centre for Genetic Resources in the Netherlands (CGN) and Wageningen University and Research. (Windig & Hulsegge, 2021). I simulated fictive populations of pedigree dogs with deleterious alleles. The alleles varied in their strength of expression, in their age of onset, and in the frequency that they occurred in the population. The frequencies and population structures were based on data used in other theses (Prosman, 2019; Stilting, 2018; Nieuwenhuis, 2021; De Gouw, 2019) and based on the resulting frequencies of single allele simulations (see *2.2. Simulation parameters*).

Once I established the populations with genetic defects, I simulated various breeding strategies. The breeding strategies included mating animals based on mean kinship, excluding individuals that carry a deleterious allele, limiting the use of sires, and mating animals that are least related to each other. These strategies are described in the 2016 study by Lewis and Windig and are influenced in part by the possibilities in Pointer (Lewis & Windig, 2016; Windig & Huslegge, 2021).

The results described changes in frequencies for the genetic defects, number of eliminated and fixed alleles, inbreeding coefficients and rate of inbreeding, changes in population structure (litter size, population size, age of parents), and changes in survival. The Pointer output was converted to Excel, where graphs were created.

Pointer provided many parameters to simulate a custom captive breeding population. To create realistic simulations, the parameters were based on existing dog breeds. The parameters were chosen to reflect a small and a medium sized population of pedigree dogs. For the small population, the parameters were based on the Saarloos wolfdog. The medium sized population had its parameters based on the Schapendoes. Note that these simulated populations do not represent the Saarloos and Schapendoes populations as they exist in real life. These breeds merely served as anchors to ground the simulations in realistic parameters.

A large breed population was not simulated. This was due to the large number of animals involved and the smaller effect breeding strategies have on large populations. Besides the aforementioned points, most breeds fall into the small or medium population size categories (Lewis et al., 2015).

2.2. SIMULATION PARAMETERS

The small sized population **("Smallpop")** based its parameters off the Saarloos wolfdog. This is a Dutch breed with approximately 60-80 puppies born each year. However, historically this breed has had an extremely small number of breeding animals. The 10 most popular sires used to contribute over 90% of the next generation. This was in part because there were little more than 10 males available for breeding. In recent years, the number of available breeding males has increased from approximately 15 individuals to the current 46 at this moment of writing. To simulate a worst-case scenario in an extremely small breed, I emulated the situation as it existed for the Saarloos around 1990. Therefore 2 champion sires were set to sire 50% of the offspring. The number of breeding males in the simulation were set at 15, the number of breeding females were set at 20, and the number of litters per year at 11. The effective population size was therefore:

$(4 * 15 * 20) / (15 + 20) = 34.3$

The medium sized population **("Midpop")** based its parameters off the Schapendoes. This is a Dutch breed with 900-1000 puppies born each year. 112 males and 150 females are used for breeding. The 10 most popular sires contributed 24% of the next generation. The number of litters born per year was set at 150. The effective population size was:

(4 * 112 * 150) / (112 + 150) = 256.5

Both populations had similar starting parameters for:
- maximum number of litters per female (4 litters)
- minimum age of female at first litter (24 months)
- sires start reproducing at 2 years old (24 months).

The age structure of the parents is shown in table 1 below, loosely based on previously done breed analyses (Stilting, 2018).

Table 1: Age structure of parents

Year	Male	Female
1	0%	0%
2	18%	23%
3	26%	27%
4	24%	24%
5	15%	17%
6	11%	8%
7	4%	1%
8	2%	0%

Litter size went up to 12 newborns, and the distribution is shown in table 2 below.

Table 2: Litter size distribution

Litter size	%	Litter size	%
1	4%	2	4%
3	5%	4	9%
5	17%	6	20%
7	16%	8	11%
9	8%	10	4%
11	1%	12	1%

The starting parameters for both populations were set to the software's default for:

- maximum number of females serviced per male/year (1000)
- maximum number of females serviced per male/life (1000)
- maximum number of sons selected as breeding male per father (1000)
- maximum kinship allowed between parents (1.000)
- maximum inbreeding allowed per animal (1.000)
- no constraints of mean kinship with remainder of breed
- no minimizing of kinship between father and mother
- no optimal contributions
- no exclusion of matings between animals with common ancestors

The program simulated 100 years and each run was repeated 50 times. Two test simulations were done for both Smallpop and Midpop. These two test simulations provided a basis for the starting allele frequencies by showing how single alleles act in the two populations under 'neutral' circumstances without the application of

management strategies. In these test simulations, the population was given two test alleles: one 'neutral' and one 'disease' allele. Both alleles had an initial starting frequency of 0.5. There was no selection against the 'neutral' allele, and a selection rate of 0.5 against the 'disease' allele. The 'disease' allele's age of expression was set at 3 years, and it had a negative effect on survival.

These two test simulations were called *SmallPop Test* and *MidPop Test*. The results for the test simulations can be found in paragraph ***3.1.1. Test simulations: neutral alleles vs disease alleles***.

SmallPop Test showed that within the small sized population, neutral alleles that experience no selection either became fixed or disappeared, with a tendency to become fixed. The average allele frequency after 100 years was 0.513. Disease alleles either disappeared from the population or lingered in low frequencies. The average disease allele frequency after 100 years was 0.029, but the average frequency of disease alleles that linger in the population was 0.18. The range of frequencies found in disease alleles that lingered in the population after 100 years was 0.045, 0.083, 0.106, 0.159, 0.174, 0.295, and 0.417.

MidPop Test showed that within the medium sized population, neutral alleles that experience no selection settle around an average frequency of 0.463 after 100 years. The average disease allele frequency after 100 years was 0.053, showing no tendency to disappear. The highest attained frequency after 100 years was 0.172.

The disease allele frequencies found after 100 years were used to determine starting allele frequencies for disease alleles in the later simulations, along with information in De Gouw's thesis (2019) called 'Evolution of the frequencies of genetic defects of various strengths' (De Gouw, 2019). De Gouw found that the lower the strength of a disease allele, the more its frequency approaches the frequencies of neutral alleles. Therefore the starting allele frequencies chosen for this thesis were 0.42 for weak disease alleles, 0.05 for strong disease alleles, and 0.17 for intermediary disease alleles.

To see what the effect was of age of onset on disease allele frequency, multiple simulations were done for both Smallpop and Midpop. The results for the age of onset simulations can be found in Paragraph ***3.1.2. Effect of age of onset***. The disease had a strength of 1 (maximum selection against the allele), and a low starting frequency of 0.05 that you would expect for a strong disease allele. Age of onset was varied at 1-8 years old. The ages of onset of 0, 1, 4, and 8 were chosen for the later simulations, which reflect an allele that presents itself at birth (0), an allele that presents itself before the age of reproduction (1), within the age of reproduction (4), and after the age of reproduction (8).

2.3. DISEASE ALLELES

The simulations in this thesis required a mix of disease alleles. Disease alleles were simulated that differ in age of onset and strength of expression. The final allele mix consisted of 60 alleles of various strengths and ages of onset. The 60 alleles had appropriate starting frequencies that were based on the disease frequencies found after 100 years in the standard population simulations. The chosen starting frequencies also take into account that a disease allele has a higher frequency in a population when the age of onset is later in life, when breeding animals have already reproduced. Most alleles in this mix had a low strength. This is because alleles with a high strength of expression experience strong selection, which reduces the number of disease alleles in a population with high strength. As a result, most disease alleles were of low strength, which is reflected in the chosen allele mix (see Table 3). The 60 alleles were placed on 60 loci and were divided alternately over 2 chromosomes.

Every allele received a code. Each code followed the same structure: Allele number – strength & age of onset. As an example, take allele 1. This allele was coded 1-0010:
1 (allele number) – **001** (selection coefficient used for strength is 0.01) **0** (age of onset is 0 years old).

Table 3: The chosen allele mix for the final simulations. 60 alleles will be simulated. Most of the alleles have a low strength of expression, ranging from 0.01 to 0.1. Some alleles have a larger strength of expression, ranging from 0.25 to 1.

Allele	Code	Starting frequency	Strength	Age of onset (years)	Allele	Code	Starting frequency	Strength	Age of onset (years)
1	1-0010	0.42	0.01	0	***31***	31-011	0.42	0.1	1
2	2-0010	0.42	0.01	0	***32***	32-011	0.42	0.1	1
3	3-0011	0.42	0.01	1	***33***	33-011	0.42	0.1	1
4	4-0011	0.42	0.01	1	***34***	34-011	0.42	0.1	1
5	5-0011	0.42	0.01	1	***35***	35-014	0.42	0.1	4
6	6-0011	0.42	0.01	1	***36***	36-014	0.42	0.1	4
7	7-0014	0.42	0.01	4	***37***	37-014	0.42	0.1	4
8	8-0014	0.42	0.01	4	***38***	38-014	0.42	0.1	4
9	9-0014	0.42	0.01	4	***39***	39-018	0.42	0.1	8
10	10-0014	0.42	0.01	4	***40***	40-018	0.42	0.1	8
11	11-0018	0.42	0.01	8	***41***	41-018	0.42	0.1	8
12	12-0018	0.42	0.01	8	***42***	42-018	0.42	0.1	8
13	13-0018	0.42	0.01	8	***43***	43-0250	0.17	0.25	0
14	14-0018	0.42	0.01	8	***44***	44-0251	0.17	0.25	1
15	15-0050	0.42	0.05	0	***45***	45-0251	0.17	0.25	1
16	16-0050	0.42	0.05	0	***46***	46-0254	0.17	0.25	4
17	17-0051	0.42	0.05	1	***47***	47-0254	0.17	0.25	4
18	18-0051	0.42	0.05	1	***48***	48-0258	0.42	0.25	8
19	19-0051	0.42	0.05	1	***49***	49-0258	0.42	0.25	8
20	20-0051	0.42	0.05	1	***50***	50-050	0.05	0.5	0
21	21-0054	0.42	0.05	4	***51***	51-051	0.05	0.5	1
22	22-0054	0.42	0.05	4	***52***	52-051	0.05	0.5	1
23	23-0054	0.42	0.05	4	***53***	53-054	0.17	0.5	4
24	24-0054	0.42	0.05	4	***54***	54-054	0.17	0.5	4
25	25-0058	0.42	0.05	8	***55***	55-058	0.42	0.5	8
26	26-0058	0.42	0.05	8	***56***	56-058	0.42	0.5	8
27	27-0058	0.42	0.05	8	***57***	57-10	0.05	1	0
28	28-0058	0.42	0.05	8	***58***	58-11	0.05	1	1
29	29-010	0.42	0.1	0	***59***	59-14	0.05	1	4
30	30-010	0.42	0.1	0	***60***	60-18	0.42	1	8

2.4. GENETIC MANAGEMENT STRATEGIES

Pointer allows the simulation of numerous breeding strategies. As discussed in the introduction, breeding strategies may either use selection to reduce genetic defects, or increase the (effective) population size to reduce drift. For this thesis, I chose to simulate four different genetic management strategies to answer the research questions. One strategy that focused on selection to reduce defects, and three strategies that focus on reducing drift. A reference population was simulated where no breeding strategy was applied, for comparison's sake. Table 4 below gives a short indication of the chosen breeding strategies, and more elaborate descriptions can be found further in this paragraph.

Table 4: An overview of the simulated strategies

Strategy	Description	Simulation in Pointer
Minimise kinship between the parents. Strategy name: **"LEAST RELATED"**	Breed parents that are as unrelated to each other as possible. A standard pedigree of 3 generations shows no common ancestors.	Check 'Minimise kinship between father and mother' + 'exclude matings between animals with common ancestors in the past 3 generations' in the tab 'breeding policy'.
Constrain kinship within the breed. Strategy name: **"MEAN KINSHIP"**	Breeding animals are only allowed to breed if they have a mean kinship with all the other animals in the population that is less than or equal to the average mean kinship in the breed.	Check 'Constrain mean kinship with remainder of breed' in the tab 'breeding policy'.
Prevent popular sires. Strategy name: **"BREEDING CAP"**	Limit how many litters can be sired by a breeding male.	In the tab 'breeding policy', limit 'maximum number of females serviced by male/year' to 1 per year, and limit 'maximum number of females serviced by male/life' to 2.
Exclusion of carriers within the breed. Strategy name: **"EXCLUDE CARRIERS"**	Exclude carriers of a disease allele.	In the tab 'genome', check 'exclude carriers fathers mothers' for a) all disease alleles b) only the disease alleles with strong strengths: 0.25, 0.5, 1

'Least related' was a strategy that mimics the real life situation where breeders try to breed parents that are as unrelated to each other as possible. In addition to that, the standard 3 generation pedigree should show no common ancestors. In reality, breeders are often not able to pick the true least related mate because of various reasons. There may not be a comprehensive database or list of all available mates, kinship status may not be known, and there may be political or logistical problems that prevent the use of potential mates (Lewis & Windig, 2016; Oldenbroek & Windig, 2012). But for the sake of the simulation, I assumed breeders did have complete information available on mean kinship, and all potential mates were known and available to the breeder.

'Mean kinship' is seen by Lewis & Windig (2016) as the best breeding strategy, though only after optimal contribution selection. These two strategies do not differ very much, but optimal contribution selection is not feasible in the context of dog breeding at this time of writing. Optimal contributions require complete control of the entire breeding population, which is a requirement that is currently not met in dog breeding where

every breeder makes their own breeding decisions. Because this strategy seeks to maximise breeding values and constrain the rate of inbreeding, it is considered the most effective genetic management strategy in theory. The strategy 'Mean kinship' also constrains the rate of inbreeding by only allowing animals to breed if they have a mean kinship with all the other animals in the population that is less or equal to the average mean kinship within the breed. It requires constant computation, updating, and publishing of mean kinships, which allows breeders to take kinship status into account when making breeding decisions (Lewis & Windig, 2016).

'Breeding cap' was a breeding strategy that focuses on preventing popular sires by limiting the number of litters a sire can produce. Lewis et al (2015) states that popular sires are an important reason for the high rates of inbreeding found in purebred dogs, and Leroy (2011) even names popular sires as the most important cause for the inbreeding levels in dog breeds and their decreased genetic diversity.

'Exclude carriers' was explored using two different scenarios. In the first scenario, all individuals carrying a disease allele were excluded from breeding. This variation caused all populations to go extinct. Therefore a second variation was tried where only carriers of strong disease alleles were excluded from breeding. 'Strong disease alleles' were defined as deleterious alleles with a selection coefficient of 0.25, 0.5, and 1.

2.5. VALIDITY AND RELIABILITY

The validity of this thesis was ensured by using an appropriate simulation program that has been tried and tested for simulating population structures, breeding strategies, and disease alleles in captive populations such as dog breeds. The program even borrows its name from a purebred dog breed (Pointer). Pointer is endorsed by the Centre for Genetic Resources in the Netherlands (CGN) and Wageningen University and Research (WUR) (Windig & Hulsegge, 2021). Other theses used Pointer and its underlying program GenManSim for similar research purposes (Prosman, 2019; Stilting, 2018; Nieuwenhuis, 2021; De Gouw, 2019). The input for the population structures in the simulations were based on data from real life breeds, to ensure a realistic outcome of the results that mimic what would happen in breeds if these management strategies were applied in real time. The simulated management strategies were chosen based on a literature review. The results were compared to findings from other studies in *chapter 4: Discussion*.

The reliability of this thesis was ensured by giving a detailed description of the input parameters in chapter 2: Methods, so that anyone might be able to replicate the results of this thesis as they are found in chapter 3: Results. The simulations were done consistently and carefully and by taking the same steps each time.

3. RESULTS

3.1. SINGLE ALLELES

3.1.1. TEST SIMULATIONS: NEUTRAL ALLELES VS DISEASE ALLELES

The program Pointer did what it was designed to do and kept both populations stable over the course of 100 years. Both Smallpop and Midpop survived. The average number of puppies per year remained constant, with an average of 66 puppies for Smallpop and 888 puppies for Midpop. The number of puppies per litter remained stable at an average of 5.77, compared to Midpop at an average of 6.03. The mean age of breeding animals remained constant as well, with a mean age in years of 3.7 (males) and 3.6 (females) for Smallpop, and 3.9 (males) and 3.6 (females) for Midpop. The mean inbreeding coefficient (COI) after 100 years is 0.668 (0.656-0.679) for Smallpop, and 0.107 (0.105-0.109) for Midpop.

Both populations were simulated with a neutral allele and a disease allele which started at an initial frequency of 0.5. The simulations showed that a neutral allele retained an average allele frequency of 0.5 in both populations over the course of 100 years. In Smallpop, drift caused the neutral allele to fixate or get eliminated in 24 runs, which started in year 25 (fig. 1). In contrast, 100 years was not enough time for the same to happen in any run for Midpop (fig. 2). The disease allele presented a different scenario. The average disease allele frequencies decreased (Fig. 3 & Fig. 4). In Smallpop, the first populations free of disease appeared in year 22. In Midpop, this only happened from year 81 onwards. In Smallpop, after 100 years the neutral allele was fixed in 13 runs and eliminated in 11 runs, while the disease allele was eliminated in 42 runs (Table 5). In contrast, the neutral allele was never fixed or eliminated in Midpop, while the disease allele was eliminated in 4 runs (Table 5). On average, the disease allele tended to linger in low frequencies after 100 years (Smallpop: 0.029; Midpop: 0.053). For some populations, the disease allele frequency after 100 years was substantial. The highest disease allele frequency after 100 years was found in a Smallpop run at 0.417, which was in the same range you might find neutral allele frequencies. The fraction of homo- and heterozygotes remained fairly stable for the neutral allele. For the disease allele, the homozygous affected individuals were almost completely gone after 100 years, with a low number of disease carriers lingering in the population.

Table 5: Over 50 runs, the number of times an allele is fixed or eliminated from a population after 100 years. Two populations were simulated: a small sized population (Smallpop) and a medium sized population (Midpop). Two types of alleles were simulated: a neutral allele with a selection ratio of 0, and a disease allele with a selection ratio of 0.5. For comparison's sake, both alleles started with a starting frequency (p) of 0.5. Alleles don't tend to be fixed or eliminated in the mid sized population, in contrast to the small sized population.

	Smallpop		**Midpop**	
	Fixed	*Eliminated*	*Fixed*	*Eliminated*
Neutral allele	13	11	0	0
Disease allele	0	42	0	4

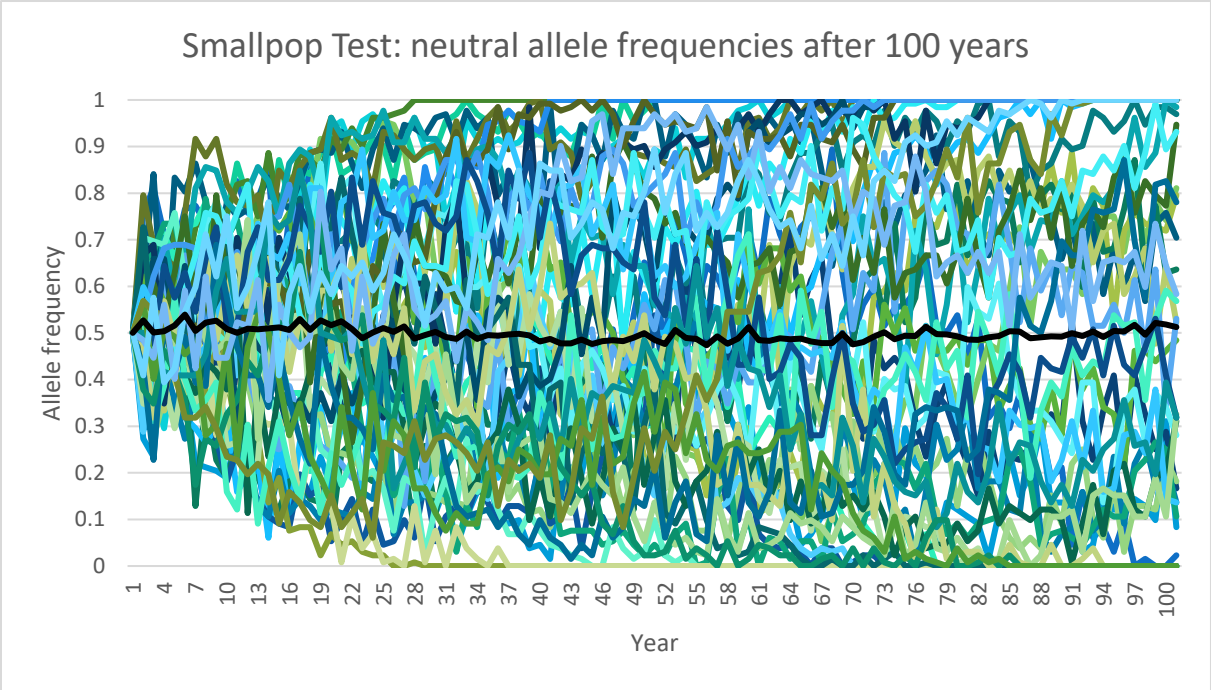


Figure 1: Neutral allele frequencies over 100 years, for 50 runs, for the small sized population (15 males, 20 females). Starting frequency = 0.5. Selection coefficient = 0. The average neutral allele frequency after 100 years is 0.513. Eliminated alleles: 11. Fixed alleles: 13.

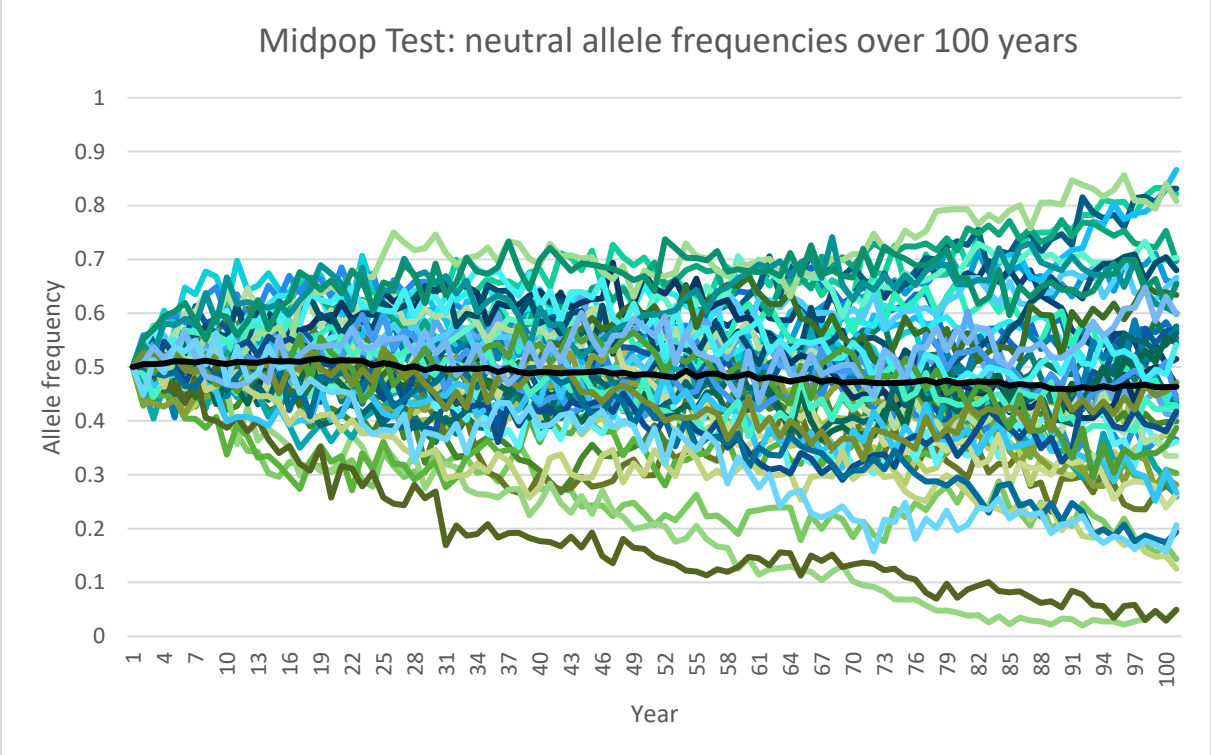


Figure 2: Neutral allele frequencies over 100 years, for 50 runs, for the medium sized population (112 males, 150 females). Starting frequency = 0.5. Selection coefficient = 0. The average neutral allele frequency after 100 years is 0.463. Alleles were not eliminated or fixed.

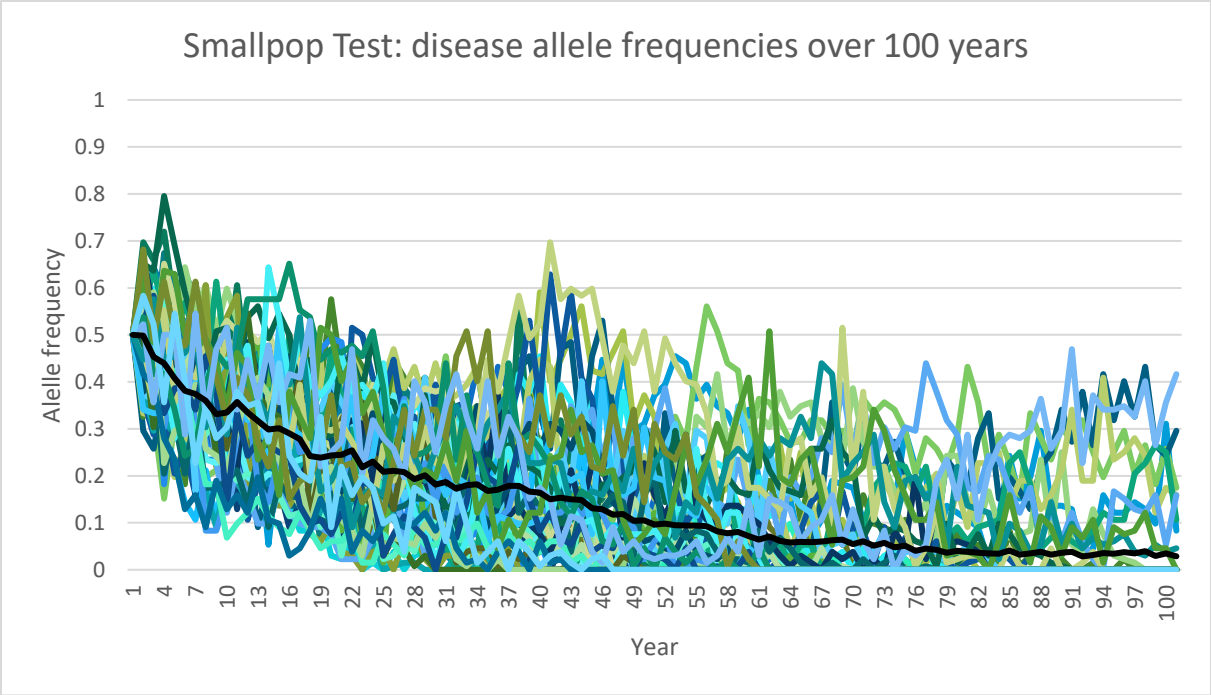


Figure 3: Disease allele frequencies over 100 years, for 50 runs, for the small sized population (15 males, 20 females). Starting frequency = 0.5. Selection coefficient = 0.5. The average neutral allele frequency after 100 years is 0.029. Eliminated alleles: 42. Fixed alleles: 0. Highest reached allele frequency after 100 years: 0.42.

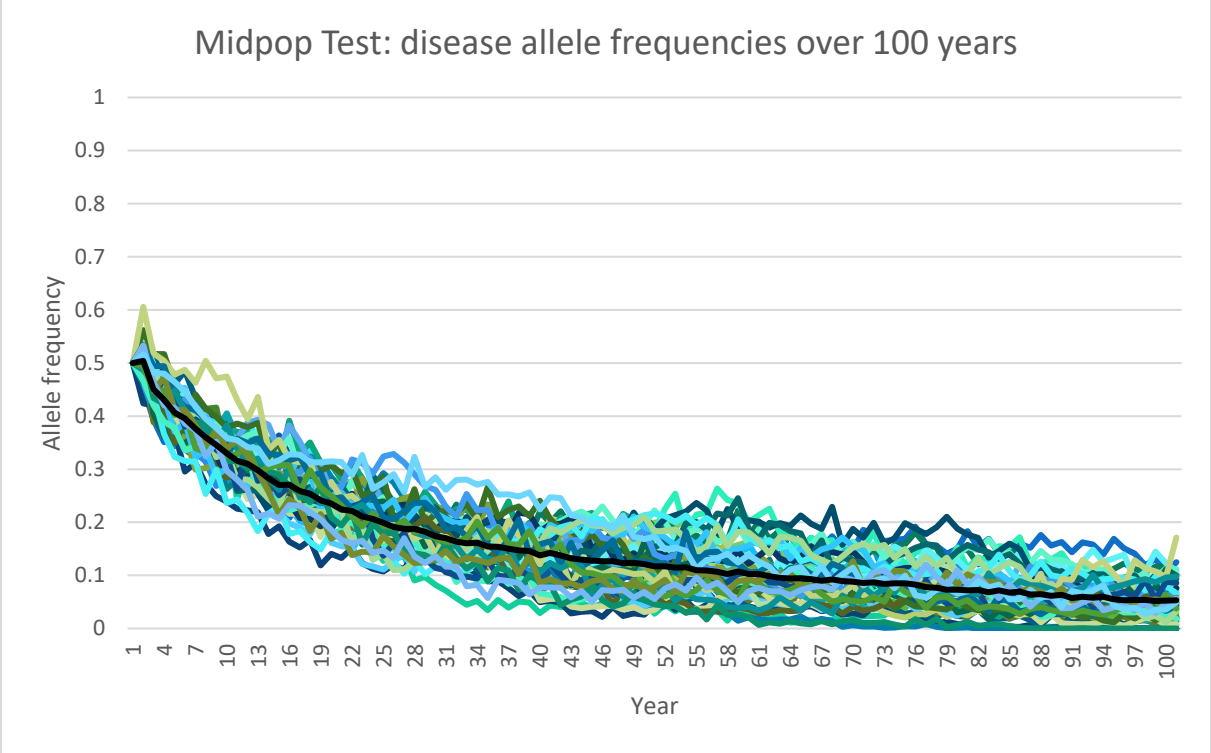


Figure 4: Disease allele frequencies over 100 years, for 50 runs, for the medium sized population (112 males, 150 females). Starting frequency = 0.5. Selection coefficient = 0.5. The average neutral allele frequency after 100 years is 0.053. Eliminated alleles: 4. Fixed alleles: 0. Highest reached allele frequency after 100 years: 0.17.

3.1.2. EFFECT OF AGE OF ONSET

The range of disease allele frequencies and the number of fixed, segregated, and eliminated runs differed depending on the age of onset of a disease allele (Table 6). Frequency graphs (Fig. 5) corresponding to different ages of onset in both Smallpop and Midpop show a visual representation of the data in Table 6.

Table 6: Disease allele frequencies after 100 years for various ages of onset of disease. This table shows the number of times an allele was fixed or eliminated in a population after 100 years and 50 runs. The disease allele had a selection coefficient of 1 and a starting frequency of 0.05. The medium sized population (112 males, 150 females, Midpop) showed no runs with a fixed allele, and in most runs the allele remained segregated in the population. The small sized population (15 males, 20 females, Smallpop) shows 1 run with a fixed allele at an age of onset of 8 years old, and in most runs the allele was eliminated.

Age of onset	**Smallpop**			**Disease allele frequencies**		
	Fixed	*Segregated*	*Eliminated*	*Highest (segregated)*	*Lowest (segregated)*	*Average over all runs*
1	0	1	49	0.038	0.038	0.001
2	0	2	48	0.091	0.008	0.002
3	0	1	49	0.015	0.015	0.000
4	0	2	48	0.030	0.023	0.001
5	0	7	43	0.258	0.091	0.026
6	0	3	47	0.742	0.083	0.028
7	0	7	43	0.955	0.091	0.066
8	1	5	44	0.099	0.038	0.028
Age of onset	**Midpop**			**Disease allele frequencies**		
	Fixed	*Segregated*	*Eliminated*	*Highest (segregated)*	*Lowest (segregated)*	*Average over all runs*
1	0	23	27	0.117	0.003	0.016
2	0	25	25	0.073	0.001	0.012
3	0	31	19	0.105	0.003	0.017
4	0	31	19	0.150	0.001	0.027
5	0	29	21	0.157	0.002	0.026
6	0	29	21	0.187	0.001	0.033
7	0	33	17	0.178	0.001	0.043
8	0	33	17	0.202	0.001	0.040

The average disease allele frequency at 100 years increased with higher age of onset. The number of populations with segregating disease alleles saw a slight increase with a higher age of onset. The population size, age of sires and dams, and litter size remained stable. When compared to Midpop, Smallpop had a much lower incidence of segregation, which was accompanied by lower average disease allele frequencies. In most Smallpop runs the disease allele was eliminated, while was only the case for less than half of the Midpop runs. The allele frequencies showed much more variation in Smallpop, and could spike to high frequencies over the course of 100 years. One Smallpop run (age of onset: 8) even saw the disease allele fixed in year 59. The allele frequencies remained much more stable in Midpop, and never approached fixation. For Smallpop, the ages of onset 5-8 showed a notable increase of segregated populations and a surge of disease allele frequencies, when

compared to ages of onset 1-4. For Midpop, there was no obviously discernible turning point, as increases in allele frequency and segregation happened gradually.

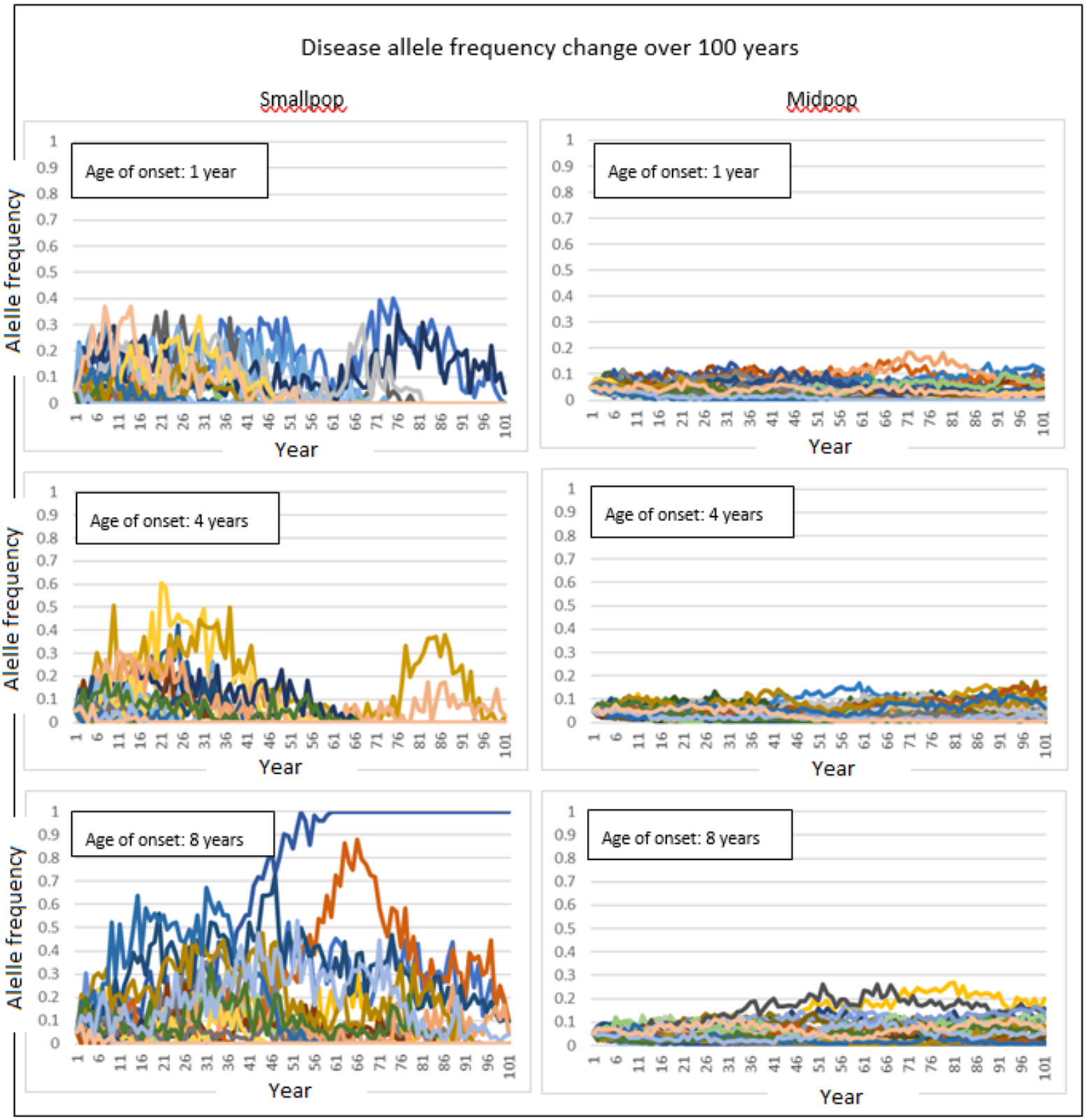


Figure 5: Change in disease allele frequency over 100 years and 50 runs shown for ages of onset 1, 4, and 8 years old. Left: small sized population (15 males, 20 females). Right: medium sized population (112 males, 150 females). Starting frequency in all graphs: 0.05. Selection coefficient in all graphs: 1. Y-axis show allele frequencies. X-axis show time in years. The Smallpop graphs show more fluctuation in allele frequencies than the Midpop graphs. In Smallpop graph: Age of onset 8 years, one disease allele becomes fixed in year 61, and another allele comes close to fixing but then decreases in frequency.

3.1.3. EFFECT OF DISEASE STRENGTH

Disease alleles behaved differently in a population over the course of 100 years depending on the strength of disease (Table 7). The starting frequency was 0.5 to enable easy comparison to the test simulations in paragraph ***3.1.1. Test simulations: Neutral alleles vs disease alleles.***

Table 7: Disease allele frequencies after 100 years for various disease strengths. This table shows the number of times an allele was fixed or eliminated in a population after 100 years and 50 runs. The disease alleles differ in their selection coefficient that represents the various strengths, with 0.01 being the lowest strength and 1 being the highest. All simulations have a starting frequency of 0.05. The higher the strength of disease, the lower the average disease allele frequency found in the population, and the more likely it is the disease allele is eliminated.

Strength	**Smallpop**			**Disease allele frequencies**		
	Fixed	*Segregating*	*Eliminated*	*Highest (segregating)*	*Lowest (segregating)*	*Average over all runs*
0.01	7	32	11	0.978	0.038	0.443
0.05	7	34	9	0.962	0.008	0.446
0.1	4	29	17	0.924	0.023	0.324
0.25	1	27	22	0.849	0.046	0.212
0.5	0	28	32	0.841	0.023	0.073
1	0	11	39	0.394	0.008	0.035
Strength	**Midpop**			**Disease allele frequencies**		
	Fixed	*Segregating*	*Eliminated*	*Highest (segregating)*	*Lowest (segregating)*	*Average over all runs*
0.01	0	50	0	0.773	0.123	0.478
0.05	0	50	0	0.694	0.059	0.357
0.1	0	50	0	0.503	0.110	0.292
0.25	0	50	0	0.330	0.010	0.158
0.5	0	50	0	0.246	0.002	0.092
1	0	47	3	0.127	0.006	0.055

The simulations for both Smallpop and Midpop showed that the average allele frequency after 100 years was lower when the strength of a disease allele was larger. Smallpop once again showed more fixed and eliminated runs when compared to Midpop, where most runs retained segregating alleles. Disease alleles with a low strength (0.01, 0.05) behaved similarly to neutral alleles. The graphs for Smallpop and Midpop with a disease allele strength of 0.05 looked very similar to the neutral allele frequencies shown in Figures 1 & 2 (Fig. 6).

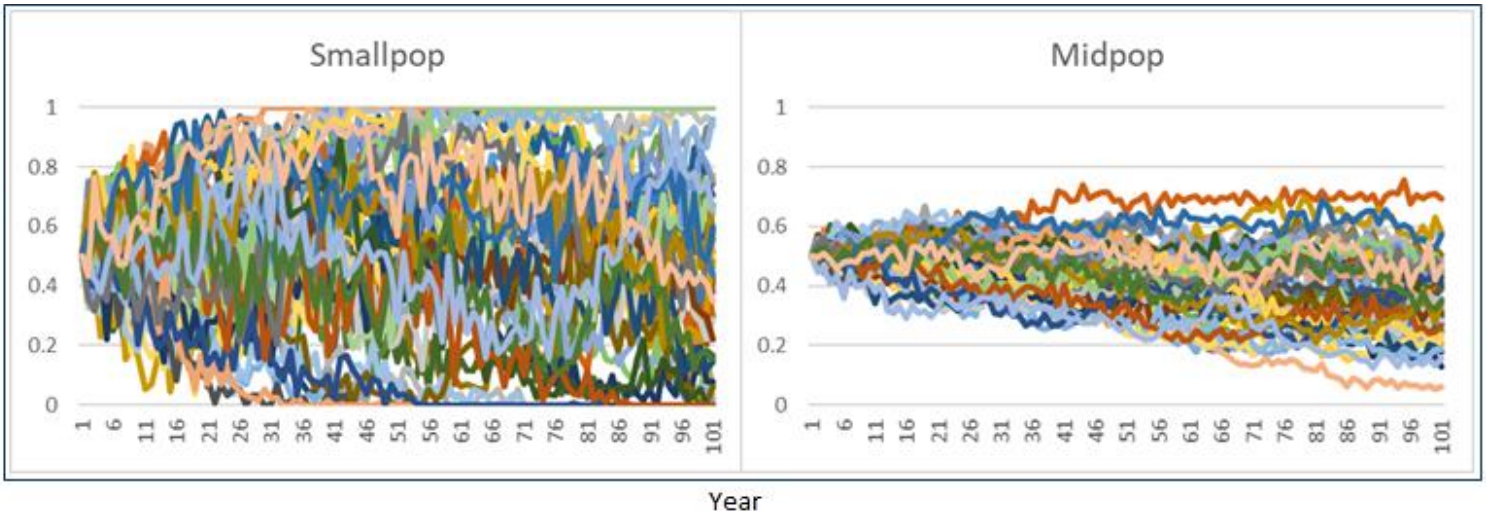


Figure 6: Change in disease allele frequency over 100 years and 50 runs shown for the small sized population (left, 15 males and 20 females) and medium sized population (right, 112 males, 150 females). Starting frequency: 0.5. Selection coefficient (disease strength): 0.05. These graphs show that the disease alleles behave similarly to neutral alleles in the test simulation, shown in Figures 1 & 2.

3.2. MULTIPLE ALLELES: REFERENCE POPULATION

3.2.1. SMALL SIZED POPULATION

A mix of 60 alleles was simulated for both Smallpop and Midpop. These simulations functioned as the reference populations for multiple alleles. The results are described below, and tables containing simulation data can be found in Appendices A and B.

None of the Smallpop populations went extinct. Within five years, litter size jumped up from 4.9 to 5.5 pups per litter, where it remained. This litter size of 5.5 was a bit lower than the litter size found in Smallpop when there were no disease alleles present (Table 8). The breeding population remained stable with on average 6 fathers and 12 mothers per year. The age of mothers and fathers was consistently around 3.5 years old for both. The inbreeding level of the population after 100 years was 0.67, and the average kinship for breeding animals was 0.70.

Smallpop saw many runs where disease alleles became fixed (Table 9). Higher ages of onset of disease were accompanied by higher rates of fixation of disease alleles. All diseases with age of onset 8 showed fixation. Lower strengths of disease were also accompanied by higher rates of fixation of disease alleles. For instance, disease strength 0.01 showed fixation among all ages of onset, even age of onset 0. But disease strength 1 only showed some fixation of disease alleles at age of onset 8. The highest number of populations where the disease allele was fixed was 8, for allele 12 (age of onset 8, strength 0.01).

Elimination of disease alleles occurred very commonly. All ages of onset and all strengths of disease showed eliminated populations. Highest rates of elimination were shown among lower ages of onset and higher strengths of disease. The highest number of populations where the disease allele was eliminated was 48, for allele 50 (age of onset 0, strength 0.5).

Higher ages of onset and lower strengths of disease were accompanied by more segregating populations, which was also where the highest disease allele frequencies are found. The highest average disease allele frequency found was 0.43 (age of onset 8, strength 0.01). The lowest average frequency found was 0.004 (age of onset 0, strength 0.5).

3.2.2. MEDIUM SIZED POPULATION

None of the Midpop populations went extinct either. Litter size decreased in year two from 6.7 to 5.6 pups per litter, and settled at 6 puppies per litter. The breeding population remained stable with on average 71 fathers and 142 mothers per year. The age of mothers was consistently 3.5 years old and the age of fathers remained 4 years old. The inbreeding level of the population after 100 years was 0.12, and the average kinship for breeding animals was 0.13.

Midpop saw no runs where disease alleles became fixed (Table 5). Elimination of disease alleles occurred more often. The highest rates of elimination were shown among lower ages of onset and higher strengths of disease. The highest number of populations where the disease allele was eliminated was 30, for allele 51 (age of onset 1, strength 0.5).

Midpop saw high rates of segregating disease alleles in populations. Higher ages of onset and lower strengths of disease were accompanied by more segregating populations, which was also where higher disease allele frequencies were found. The highest average disease allele frequency found was 0.427 (age of onset 8, strength 0.05). The lowest average frequency found was 0.01 (age of onset 0, strength 1).

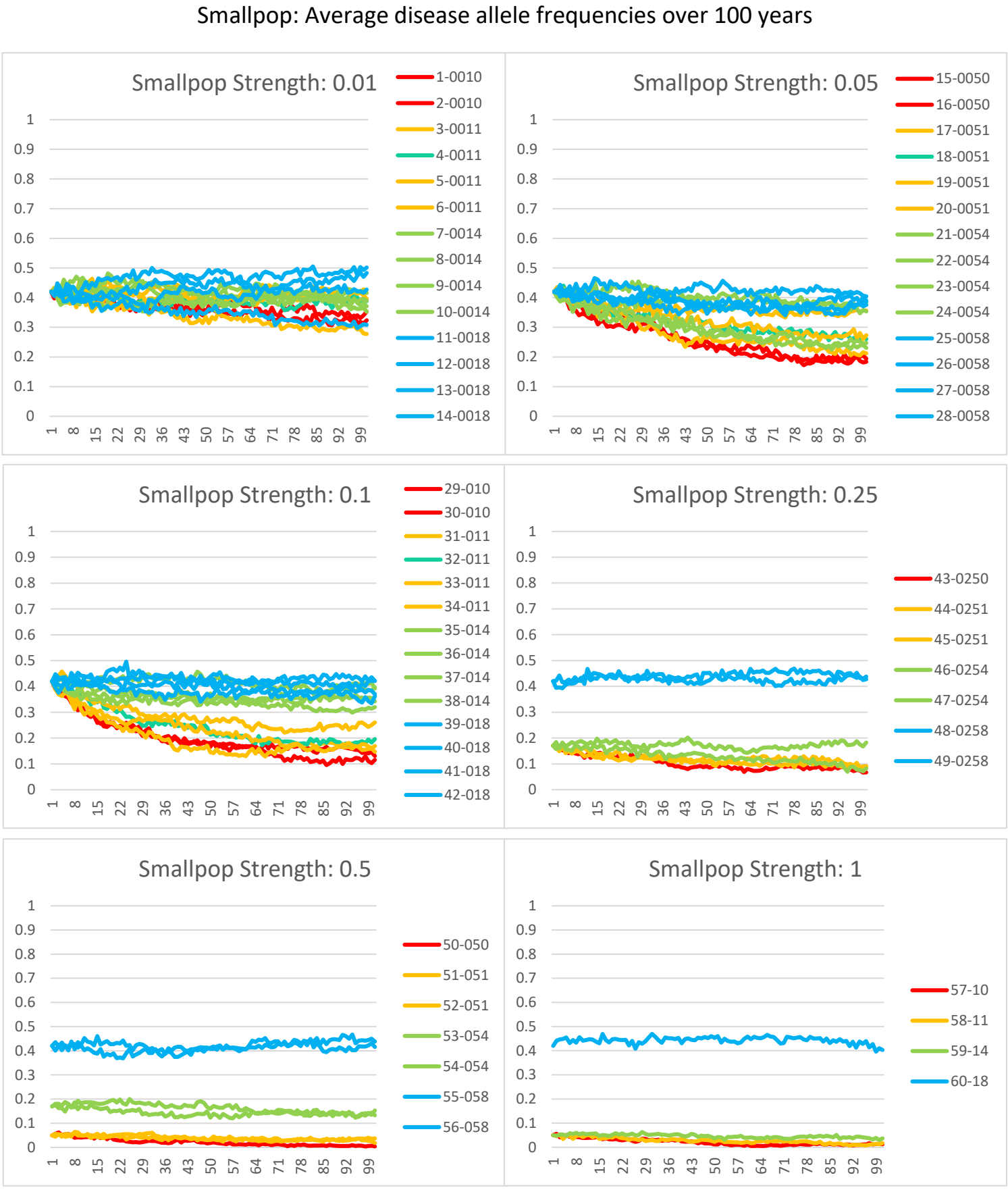
Table 8: Litter size comparison between the reference population (with 60 disease alleles) and the test simulation (without any disease alleles). Litter size for the medium sized population (112 males, 150 females) remained unchanged, but litter size for the small sized population (15 males, 20 females) decreased with 0.3 puppy in the reference population.

	Test simulation (neutral allele)	**Reference population (60 disease alleles)**
Smallpop	5.8	5.5
Midpop	6	6

Table 9: The total number of runs where a disease allele was fixed or eliminated in the small sized reference population (15 males, 20 females) and medium sized reference population (112 males, 150 females). In the medium sized population, no alleles were fixed, and 171 alleles were eliminated. In the small sized population, 44 alleles were fixed and 589 alleles were eliminated.

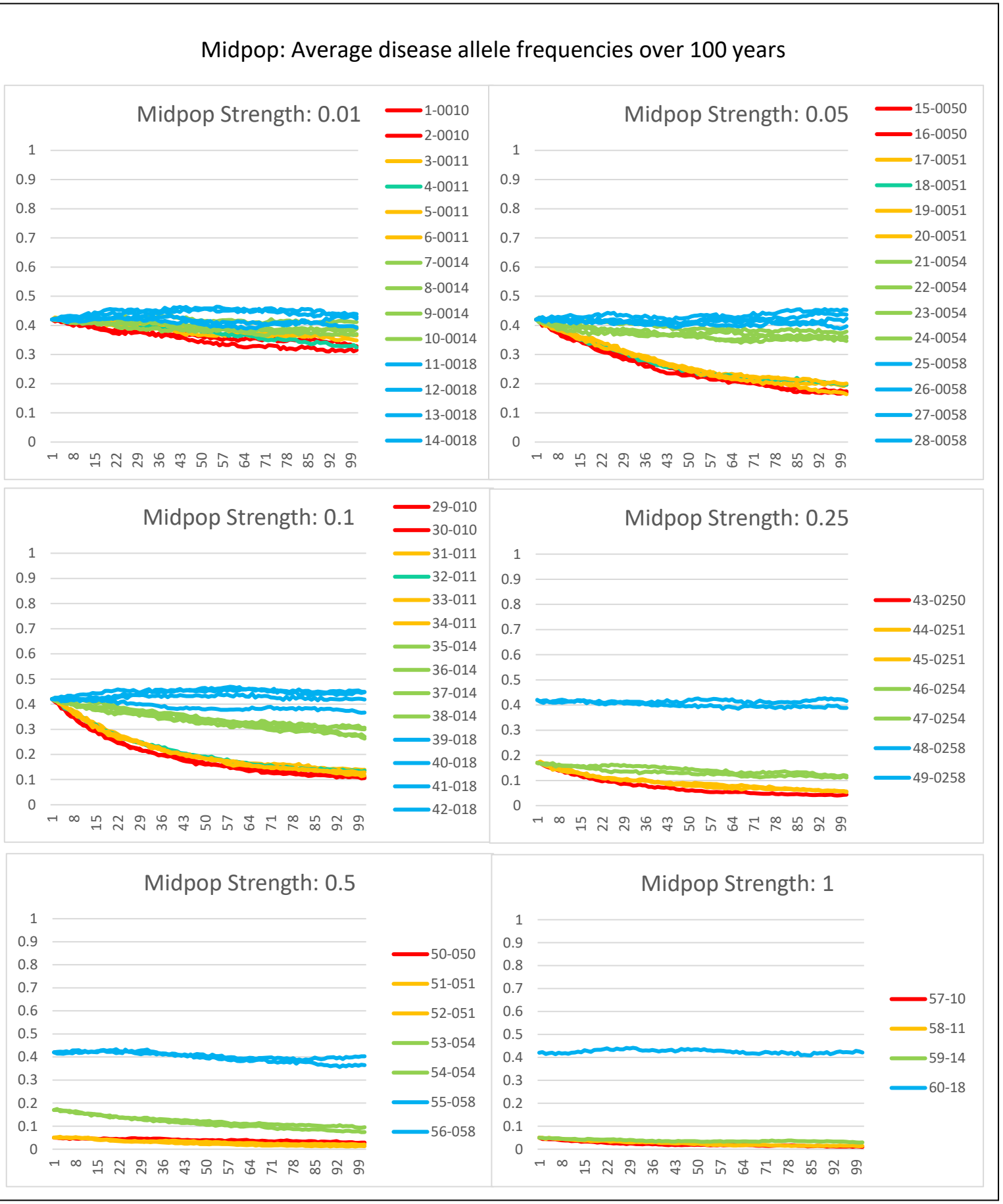
	Total runs with fixed alleles	**Total runs with eliminated alleles**
Smallpop reference population	44	589
Midpop reference population	0	171

Smallpop showed more fluctuation in allele frequencies compared to Midpop. Disease alleles with an age of onset 8 (blue lines) remained consistently at the frequency of a neutral allele, regardless of the strength of disease. Disease alleles with an age of onset 0 (red lines) were consistently found at the lowest frequencies, followed by age of onset 1 (orange line), and then age of onset 4 (green line). These patterns in average allele frequencies found in all 60 alleles over 100 years were apparent in both Smallpop and Midpop (Fig. 7 & Fig. 8).



Red = age of onset 0, orange = age of onset 1, green = age of onset 4, blue = age of onset 8.

Figure 7: Average disease allele frequencies over 100 years and 50 runs for all 60 disease alleles in the small sized population (15 males, 20 females), divided by strength of disease and age of onset of disease.



Red = age of onset 0, orange = age of onset 1, green = age of onset 4, blue = age of onset 8.

Figure 8: Average disease allele frequencies over 100 years and 50 runs for all 60 disease alleles in the medium sized population(112 males, 150 females), divided by strength of disease and age of onset of disease.

3.3 BREEDING STRATEGIES

3.3.1. GENERAL OBSERVATIONS

An overview of the general observations of the results of the four different breeding strategies revealed the differences between the four breeding strategies compared to the reference population (Table 10). The information in Table 10 is discussed further down in this paragraph.

Table 10: Overview of the general observations of four different breeding strategies applied for 100 years: 'Least related', 'Mean kinship', 'Breeding cap', and 'Exclude carriers'. The results are shown for both the small sized population (15 breeding males, 20 breeding females) and the medium sized population (112 breeding males, 150 breeding females). The results of the reference (no breeding strategies applied) are shown for comparison.
*** All small sized populations went extinct in year 2 when 'Exclude carriers' was applied.**

Smallpop		*Reference*	***Least related***	***Mean kinship***	***Breeding cap***	***Exclude carriers***
Average litter size		5.5	5.5	5.6	5.7	- *
Litters per year		11.9	11.9	11.8	11	- *
Actual fathers per year		6.1	4.7	5.5	11	- *
Average age per year…	*Breeding males*	2.8	2.8	2.9	2.1	- *
	Actual fathers	3.6	3.9	3.7	2.8	- *
	Breeding females	2.8	2.8	2.8	2.8	- *
	Mothers	3.5	3.5	3.5	3.5	- *
Average inbreeding level after 100 years		0.674	0.513	0.358	0.420	- *
Midpop		*Reference*	***Least related***	***Mean kinship***	***Breeding cap***	***Exclude carriers***
Average litter size		5.9	5.9	5.9	6	5.3
Litters per year		143	143	142.8	86.8	120.9
Actual fathers per year		70.8	36.9	52.1	86.9	60.9
Average age per year…	Breeding males	3.2	3.1	3.1	2.1	2.8
	Actual fathers	3.9	4.9	3.9	2.8	3.8
	Breeding females	2.9	2.8	2.8	2.9	2.5
	Mothers	3.5	3.5	3.5	3.6	3.2
Average inbreeding level after 100 years		0.116	0.098	0.065	0.069	0.518

BREEDING STRATEGY: 'EXCLUDE CARRIERS'

This breeding strategy focused on excluding carriers of a disease. Two scenarios were simulated: one in which every individual carrying a disease allele was excluded from breeding, and one in which only individuals carrying disease alleles for high strength diseases were excluded (s= 0.25, 0.5, 1).

In the scenario where every carrier was excluded, all runs for both Smallpop and Midpop went extinct in year 2.

In the scenario where only carriers of high strength diseases were excluded, all Smallpop runs went extinct in year 2. Out of the initial 50 runs, only one single Midpop run survived after year 3. This one remaining run had a slow start in the simulation, as can be observed in figure 9 below. Very few animals made it through the initial selection. It took more than a decade to rebuild the population size, which was reflected in the number of mothers (= litters) and fathers per year.

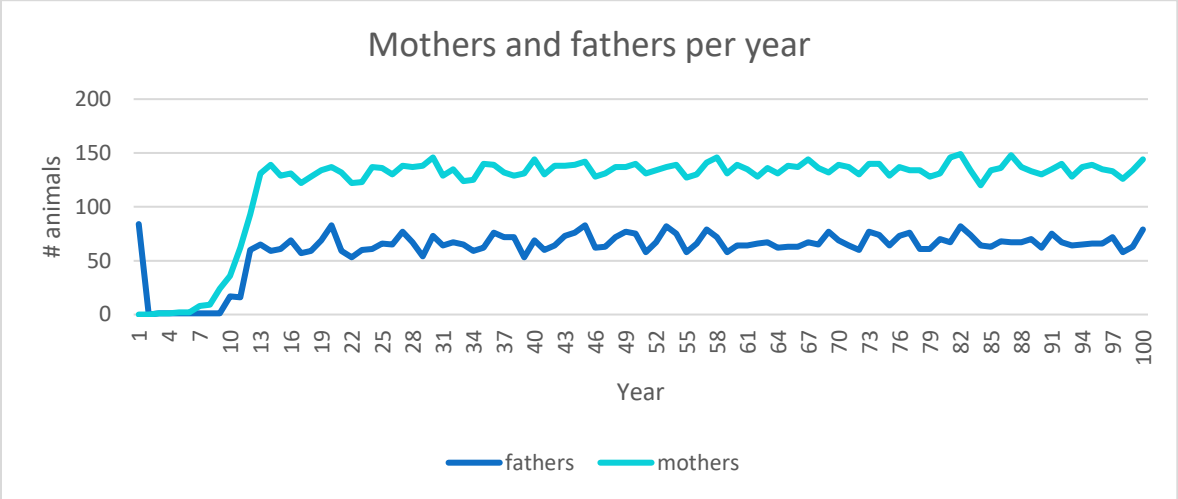


Figure 9: The number of mothers and fathers per year in the medium sized population (112 breeding males, 150 breeding females) that is subjected to the 'Exclude carriers' breeding strategy. This run almost went extinct in the first decade, with only 1 actual father throughout years 3-9. The population stabilizes again from year 13 onwards.

After the first decade, the population stabilized. Litter size became an average of 5.3 puppies (compared to 5.9 in the reference). The average number of litters per year was 120.9 (compared to 143). The number of actual fathers per year was 60.9 (compared to 70.8). The average ages of breeding males, fathers, breeding females, and mothers was approximately the same as those found in the reference (2.8, 3.8, 2.5, 3.2 years old, compared to 3.2, 3.9, 2.9, 3.5 in the reference). A striking observation was the inbreeding level in the single surviving run, which was much higher than in the reference or in any of the other breeding strategies: 0.518 (reference: 0.116).

Because only one Midpop run survived 100 years, there was no average disease allele frequency. The data for this single surviving run can be found in Appendix I. Two alleles were fixed: both alleles had an age of onset of 8, and the disease strengths were 0.05 and 0.1. Twenty-two disease alleles were eliminated from this population. These eliminated alleles include all alleles with disease strength 0.25, 0.5, and 1; three alleles with disease strength 0.05 and ages of onset 1 and 4; and one allele with disease strength 0.01 and age of onset 8.

BREEDING STRATEGY: 'LEAST RELATED'

This strategy focused on minimizing kinship between parents, in addition to excluding common ancestors in the past 3 generations. All Smallpop and Midpop runs survived. The average litter size and number of litters per year was respectively 5.5 puppies and 11.9 litters for Smallpop, and 5.9 puppies and 143 litters for Midpop. These numbers were exactly the same as in the reference. Compared to the reference, the number of actual

fathers per year was lower for both Smallpop and Midpop. Smallpop had 4.7 fathers per year (base pop: 6.1), and Midpop had 36.9 fathers per year (base pop: 70.8). The age of breeding animals in Smallpop was on average similar to those found in the reference (breeding males 2.8, actual fathers 3.9; breeding females 2.8, mothers 3.5 years old). The same was true for the age of breeding animals in Midpop, which were also similar to those found in the reference (breeding males 3.1; breeding females 2.8, mothers 3.5 years old). The exception was the average age of actual fathers, which was 4.9 instead of 3.9 years old. The average inbreeding level in the populations after 100 years was lower than in the reference, with 0.513 for Smallpop (compared to 0.674), and 0.098 for Midpop (compared to 0.116).

Smallpop showed more runs with a fixed disease allele than Midpop, where disease alleles were never fixed. Smallpop also showed more runs where the disease allele was eliminated. This occurred at all ages of onset and at all disease strengths. In contrast, disease alleles were usually eliminated in Midpop runs if age of onset was low and disease strength was high. The largest number of Smallpop runs where the disease allele was fixed was 8, which happened for an allele with age of onset 8 and disease strength 0.25 (allele 49). The largest number of Smallpop runs where the disease allele was eliminated was 48, which happened for alleles with ages of onset 0 and 1 and disease strength 1 (alleles 47 and 58). The lowest average disease allele frequency for Smallpop was 0.007, and the highest allele frequency was 0.481 (age of onset 8, disease strength 1). For Midpop, the lowest average disease allele frequency was 0.017, and the highest allele frequency was 0.448 (age of onset 8, disease strength 0.05). All allele frequency data for the breeding strategy 'Least related' can be found in Appendices C and D for Smallpop and Midpop respectively.

BREEDING STRATEGY: 'MEAN KINSHIP'

This strategy focused on constraining kinship within the population. All Smallpop and Midpop runs survived. The average litter size was 5.6 and 5.9 puppies for respectively Smallpop and Midpop, which was similar or close to litter size in the reference. The same was true for number of litters per year, which was 11.8 litters for Smallpop and 142.8 for Midpop. The number of actual fathers was lower for both populations when compared to the reference. The actual fathers were 5.5 for Smallpop, which was slightly lower than in the reference (6.1). A larger difference was seen in Midpop, where the actual fathers numbered 52.1 (compared to 70.8). In both populations, the ages of breeding animals were approximately the same as those found in the reference (Table 4). The average inbreeding level after 100 years was low compared to the reference: 0.358 in Smallpop (compared to 0.674) and 0.065 in Midpop (compared to 0.116).

Smallpop showed more runs with a fixed disease allele than Midpop, where disease alleles were never fixed. Smallpop also showed more runs where the disease allele was eliminated. This occurred at all ages of onset and at all disease strengths. In contrast, disease alleles were usually eliminated in Midpop runs if age of onset was low and disease strength was high. The largest number of Smallpop runs where the disease allele was fixed was 5, which happened for two alleles: one with age of onset 8 and disease strength 0.01 (allele 12), and one with age of onset 8 and disease strength 0.5 (allele 55). The largest number of Smallpop runs where the disease allele was eliminated was 46, which happened for allele 52 with an age of onset of 1 and disease strength 0.5. The lowest average disease allele frequency for Smallpop was 0.011, and the highest allele frequency was 0.477 (age of onset 8, disease strength 01). For Midpop, the lowest average disease allele frequency was 0.021, and the highest allele frequency was 0.428 (age of onset 8, disease strength 0.05). All allele frequency data for the breeding strategy 'Mean kinship' can be found in Appendices E and F for Smallpop and Midpop respectively.

BREEDING STRATEGY: 'BREEDING CAP'

This breeding strategy focused on limiting popular sires by restricting their total number of breedings per life, and by restricting the number of serviced females per male per year. In this scenario, all runs survived in both Smallpop and Midpop. The average litter size was approximately the same compared to the reference, with 5.7 puppies for Smallpop and 6 puppies for Midpop. However, the average number of litters per year was lower,

with 11 litter for Smallpop (compared to 11.9 in the base population), and 86.8 litters for Midpop (compared to 143 litters). The number of actual fathers per year was the same as the number of yearly litters per year (11 for Smallpop, and 86.8 for Midpop), which was higher than the actual fathers in the reference (6.1 for Smallpop, 70.8 for Midpop). The average ages for breeding females and mothers was similar to those found in the reference for both Smallpop and Midpop (respectively 2.8 and 3.5, 2.9 and 3.6 years old). However, the average ages for breeding males and fathers was lower than in the reference, with 2.1 and 2.8 for both Smallpop and Midpop. The average inbreeding levels were also lower than in the reference: 0.069 for Midpop (compared to 0.116), and 0.420 for Smallpop (compared to 0.674).

Just like with the previous breeding strategies, Smallpop showed more runs with a fixed disease allele than Midpop, where disease alleles were never fixed. Smallpop also showed more runs where the disease allele was eliminated. This occurred at all ages of onset and at all disease strengths. In contrast, disease alleles were only eliminated in Midpop runs if age of onset was low and disease strength was high. The one exception was a single run with age of onset 8 and disease strength 0.01, where the disease allele was eliminated. The largest number of Smallpop runs where the disease allele was fixed was 5, which happened for an allele with age of onset 8 and disease strength 0.01 (allele 13). The largest number of Smallpop runs where the disease allele was eliminated was 49, which happened for an allele with age of onset 1 and disease strength 0.5 (allele 51). The lowest average disease allele frequency for Smallpop was 0.008 (age of onset 0, disease strength 1) and the highest allele frequency was 0.472 (age of onset 8, disease strength 1). For Midpop, the lowest average disease allele frequency was 0.012 (age of onset 1, disease strength 1) and the highest allele frequency was 0.442 (age of onset 8, disease strength 0.1). All allele frequency data for the breeding strategy 'Breeding cap' can be found in Appendices G and H for Smallpop and Midpop respectively.

Table 11: Rate of inbreeding (Delta F), shown for all four breeding strategies, for the small sized population (Smallpop, 15 breeding males and 20 breeding females)) and the medium sized population (Midpop, 112 breeding males and 150 breeding females).

	Reference	**Least related**	**Mean kinship**	**Breeding cap**	**Exclude carriers**
Smallpop	*0.039*	0.026	0.016	0.017	-
Midpop	*0.005*	0.004	0.002	0.002	0.025

The rate of inbreeding differed depending on the breeding strategy that was applied and on the effective population size (Table 11). Compared to the reference, where no breeding strategy was employed, all breeding strategies showed a lower rate of inbreeding with the exception of strategy 'exclude carriers'. The rate of inbreeding seen in the single surviving Midpop run was 0.025, which was higher than those found in the other Midpop strategies and was more in line with the rates found in Smallpop.

3.3.2. BREEDING STRATEGIES & TOTAL FIXED AND ELIMINATED ALLELES

Looking at the total amount of times an allele was fixed or eliminated in a run, the simulations showed notable differences between Smallpop and Midpop (Fig. 10). Contrary to Smallpop, disease alleles were never fixed in Midpop, no matter the breeding strategy in place. In Smallpop, breeding strategy 'Breeding cap' showed the smallest total number of runs with fixed alleles (74 runs), followed closely by 'Mean kinship' (76 alleles), and the reference showed the largest amount (146 runs).

As for the eliminated alleles, Midpop scored lower than Smallpop as well. The largest total number of runs with eliminated alleles in Midpop was found in breeding strategy 'Least related' (238 runs), and the smallest number was 133 runs in 'Mean kinship'. For Smallpop, the largest total number of runs with eliminated alleles was found in the reference (1147 runs) and the smallest number was 825 runs in 'Mean kinship'. The tables containing the number of eliminated and fixed alleles can be found in Appendices J and K for Smallpop and Midpop respectively.

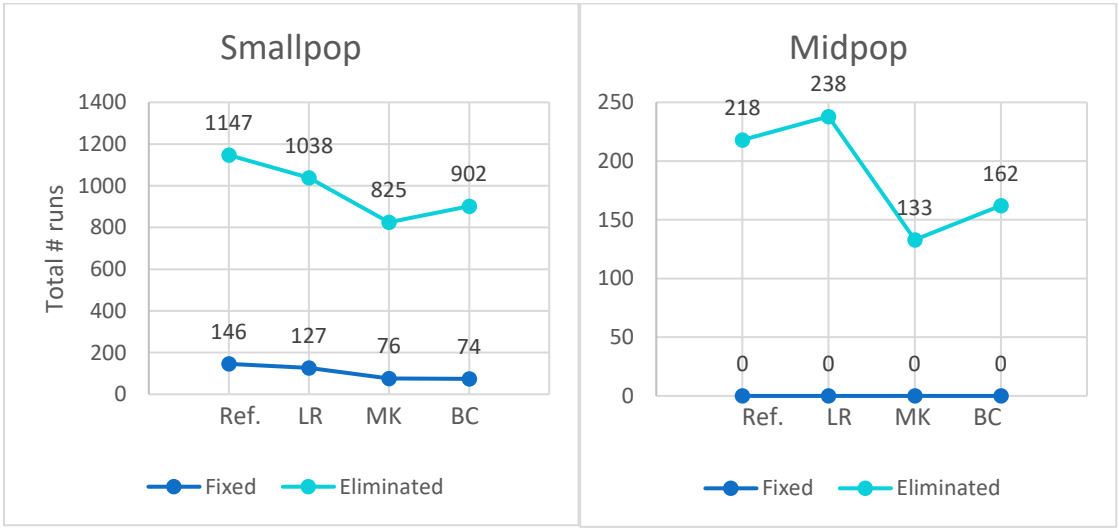


Figure 10: The total number of runs where a disease allele was fixed or eliminated per breeding strategy. The breeding strategies are Least Related (LR), Mean Kinship (MK), Breeding Cap (BC), and the reference population (Ref.) shown for both the small sized population (Smallpop, 15 males, 20 females) and medium sized population (Midpop, 112 males, 150 females).

3.3.3. BREEDING STRATEGIES & AGE OF ONSET OF DISEASE

The allele mix consisted of alleles with four different ages of onset of disease: 0, 1, 4, and 8 years old. To compare these different ages of onset, the total number of runs with eliminated and fixed disease alleles are shown for each breeding strategy per different age of onset in Tables 12 (for Smallpop) and 13 (for Midpop). The numbers are visually depicted in Fig. 11 - 13 The ages of onset 0 and 1 showed similar levels of eliminated alleles. At an age of onset of 4 years old, fewer runs showed eliminated alleles. An age of onset of 8 years old showed the least amount of runs with eliminated alleles. For Midpop, an age of onset of 8 years meant 0 runs with eliminated alleles, regardless of the breeding strategy in use. In short, these figures show that higher ages of onset went hand in hand with fewer runs with eliminated alleles. This trend could be observed for all breeding strategies. Of the three breeding strategies, 'Least related' consistently showed the highest number of runs with eliminated alleles, while 'Mean kinship' consistently showed the lowest number.

Table 12: The total number of runs with eliminated and fixed alleles in the small sized population (15 males and 20 females), shown for each breeding strategy and the reference population, divided by age of onset.

	Reference		**Least related**		**Mean kinship**		**Breeding cap**	
	Fixed	Eliminated	Fixed	Eliminated	Fixed	Eliminated	Fixed	Eliminated
Age of onset: 0	2	189	4	180	1	151	0	159
Age of onset: 1	6	177	3	171	1	152	1	165
Age of onset: 4	7	151	9	131	6	106	5	114
Age of onset: 8	29	72	23	56	20	33	19	44

Table 3: The total number of runs with eliminated and fixed alleles in the medium sized population (112 males and 150 females), shown for each breeding strategy and the reference population, divided by age of onset.

	Reference		**Least related**		**Mean kinship**		**Breeding cap**	
	Fixed	Eliminated	Fixed	Eliminated	Fixed	Eliminated	Fixed	Eliminated
Age of onset: 0	0	67	0	72	0	47	0	53
Age of onset: 1	0	69	0	63	0	44	0	56
Age of onset: 4	0	35	0	42	0	14	0	14
Age of onset: 8	0	0	0	0	0	0	0	0

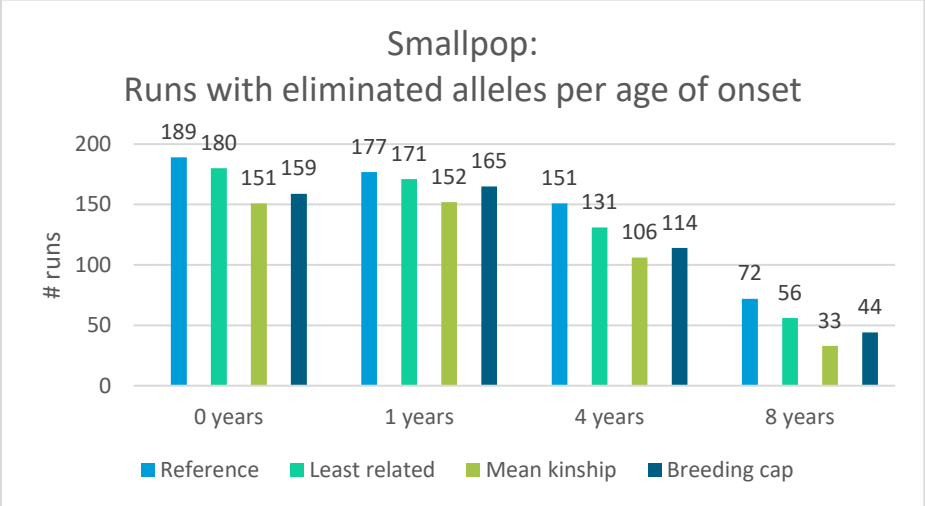


Figure 11: Number of runs with eliminated alleles per age of onset, for each breeding strategy, shown for the small sized population (15 males, 20 females).

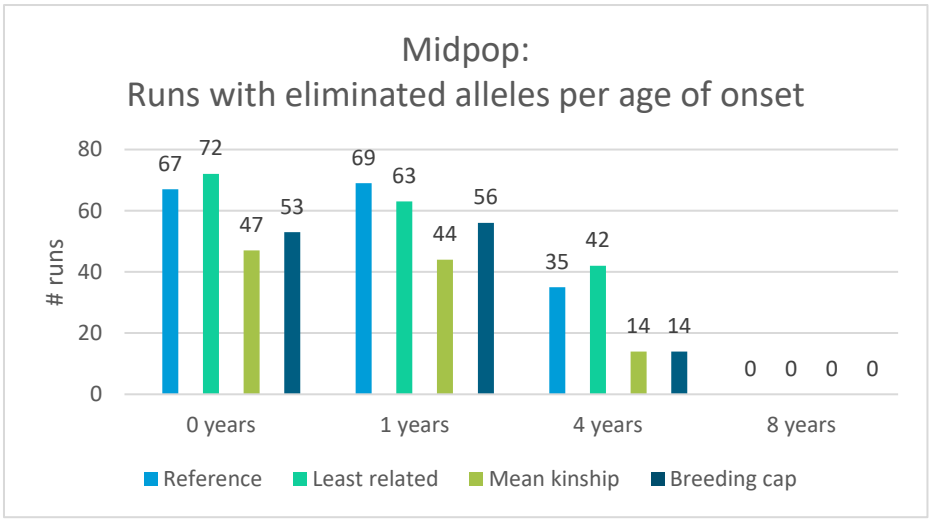


Figure 12: Number of runs with eliminated alleles per age of onset, for each breeding strategy, shown for the medium sized population (112 males, 150 females).

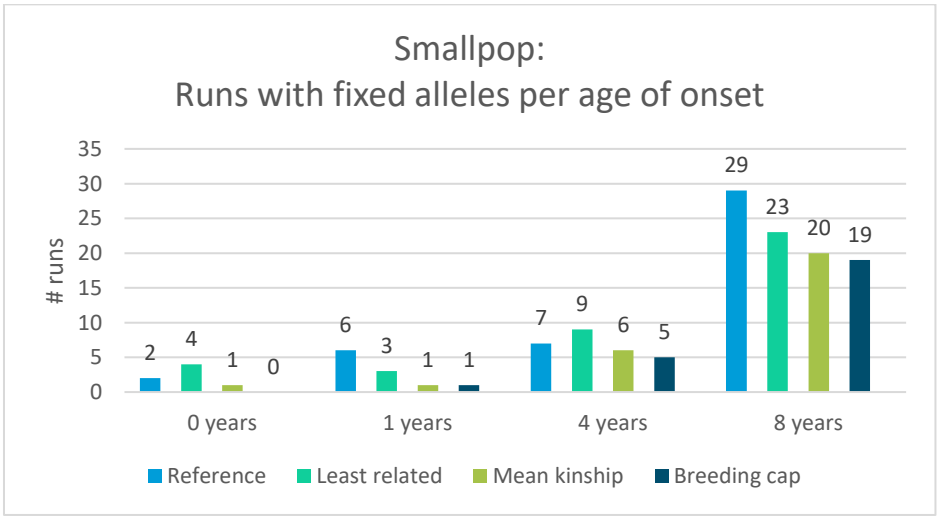


Figure 7: Number of runs with fixed alleles per age of onset, for each breeding strategy, shown for the small sized population (15 males, 20 females).

Contrary to what appeared to be the case for eliminated alleles, the number of runs with fixed alleles increased with a higher age of onset of disease (Fig. 13). This trend could be observed for all breeding strategies. At an age of onset of 8 years old, the breeding strategy 'Least related' showed the highest number of runs with fixed alleles (23 runs), whereas 'Breeding cap' showed the lowest number (19 runs), followed closely by 'Mean kinship' (20 runs).

3.3.4. BREEDING STRATEGIES & STRENGTH OF DISEASE

The allele mix contained alleles of six different disease strengths : 0.01, 0.05, 0.1, 0.25, 0.5, and 1. Disease alleles were more likely to be eliminated when disease strength was high (Fig. 14 & Fig. 15). This trend appeared in all breeding strategies. Of the three strategies, 'Least related' almost always showed the highest number of runs with eliminated alleles. 'Mean kinship' generally showed the lowest number of runs with eliminated alleles. In Midpop, disease alleles with low disease strengths (0.01, 0.05, 0.1) hardly experienced any elimination.

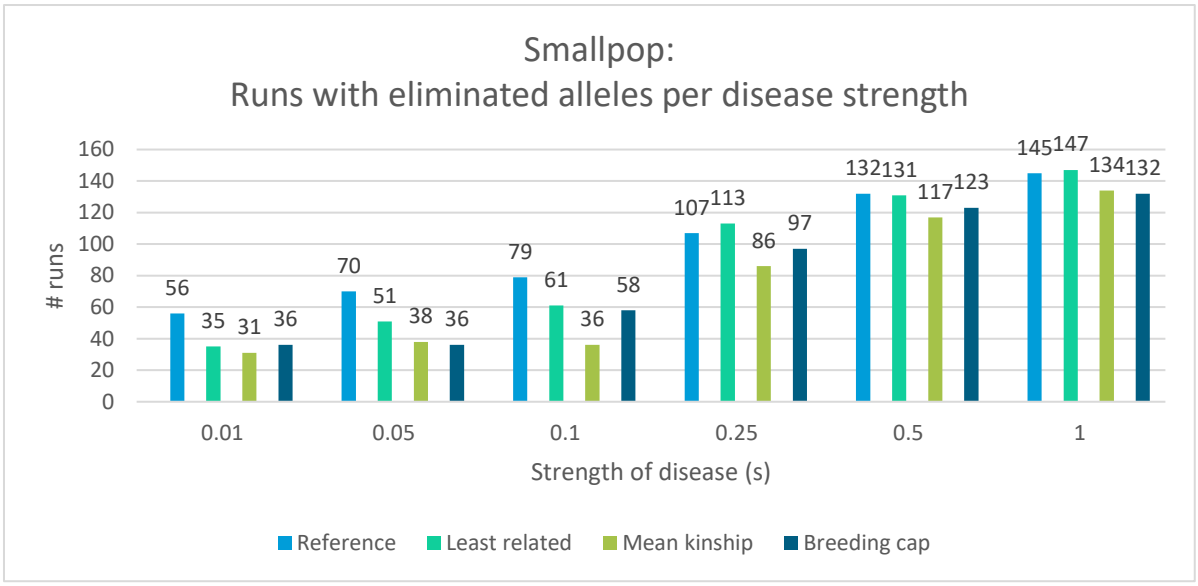


Figure 84: Number of runs with eliminated alleles per disease strength , for each breeding strategy, shown for the small sized population (15 males, 20 females).

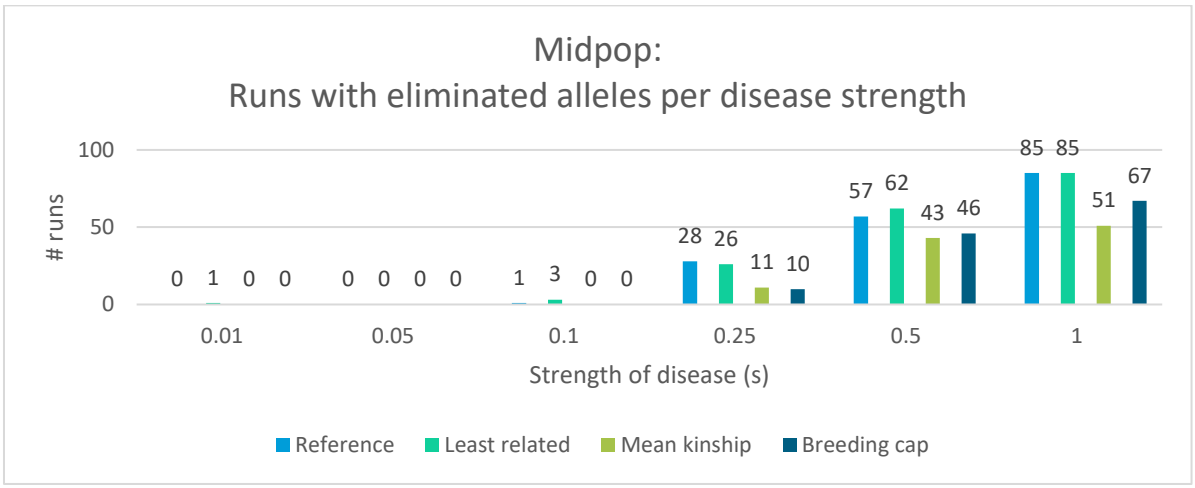


Figure 15: Number of runs with eliminated alleles per disease strength , for each breeding strategy, shown for the medium sized population (112 males, 150 females).

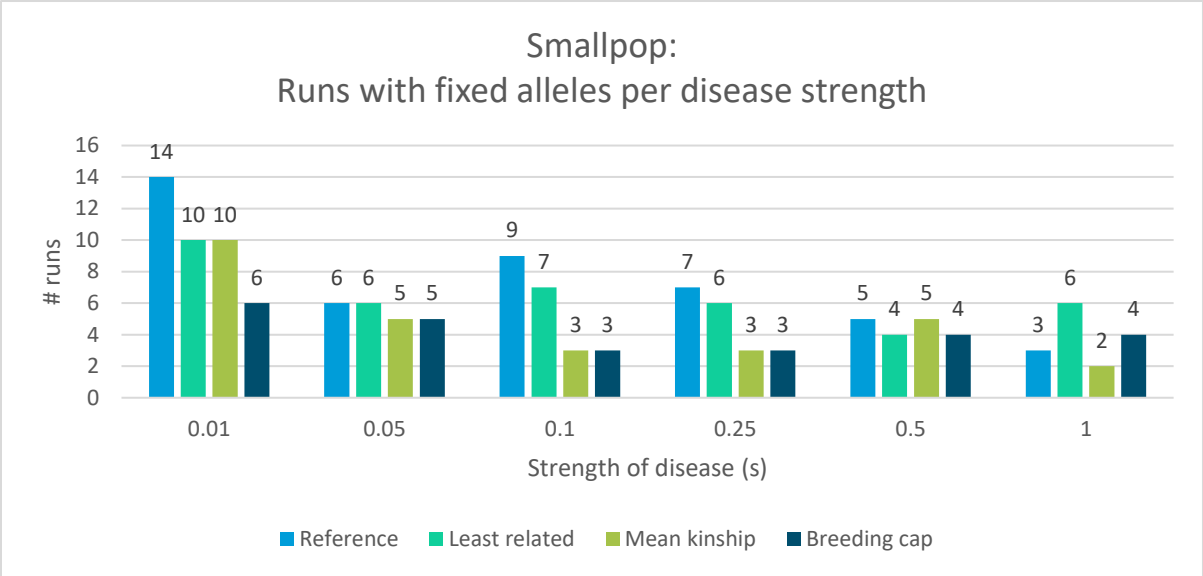


Figure 16: Number of runs with fixed alleles per disease strength , for each breeding strategy, shown for the small sized population (15 males, 20 females).

The number of runs with fixed alleles decreased a bit as the strength of disease increased (Fig. 16). However, the number of fixed alleles was generally quite low. It was usually highest in the reference population, followed by breeding strategy 'Least related'.

3.3.5. BREEDING STRATEGIES AND DISEASE ALLELE FREQUENCIES

The highest and lowest average allele frequencies were not consistently found in any one single breeding strategy (Tables 14 & 15). For Smallpop, the lowest disease allele frequency was often found in 'Breeding cap' (8 times), followed by 'Least related' (5), and 'Mean kinship' (4). The highest disease allele frequency was often found in 'Least related' (8), followed by 'Mean kinship' (6), and 'Breeding cap' (4). Note that the reference had the lowest disease allele frequency 6 times, and the highest disease allele frequency 5 times.

For Midpop, the lowest disease allele frequency was often found in 'Breeding cap' (9 times), followed by 'Mean kinship' (7), and 'Least related' (5). The highest disease allele frequency was often found in 'Breeding cap' as well (9), followed closely by 'Least related' (8), and then by 'Mean kinship' (4). Note that the reference had the lowest disease allele frequency 4 times, and the highest disease allele frequency 3 times.

Table 14: Smallpop, average disease allele frequency for breeding strategies 'Least related', 'Mean kinship', and 'Breeding cap', and the reference. Green and blue mark respectively the lowest and the highest frequencies found for a breeding strategy. The alleles are described on the left with their number, strength (s), and age of onset (y).

Smallpop (alleles)	*Reference*	***Least related***	***Mean kinship***	***Breeding cap***
1-2 (s=0.01, y=0)	0.338	0.375	0.385	0.389
3-6 (s=0.01, y=1)	0.369	0.370	0.367	0.345
7-10 (s=0.01, y=4)	0.382	0.407	0.431	0.397
11-14 (s=0.01, y=8)	0.430	0.414	0.477	0.433
15-16 (s=0.05, y=0)	0.095	0.254	0.207	0.207
17-20 (s=0.05, y=1)	0.276	0.274	0.267	0.218
21-24 (s=0.05, y=4)	0.303	0.417	0.359	0.393
25-28 (s=0.05, y=8)	0.389	0.413	0.432	0.381
29-30 (s=0.1, y=0)	0.124	0.161	0.136	0.115
31-34 (s=0.1, y=1)	0.193	0.210	0.185	0.161
35-38 (s=0.1, y=4)	0.363	0.331	0.310	0.371
39-42 (s=0.1, y=8)	0.394	0.425	0.426	0.355
43 (s=0.25, y=0)	0.067	0.063	0.047	0.052
44-45 (s=0.25, y=1)	0.09	0.054	0.054	0.049
46-47 (s=0.25, y=4)	0.132	0.113	0.131	0.158
48-49 (s=0.25, y=8)	0.433	0.448	0.436	0.411
50 (s=0.5, y=0)	0.004	0.007	0.017	0.019
51-52 (s=0.5, y=1)	0.028	0.019	0.011	0.013
53-54 (s=0.5, y=4)	0.143	0.081	0.135	0.087
55-56 (s=0.5, y=8)	0.427	0.470	0.459	0.436
57 (s=1, y=0)	0.013	0.007	0.018	0.008
58 (s=1, y=1)	0.017	0.007	0.024	0.009
59 (s=1, y=4)	0.036	0.031	0.020	0.031
60 (s=1, y=8)	0.403	0.481	0.432	0.472

Table 15: Midpop, average disease allele frequency for breeding strategies 'Least related', 'Mean kinship', and 'Breeding cap', and the reference. Green and blue mark respectively the lowest and the highest frequencies found for a breeding strategy. The alleles are described on the left with their number, strength (s), and age of onset (y).

Midpop (alleles)	*Reference*	***Least related***	***Mean kinship***	***Breeding cap***
1-2 (s=0.01, y=0)	0.322	0.330	0.332	0.310
3-6 (s=0.01, y=1)	0.357	0.355	0.336	0.343
7-10 (s=0.01, y=4)	0.403	0.410	0.401	0.415
11-14 (s=0.01, y=8)	0.421	0.427	0.409	0.423
15-16 (s=0.05, y=0)	0.172	0.164	0.180	0.146
17-20 (s=0.05, y=1)	0.190	0.206	0.200	0.186
21-24 (s=0.05, y=4)	0.361	0.311	0.360	0.357
25-28 (s=0.05, y=8)	0.427	0.448	0.320	0.421
29-30 (s=0.1, y=0)	0.109	0.101	0.111	0.092
31-34 (s=0.1, y=1)	0.130	0.142	0.122	0.113
35-38 (s=0.1, y=4)	0.285	0.282	0.316	0.318
39-42 (s=0.1, y=8)	0.420	0.435	0.404	0.442
43 (s=0.25, y=0)	0.044	0.050	0.043	0.036
44-45 (s=0.25, y=1)	0.055	0.057	0.053	0.037
46-47 (s=0.25, y=4)	0.115	0.096	0.096	0.126
48-49 (s=0.25, y=8)	0.402	0.422	0.399	0.429
50 (s=0.5, y=0)	0.028	0.023	0.026	0.013
51-52 (s=0.5, y=1)	0.017	0.029	0.021	0.019
53-54 (s=0.5, y=4)	0.085	0.072	0.084	0.089
55-56 (s=0.5, y=8)	0.220	0.390	0.417	0.433
57 (s=1, y=0)	0.010	0.017	0.021	0.014
58 (s=1, y=1)	0.014	0.024	0.022	0.012
59 (s=1, y=4)	0.030	0.037	0.031	0.047
60 (s=1, y=8)	0.421	0.359	0.381	0.427

However, to get a better look at the effect of the different breeding strategies on the disease alleles, it was important to zoom in and see how the allele frequencies changed over the course of 100 years. Two types of graphs could be distinguished. Erratic graphs showed haphazard lines that fly up and down. They reached higher and lower frequencies and they reached them sooner. As a result, disease alleles tended to get fixed or eliminated faster than in centered graphs. Centered graphs showed allele frequencies that stayed much closer to the average of all runs (black line). As a result, disease alleles tended to take longer to get fixed or eliminated. For example, the first disease allele was permanently eliminated in year 23 in the erratic Smallpop graph, but only in year 34 in the centered Smallpop graph. While no disease alleles were eliminated in the Midpop graphs, these graphs do show the difference between erratic and centered very clearly.

The three breeding strategies consequently differed in the type of graph they showed. First of all, the reference consistently showed the most erratic graphs. Of the breeding strategies, 'Least related' showed the most erratic graphs, whereas 'Mean kinship' consistently showed the most centered graphs. 'Breeding cap' tended towards erratic graphs.

Two examples of erratic graphs are the allele frequencies found for allele 17-0051 (strength=0.05, age of onset=1) for Smallpop and Midpop that belong to breeding strategy 'Least related' (Fig. 17 & Fig. 18). An example of a centered graph are the allele frequencies found for allele 17-0051 for Midpop that depicts breeding strategy 'Mean kinship' (Fig. 19).

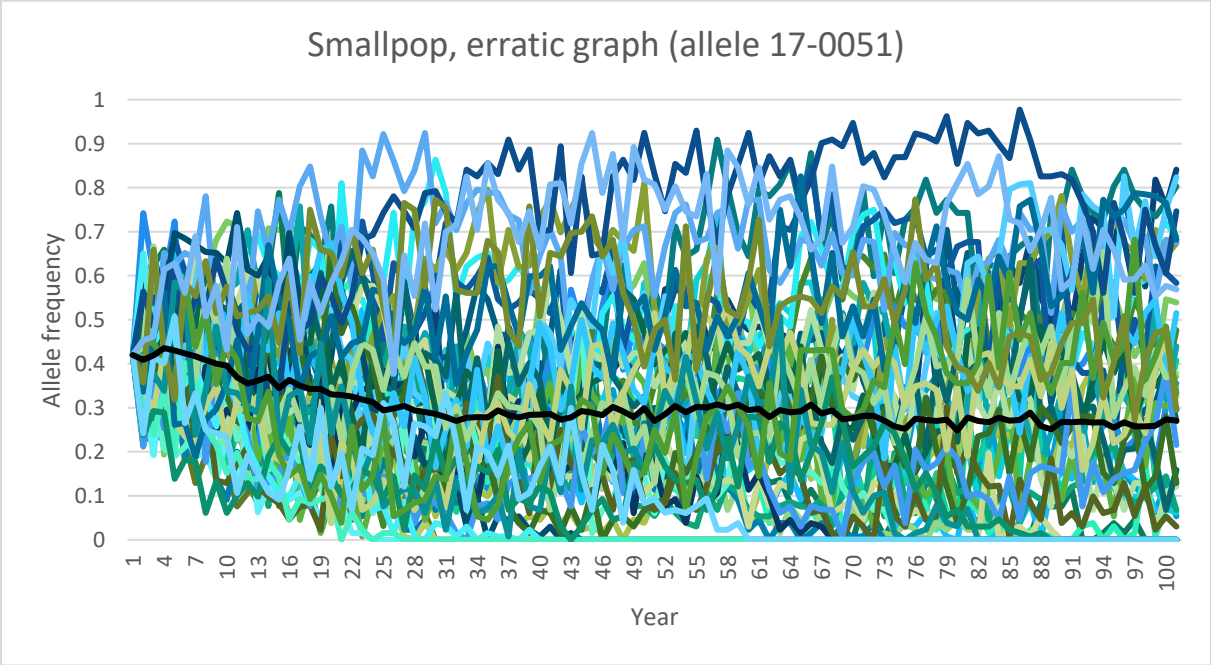


Figure 17: A Smallpop graph with erratic allele frequency lines, shown for 50 runs over the course of 100 years. This graph shows allele 17-0051 for breeding strategy 'Least related'.

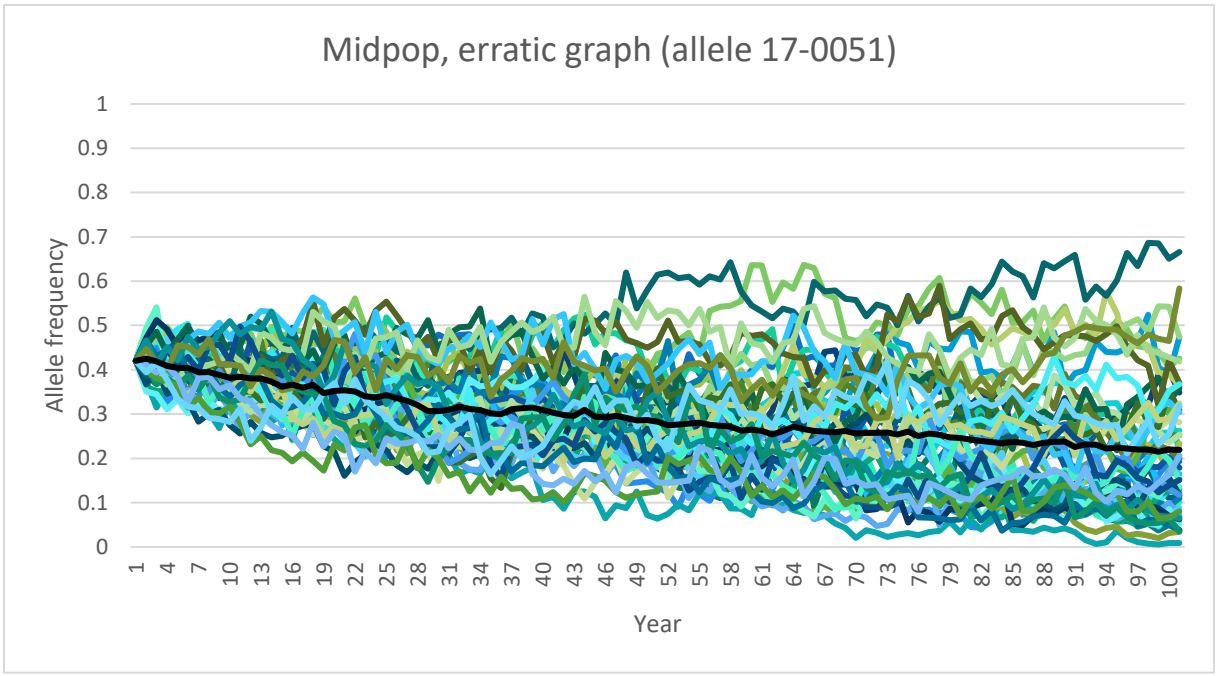


Figure 18: A Midpop graph with erratic allele frequency lines, shown for 50 runs over the course of 100 years. This graph shows allele 17-0051 for breeding strategy 'Least related'.

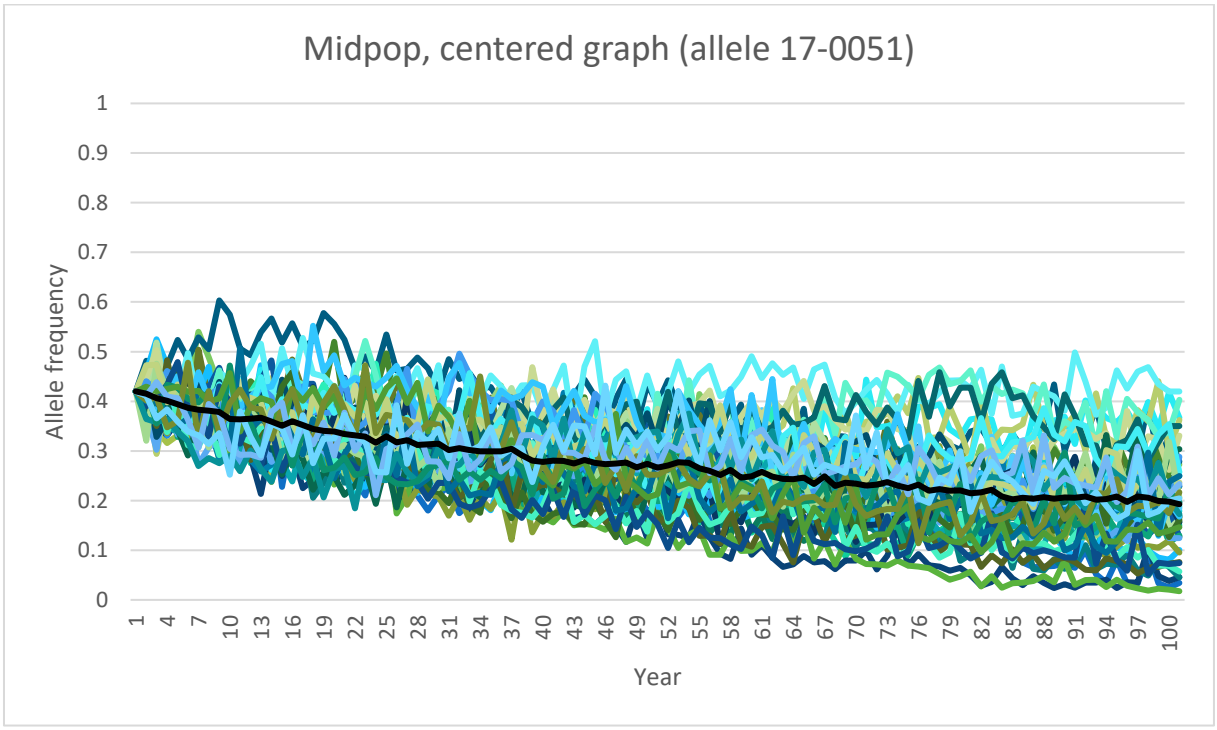


Figure 19: A Midpop graph with centered allele frequency lines, shown for 50 runs over the course of 100 years. This graph shows allele 17-0051 for breeding strategy 'Mean kinship'.

4. DISCUSSION

4.1. RECAP

The aim of this research was to find out how different kinds of genetic management strategies affect the allele frequencies of genetic defects, and see which strategy reduces the number of defects most effectively. The sub-questions focused on the effects of the age of onset of a disease allele, strength of a disease allele, and population size. And, because every individual carries many disease alleles, whether disease alleles behave differently when simulated as a group rather than individually.

The findings suggest that genetic defects with an early age of onset, and defects with a strong disease strength, are more likely to be eliminated from a population over the course of 100 years than defects with a late onset or having a weak disease strength. Disease alleles did not appear to behave differently when simulated as a single allele as opposed to when simulated in a group. Furthermore, the data also indicate that populations with a small effective population size are affected more by drift than populations with a larger effective population size. The genetic management strategy 'Exclude carriers' focused on selection to eliminate defects, which it did successfully if the population did not go extinct—but it did go extinct almost every single time. The strategies 'Mean kinship', 'Least related', and 'Breeding cap' never went extinct. They were less successful in eliminating defects, but they were successful in preventing fixed disease alleles, keeping overall disease allele frequencies low, and keeping inbreeding levels lower.

4.2. RESULT ANALYSIS

4.2.1. STRENGTH OF DISEASE AND AGE OF DISEASE ONSET

The data in this thesis indicate that higher ages of onset and lower strength of disease result in higher disease allele frequencies: the higher the age of onset of disease alleles, the more disease alleles are found in the population. This includes an increased number of fixed alleles and a decreased number of eliminated alleles. The same is true for disease alleles with a low strength. In other words, the results of this thesis suggest that when the age of onset of a disease is high and/or when the strength of disease is low, disease alleles are more likely to remain in the population and may even become fixed. Keightley et al (1998) supports these findings with experimentations, stating that mutations with a strong effect are quickly eliminated from a population, and that the mutations with a smaller deleterious effect are a far greater worry because the force of selection against them is weaker, and so populations have trouble ridding themselves of these alleles with a weak disease strength.

Two recently published studies present ideas regarding age of disease onset and selection pressures. Pavard and Coste (2021) looked at alleles that are expressed after female reproduction has ended. They state that alleles coding for late onset diseases are often considered neutral in scientific literature, which is also in line with the results found in this thesis. However, when they calculated selection coefficients for late onset diseases using an evolutionary demography model in humans, taking sociocultural factors into account, their results showed that there is negative selection against late onset diseases. They conclude that neutrality is probably the exception for late onset diseases, when you take sociocultural factors into consideration. In other words, even if deleterious alleles take effect after reproduction, they may still be experiencing selection pressure (Pavard & Coste, 2021). Though said study was done with a human population in mind, it might be translatable to dogs as well. For instance, breeders may look at disease occurrence in ancestors and use that to help guide breeding decisions, which adds a negative selection pressure. Or if mothers die soon after birth, the fitness of their offspring is lower due to the absence of maternal care. Still, the simulations in this thesis suggest that even if selection pressure is high (selection coefficient of 1), the alleles for a late onset disease still behave almost like a neutral allele—as if they are not affected by negative selection pressure. Rapaport et al (2021) looked at factors that impact the intensity of negative selection in humans. In line with the findings of

this thesis, that when the age of disease onset is low and when the strength of disease is high, disease alleles are strongly affected by selection. They found that genes experience stronger negative selection pressure when the age of disease onset is before reproductive maturity, and there is also a strong negative selection pressure when the disease comes with a loss of function (Rapaport et al., 2021).

In this thesis, the simulated management strategies appeared to be effective at reducing average disease allele frequencies, reducing the number of fixed alleles, and mostly effective at keeping inbreeding levels lower. However, none of the strategies seemed very effective at reducing disease alleles with a higher age of onset or lower strength of disease. The studies discussed in this chapter so far provide an idea as to why this is. For instance, they indicate that whether an individual is at the end of its reproductive age is an important factor impacting selection against a disease allele. Selection pressures may have already removed individuals with an early onset disease before they get to reproduce and pass on their genes. And if disease strength is high, strong selection pressures also remove the individual from the population. In both of these cases, disease allele frequency in the population will decrease due to the negative selection pressure. However, when the age of disease onset is high, individuals will already have reproduced and passed on their genes by the time the disease is expressed. And when the strength of disease is low, a disease allele is affected very little by selection pressures. In both of these cases, the disease allele frequency in the population is either unaffected or decreases very little. This means that in the long term, deleterious alleles with a small effect and alleles with a late age of onset could have a much larger impact on the genetic health of a population than strongly deleterious alleles and alleles that come to expression early in life.

4.2.2. GENETIC DRIFT AND POPULATION SIZE

When selection has little to no effect on disease alleles, the changes you see in allele frequencies are mainly caused by genetic drift. In other words, any observed change in allele frequency may simply be due to random chance (Masel, 2011). Under the influence of genetic drift, alleles with a small deleterious effect may become fixed, especially in conjunction with high inbreeding rates. This effect is most strongly seen in small populations (Keightley et al., 1998; Masel, 2011). Due to genetic drift, the frequency of any one allele might increase or decrease, causing it to become fixed or eliminated. This is potentially harmful to a population if a fixed allele is deleterious or if an eliminated allele is associated with survival (Odom, 2022).

The results of this thesis showed that the population with a small effective population size was affected more by genetic drift than the medium sized population. This observation is exactly in line with what the theory predicts and what other studies have found (Keightley et al., 1998; Lynch et al., 1995). A recent study by Mathur & DeWoody in 2021 compared the genetic load of deleterious mutations of a small effective population with a large effective population. Though their study looked at a wild species (Montezuma quail) and not a captive species, it yielded similar results. It found that the large effective population carried more deleterious mutations (potential load), but the small effective population had more fixed deleterious mutations with a small effect (realized load). The study named genetic drift and increased inbreeding levels as the driving forces behind the increased realized load (Mathur & DeWoody, 2021). Due to the increased genetic load, a small sized population may become caught in a so called ‘mutational meltdown’ (Keightley et al., 1998). Deleterious mutations with a small effect accumulate in a small sized population, which increases inbreeding depression, which decreases the effective population size, which in turn further increases the probability of fixation of deleterious mutations through genetic drift. This cycle eventually results in the extinction of the population. The ‘extinction vortex’ (Gilpin & Soulé, 1986) describes a similar path to extinction, and Fagan & Holmes (2006) adds that the time to extinction depends on the effective population size. The smaller the effective population size, the shorter the road to extinction.

The effective population size of many dog breeds is unfortunately quite small. Wijnrocx et al (2016) analysed 23 dog breeds ranging from popular (Labrador retriever, Australian shepherd, German shepherd) to rare (Belgian Laekenois, Griffon Belge, Bouvier des Ardennes). They found that the effective population sizes ranged from

829.1 (Australian shepherd) to 3.2 (Bouvier des Ardennes). Of the 23 populations, 22 had an effective population size smaller than 500 individuals. Of these, 16 effective populations were smaller than 100 individuals, and 9 breeds had less than 50 individuals (Wijnrocx et al., 2016). The two populations in this thesis had an effective population size of 34.3 (small sized population) and 256.5 (medium sized population), which falls within the range found by Wijnrocx et al (2016).

In wild animal populations, the minimum effective population size to ward off inbreeding depression and maintain long term evolutionary potential—a population's adaptive ability—is known as the 'minimum viable population size', or MVP (Rosenfeld, 2014; Frankham et al., 2014; Jamieson & Allendorf, 2012; Flather et al., 2011; Traill et al., 2011). Originally, the MVP was viewed as the the smallest effective population size necessary to keep an isolated population extant for 1000 years with an extinction risk of 1%. Nowadays the timeline for a population to remain extant is 100 years with an extinction risk of 5% (Flather et al., 2011). As a rule of thumb, the MVP is set at 50 individuals to prevent inbreeding depression and 500 individuals to keep a population alive in perpetuity (Jamieson & Allendorf, 2012). However, Traill et al (2011) suggests that this rule of thumb may be far too conservative, and that not hundreds, but rather thousands of individuals are required to prevent a population from going extinct. Frankham et al (2014) state that an effective population size of 50 does not prevent inbreeding depression. Instead they propose that an effective population size of 100 will limit inbreeding depression to 10% over 5 generations, and an effective population size of 1000 individuals is required to keep a population viable in the long term (Frankham et al, 2014). The Centre of Genetic Resources in the Netherlands (CGN) uses a 'traffic light' system that indicates the risk status of a captive population (Oldenbroek, 2016). An effective population size of less than 50 individuals runs the risk of extinction through an accumulation of genetic defects; between 50-100 individuals it is certain that genetic defects will occur; between 100-200 individuals there is a chance that genetic defects will occur; and over 200 individuals means a small chance that genetic defects will occur. This means that out of the 23 breeds analysed by Wijnrocx et al (2016) only 7 breeds may avoid inbreeding depression, and only 1 breed is viable in the long term. Similarly, the two populations simulated in this thesis would not survive in perpetuity by either the 50/500 or 100/1000 standard, but the medium sized population does meet the CGN's standard. Furthermore, the small sized population would not be expected to stave off inbreeding depression.

Of course, MVPs are used in conservation genetics to prevent species from going extinct. While dog breeds are isolated, there is always the option of genetic rescue by introduction of individuals from outside the breed—in other words, through outcross (Lewis & Windig, 2017; Yordy, 2019; Windig & Doekes, 2018).

4.2.3. INBREEDING LEVELS AND INBREEDING DEPRESSION

As previously stated, genetic drift and inbreeding levels are driving factors behind an increase in fixed deleterious mutations in a population (Mathur & DeWoody, 2021). Inbreeding is unavoidable in closed populations (Wright et al., 2008). Purebred dog breeds are dog populations bred within a closed gene pool, and as such are considered an isolated population. Lewis & Windig (2017) state that most dog breeds originate from a small number of founders, have known high levels of inbreeding since their creation in order to fix breed typical traits, and have small effective population sizes. Only a few animals are chosen to pass on their genes to the next generation. When a population is large enough, the worst deleterious alleles are purged from a population through natural selection, and most other deleterious alleles linger in low frequencies and only in heterozygous states where they do not cause an issue. However, deleterious alleles may reach high frequencies when only very few animals contribute to the next generation, as is the case with purebred dogs, especially in breeds with small effective population sizes. Frequencies may rise so quickly that natural selection is unable to purge the defects (Lewis & Windig, 2017). This is likely what happened in the simulated populations in this thesis where a deleterious allele ended up fixed.

Purebred dogs have historically known high selection pressures from breeders, gone through bottlenecks, and the effect of breed management practices such as the overuse of popular sires and deliberate close inbreeding.

This is why dog breeds show diminished genetic diversity compared to other species (Leroy, 2011). Paired with a small effective population size, the high rates of inbreeding increase genetic drift, which can cause high frequencies and even fixation of recessive deleterious alleles (Lewis & Windig, 2017). When deleterious alleles are fixed for fitness traits, inbreeding depression is seen in a population, and inbred populations have a greater extinction risk (Wright et al., 2008; Hedrick & Garcia-Dorado, 2016; Charlesworth & Willis, 2009; Yordy et al., 2019).

In short, high rates of inbreeding are very undesirable, but unfortunately many dog breeds do know high inbreeding rates. The FAO recommends a rate of inbreeding of 0.5% - 1% per generation to avoid the aforementioned negative consequences of inbreeding (Lewis & Windig, 2017). The aforementioned 'traffic light' system of the Centre of Genetic Resources in the Netherlands notes that an inbreeding rate of >1% guarantees extinction due to accumulated genetic defects, and <0.25% means there is only a small chance of genetic defects. In this thesis, the population with a medium effective population size complied with this recommendation in all simulated breeding strategies except 'Exclude carriers' (which was a rate of 2.5% in the single surviving population). However, the small effective population size did not meet this recommendation at 2.6% ('Least related'), 1.6% ('Mean kinship'), and 1.7% ('Breeding cap'). The rate of inbreeding is still too high, which is another indicator that the effective population size needs to increase to ward off the negative consequences of inbreeding and genetic drift and in order to keep the population viable in the long run.

The average inbreeding level of 112 dog breeds was 31%, ranging between 6.5% for the Sloughi and 86.8% for the Norwegian Lundehund (Dreger et al., 2016). Important to note is that Dreger et al looked at genetic COI, while the program Pointer used for this thesis looked at pedigree COI, which tends to be less accurate and shows lower values. This is because pedigree COI assumes that each allele is passed down equally (50%), while in reality this is not the case. Pedigree COI also depends on the number of generations included in the calculations. The data files for Dreger et al's research show what these differences look like in practice. Take for instance the Labrador retriever: Pedigree COI for 5 generations is 2.6%, for 10 generations is 7.3%, and for all generations is 8.2%, while the genetic COI is 21.1% (Dreger et al., 2016).

In this thesis, the inbreeding level of the medium sized population after 100 years of simulations is 9.8% (Least related), 6.5% (Mean kinship), and 6.9% (Breeding cap). In contrast, the inbreeding levels of the small sized population is very high after 100 years: 51.3% (Least related), 35.8% (Mean kinship), and 42% (Breeding cap). Note that this is pedigree COI, so in reality the inbreeding levels are likely much higher. You would expect a population with such high inbreeding levels to experience the effects of inbreeding depression. As it happens, the Saarloos wolfdog –which the small sized population in this thesis is based on— does in fact suffer from the effects of inbreeding depression, which is why they opened their breed's gene pool and started genetic rescue through an outcross program (Eggink & Oldenbroek, 2014).

4.2.4. GENETIC MANAGEMENT STRATEGIES

The genetic management strategy that was most effective at reducing deleterious alleles was 'Exclude carriers'. This was the strategy that focused on selection to reduce disease alleles. However, the selection pressure was so high that all of the small sized populations went extinct, and only one single medium sized population survived. As previously mentioned, the rate of inbreeding in this population was very high at 2.5% per generation, and the level of inbreeding was also very high at 51.8%. The reason all small sized populations and most medium sized populations went extinct is because no animal was left to breed with after the initial severe purge of all individuals carrying a disease allele. A 2012 study by Leroy and Rognon (2012) supports the idea that excluding carriers is not a good idea in a small effective population size, as it has a negative effect on genetic diversity (Leroy & Rognon, 2012).

Taking all the aforementioned studies into consideration: that dog breeds exist having a limited gene pool, with a small effective population size, and with already high average inbreeding levels, it seems unwise to focus on

selection to reduce all disease alleles. The increased selection pressure would add further strain on the effective population size and the level of inbreeding. Selection by eliminating individuals means a decrease in effective population size, increasing drift and inbreeding levels, which increases the chance of weak disease alleles reaching high frequencies and even becoming fixed. Selection is only effective against the strong disease alleles and early onset disease alleles, while drift influences disease alleles with a small effect and late onset disease alleles. Eliminating one allele or a few alleles through selection is possible, but only if the effective population size is large enough.

The other three strategies in this thesis focused on inbreeding rate and therefore reducing drift. According to Leroy (2011) popular sires and close inbreeding are the most important breeding practices that reduce diversity within a breed. Popular sires are particularly problematic and have led to the spread of a large number of deleterious alleles (Leroy, 2011). Putting a limit on the amount of litters a male can sire, therefore seems like a practical idea. The strategy 'Breeding cap' put a limit of 2 litters per sire per lifetime, and was effective at increasing the population size, and decreasing the rate of inbreeding, average inbreeding level after 100 years, decreasing the number of fixed alleles, and decreasing the average disease allele frequencies. The strategies 'Least related' and 'Mean kinship' did the same things, but out of all three strategies, 'Mean kinship' did it best. This comes with little surprise, as this breed management technique is touted as one of the most effective strategies by Lewis & Windig (2017), coming only after optimal contributions (see next subchapter: Limitations). While these strategies did not focus on eliminating alleles, many disease alleles were still eliminated all the same. This is likely because there are two types of selection at play: artificial selection performed by breeders, and natural selection that always acts against alleles with a negative effect on fitness (Keightley et al., 1998). This effect is also seen in the reference population, where no breeding strategy was applied but where disease allele frequencies went down and where many disease alleles were eliminated.

The results of this thesis, supported by the studies mentioned in this chapter, therefore suggest that focusing on reducing inbreeding and drift successfully prevents deleterious mutations from fixing, brings down disease allele frequencies and keeps them at low frequencies where they pose no harm in a heterozygous state. Selection against disease alleles provides elimination in the short term, but at the cost of increasing inbreeding levels and decreasing the effective population, thereby increasing drift and increasing the chances of fixing disease alleles. Whereas reducing drift by increasing the effective population size provides a decrease in disease allele frequencies and helps prevent fixation of disease alleles, but at the cost of less disease alleles eliminated in the short term.

4.3. LIMITATIONS

This thesis looked at four different types of genetic management strategies:
'Least related', where individuals are mated that are as unrelated to each other as possible;
'Mean kinship', where individuals are allowed to mate when they have a mean kinship with all the other animals in the population that is less than or equal to the average mean kinship in the breed;
'Breeding cap', which limits the number of times a male is used for breeding;
and 'Exclude carriers', which excludes individuals that carry an allele for a genetic defect.

However, there are more types of genetic management strategies that could be interesting to look at, such as optimal contributions and outcross. Optimal contributions is theoretically the best strategy to restrict inbreeding rates in closed populations, but because it requires complete control over all breeding in a population it is nearly impossible to achieve in purebred dogs (Lewis & Windig, 2017). While it is possible to simulate optimal contributions in Pointer, the program used for this thesis, I chose not to simulate optimal contributions due to its current unfeasibility in real life.

The disease alleles in this thesis only affected survival traits and not fertility traits, though that is also a possibility in Pointer. The expectation is that the results would be roughly the same for fertility traits, as fertility

is a fitness trait just like survival is. Therefore it stands to reason that the force of natural selection would act against deleterious fertility alleles as well. But because a deleterious survival trait results in death, which is a far more severe consequence of a deleterious allele reaching a homozygous state than the inability to procreate, the time that deleterious fertility alleles are eliminated probably takes a bit longer. In other words, I expect deleterious fertility alleles to behave like deleterious alleles with a low strength, as simulated in this thesis.

I only looked at monogenic recessive diseases in this thesis, and not at polygenic diseases. A polygenic disease is often comprised of many different alleles, each likely with a small effect, though little is known yet about the inheritance of these complex traits (Lewis & Windig, 2017). I therefore chose to focus only on monogenic recessive diseases.

This thesis focuses on a small sized population and a medium sized population. While a large sized population is possible to simulate in Pointer, such simulations take far longer for Pointer to process. In addition to that, most dog breed populations have relatively small effective populations ((Lewis et al., 2015; Wijnrocx et al., 2016; Dreger et al., 2016). The rough results of the large sized population may be rather easily deduced, however. As the effective population size increases, the effect of genetic drift decreases. This means that a large sized population is likely to see lower average inbreeding levels, a lower rate of inbreeding, no fixed disease alleles, fewer eliminated disease alleles and strong disease alleles take longer to eliminate, more disease alleles lingering in the population at very low frequencies, and less fluctuation in the allele frequencies. The only problems for a large population may be expected when a population experiences a bottleneck event, where the resulting smaller population ends up with an increased load of deleterious alleles. These deleterious alleles used to linger at low frequencies when the population was large, but may reach high frequencies after the bottleneck event (Yordy, 2019; Mathur & DeWoody, 2021).

Lastly, I did not look specifically at purging as a breeding strategy, though the phenomena was at play during the simulations. Purging is when a deleterious allele is removed from the population through inbreeding, drift, or natural selection (Mathur & DeWoody, 2021; Lewis & Windig, 2016; Hedrick & Garcia-Dorado, 2016). Natural selection purges highly deleterious alleles, as could be seen in the reference simulations in this thesis where no breeding strategy was applied, but where the frequencies of disease alleles went down regardless. Drift can purge an allele through chance, as could be seen in several simulations in the small sized population where a disease allele disappeared over time. And inbreeding can purge deleterious alleles by exposing them in a homozygous state to selection pressures. However, induced purging by inbreeding is not without danger, and therefore it is a questionable technique to improve a population's health (Hedrick & Garcia-Dorado, 2016).

4.4. RECOMMENDATIONS

The simulations for 'Exclude carriers' were largely unsuccessful in this thesis, with almost all the populations going extinct. While the selection simulated in this strategy was very rigorous, selection against defects is widely practices by dog breeders, for instance with the use of genetic tests to identify individuals with known disease alleles (Lewis & Windig, 2017). The results founds in this thesis suggest that selection against defects by removal of carrier individuals reduces the effective population size and increases the effect of drift. It would be interesting to look at ways to both select against defects yet also maintain a sufficiently large effective population. The aim here would be to aid the elimination of disease alleles, yet also keep the effects of drift at bay.

Despite the potential for opposition within the breed communities, outcross is a strategy that successfully reduces inbreeding levels in a breed by opening up the gene pool, though it requires extra measures in conjunction with outcross to prevent a return to previous inbreeding rates the moment the outcross project concludes (Lewis & Windig, 2017). At this time of writing, several breeds in the Netherlands and elsewhere have outcross projects underway, such as the Norwegian lundehund, Saarloos wolfdog, and Wetterhoun (Melis et al, 2022; AVLS, 2021; NVSW, 2022). The Saarloos wolfdog outcross program started in 2012 and was

analysed in 2018 by Windig and Doekes, who found that the inbreeding rate was indeed reduced after breeding outcross litters. The inbreeding rate saw a smaller reduction when the outcross was followed by backcrosses, and continuous outcrosses were more effective than a single or few outcrosses. Their research also found that introduction of severely deleterious alleles through outcross occurred rarely, while introduction of neutral alleles or slightly deleterious alleles occurred more frequently. The study concluded that outcross is most effective when done continuously, and that it can also be used to buy time to increase the effective population size (Windig & Doekes, 2018).

The results of this thesis suggest that the weak disease alleles in a breed with a small effective population size are much affected by drift, which causes increased fluctuation of disease allele frequencies and increases the risk of fixed alleles. Fixed alleles are impossible to get rid of in a closed gene pool. When too many disease alleles with a small effect are fixed it may cause inbreeding depression, which shows as lack of fertility and vitality. This is what has likely happened in the aforementioned breeds (Lundehund, Saarloos wolfdog, wetterhoun), where breeders noticed an increase in unsuccessful breedings, empty females, and small litters. Outcross opens up the gene pool and provides an influx of new alleles, which can bring immediate relief with regard to infertility and vitality issues (Lewis & Windig, 2017). An example of this is with the Saarloos wolfdog outcross program. A Saarloos wolfdog dam was bred to a Siberian husky sire. The Saarloos dam birthed a single F1 puppy. This F1 outcross was bred to a Saarloos wolfdog male, and gave birth to a litter of 10 puppies (AVLS, 2020). It is possible to simulate outcross in Pointer, so this is an interesting area of exploration for future research.

The theoretically most effective strategy is optimal contributions. Despite this management strategy being not feasible at this point in time, due to its success in other species it is worth further exploration. Optimal contributions requires information on every animal in a population and full cooperation of breeders. If this hurdle can be taken, the next step is to create an index with weighted breeding traits that include heritable diseases. A tentative first exploration was done for the Norwegian lundehund to see how it might help the breed (Kettunen et al., 2017; Lewis & Windig, 2017).

Of the simulated genetic management strategies, ‘Mean Kinship’ is most effective at keeping inbreeding levels low and thereby genetic drift as small as possible, and also maintaining low disease allele frequencies. Implementation requires accurate record keeping, information on every animal, and constant computation and publication of mean kinships (Lewis & Windig, 2017). Constant computation is no longer a problem nowadays, as is shown by the successful implementation by the Dutch Schapendoes club (Vereniging de Nederlandse Schapendoes, 2017). So if it was made available to them, breeders could use mean kinship to identify the genetically most important animals at any moment in time. Breeders base their breeding decisions not on any one factor, but on many factors, and the additional information on mean kinships could go a long way in reducing the effect of drift within a breed (Lewis et al., 2015).

While the strategy ‘Mean kinship’ may be the most effective at keeping the forces of genetic drift low, there is another genetic management strategy that showed similar results and came close to the same values found in ‘Mean kinship’. The strategy ‘Breeding cap’ is a fairly straightforward strategy that is easy to implement. All it requires is a limitation on the number of litters per sire. Some research suggests that restrictions on the number of litters per sire may be difficult to implement due to strong opposition from breeders, who might view such a measure as a potential loss of income (Windig & Oldenbroek, 2015). Even so, plenty of breed clubs already implement sire restrictions. Taking Dutch breeds as an example: only 1 out of 9 breeds does not have a restriction on the maximum number of litters per sire (this breed is the Schapendoes). The other eight breeds already have breeding protocols in place that enforce a maximum number of litters per sire of 3 (Saarloos wolfdog), 4 (Smoushond, Dutch shepherd), 5 (Markiesje), 6 (Wetterhoun), 9 (Drentsche Patrijshond), 10 (Stabyhoun), and 25 (Kooikerhondje) (Vereniging voor de Hollandse Herder, 2020; NVSWH, 2002; AVLS, 2021; DPHCN, 2014; Vereniging de Drentsche Patrijshond, 2021; NVSW, 2016; NVSW, 2019; Markiesjesvereniging, 2021; Markiesjes Welzijnsvereniging, 2021; De Schapendoes Club, 2021; Vereniging Het Nederlandse

Kooikerhondje, 2022; De Hollandse Smoushondenclub, 2012). This suggests that breed clubs in general may be open to this genetic management strategy. However, the sire restriction in this thesis was 2 litters per sire. None of the sire restrictions currently in place in the aforementioned examples has a restriction in place that is so severe. So while implementing a breeding cap on sires may appear to be the most practical strategy out of those simulated in this thesis, it may not be easily implemented.

Dog breeding is characterized by its complicated nature due to its uncoordinated breeding efforts and the many factors influencing breeding decisions (Wang et al, 2018). This makes it difficult to implement an efficient strategy that is both effective at ensuring a breed's health and is endorsed by breeders. While most research focuses on the effectiveness of breeding strategies, there is a notable lack of research exploring how breeders feel about different genetic management strategies and their willingness to accept them. Baublys & Tubelyte (2011) did one of the few studies that looks into where breeders get their information from to help guide their breeding decisions. It would be very helpful to shine a light on this subject, which would help in deciding which genetic management strategy is most efficient (both effective and supported by breeders). It would also help guide communication efforts towards breeders and breed clubs, by showing which strategies may count on breeder endorsement and which strategies may need extra measures taken for them to be accepted by the dog breeding community. And in the latter case: exactly which measures would result in sufficient breeder endorsement? How to convince breeders to accept and implement effective genetic management strategies? Research may find the best genetic management strategies, but they will not work if breeders do not embrace them.

5. CONCLUSION

Four sub-questions were formulated to answer the main research question. These sub-questions focused on the effect of the age of disease onset, the effect of the strength of expression of a disease allele, the effect of the effective population size, and whether disease alleles behave differently as a single allele versus in a group. The results in this thesis indicated that deleterious alleles with a higher disease strength and early age of onset were more likely to be eliminated from a population than those with a weak disease strength and late age of onset. Strong disease alleles and early onset disease alleles were affected more by selection, while weak disease alleles, so those with a small effect, and those that are late onset, were affected more by drift. The larger the effective population size, the more likely it was that disease alleles lingered at low frequencies in the population in a heterozygous state. Disease alleles were not likely to become fixed, but it took longer to eliminate them from the population. Small effective population sizes showed higher inbreeding levels and were affected more by drift, so it was more likely that disease alleles attained high frequencies and even became fixed. Disease alleles did not appear to behave differently when simulated as a single allele versus in a group.

The main research question was: *What kinds of genetic management most effectively reduce genetic defects in dog breeding?*

'Exclude carriers' almost always resulted in swift extinction of the population. This genetic management strategy used selection to reduce defects. The single surviving medium sized population showed very high inbreeding levels and a very high rate of inbreeding. When selection against disease alleles was very rigorous, the effective population grew smaller, inbreeding levels went up, the effect of drift grew larger, and disease alleles with a small effect were more likely to become fixed in the population. An accumulation of fixed deleterious alleles with a small effect may eventually lead to inbreeding depression and extinction.

The other three management strategies 'Least related', 'Mean kinship', and 'Breeding cap' focused on reducing drift by increasing the effective population size. They all successfully reduced inbreeding levels and rate of inbreeding, lowered the amount of fixed deleterious alleles, and helped prevent the fixation of disease alleles by decreasing the average disease allele frequencies. When the effective population size was increased, some disease alleles took longer to eliminate, but they lingered at low frequencies in a heterozygous state where they did not pose a problem.

6. REFERENCES

AVLS. (2020, 10 11). *Plan van Aanpak 2.0.* Retrieved from Algemene Vereniging voor Liefhebbers van Saarlooswolfhonden: https://avls.nl/wp-content/uploads/2022/01/Plan-van-Aanpak-2.0.pdf

AVLS. (2021, 05 24). *Fokreglement VFR*. Retrieved from Algemene Vereniging voor Liefhebbers van Saarlooswolfhonden: https://avls.nl/fokreglement/

Charlesworth, D., & Willis, J. H. (2009). The genetics of inbreeding depression. *Nature reviews genetics*, *10*(11), 783-796. https://doi-org.ezproxy.library.wur.nl/10.1038/nrg2664

De Hollandse Smoushonden Club. (2012, 01 21). *Verenigingsfokreglement*. Retrieved from Smoushond.nl: https://smoushond.nl/fokreglement/

De Schapendoes Club. (2021, 08 07). *Fokreglement*. Retrieved from De Schapendoes Club: https://www.schapendoesclub.nl/fokkerij/fokreglement.htm

Dreger, D. L., Rimbault, M., Davis, B. W., Bhatnagar, A., Parker, H. G., & Ostrander, E. A. (2016). Whole-genome sequence, SNP chips and pedigree structure: building demographic profiles in domestic dog breeds to optimize genetic-trait mapping. *Disease Models & Mechanisms*, *9*(12), 1445-1460. https://doi-org.ezproxy.library.wur.nl/10.1242/dmm.027037

DPHCN. (2014, 07 02). *Verenigings fokreglement (VFR) van de Drentsche Patrijshonden Club Nederland*. Retrieved from Drentsche Patrijshonden Club Nederland: https://www.dphcn.nl/fokkerij/verenigings-fokreglement/

Eggink, M., & Oldenbroek, J. K. (2014). Eerste pups outcross Saarlooswolfhond geboren. *Zeldzaam huisdier*, *39*(4), 18-20.

Fagan, W. F., & Holmes, E. (2006). Quantifying the extinction vortex. *Ecology letters*, *9*(1), 51-60. https://doi-org.ezproxy.library.wur.nl/10.1111/j.1461-0248.2005.00845.x

Flather, C. H., Hayward, G. D., Beissinger, S. R., & Stephens, P. A. (2011). Minimum viable populations: is there a 'magic number'for conservation practitioners?. *Trends in ecology & evolution*, *26*(6), 307-316. https://doi-org.ezproxy.library.wur.nl/10.1016/j.tree.2011.03.001

Frankham, R., Bradshaw, C. J., & Brook, B. W. (2014). Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation*, *170*, 56-63. https://doi-org.ezproxy.library.wur.nl/10.1016/j.biocon.2013.12.036

Frankham, R., Brook, B. W., Bradshaw, C. J., Traill, L. W., & Spielman, D. (2013). 50/500 rule and minimum viable populations: response to Jamieson and Allendorf. *Trends Ecol. Evol*, *28*, 187-188. https://doi-org.ezproxy.library.wur.nl/10.1016/j.tree.2013.01.002

Galeta, P., Lázničková-Galetová, M., Sablin, M., & Germonpré, M. (2021). Morphological evidence for early dog domestication in the European Pleistocene: New evidence from a randomization approach to group differences. *The Anatomical Record*, *304*(1), 42-62. https://doi-org.ezproxy.library.wur.nl/10.1002/ar.24500

Gilpin, M. E. (1986). Minimal viable populations: processes of species extinction. *Conservation biology: the science of scarcity and diversity*.

Hedrick, P. W., & Garcia-Dorado, A. (2016). Understanding inbreeding depression, purging, and genetic rescue. *Trends in ecology & evolution*, *31*(12), 940-952. https://doi-org.ezproxy.library.wur.nl/10.1016/j.tree.2016.09.005

Jamieson, I. G., & Allendorf, F. W. (2012). How does the 50/500 rule apply to MVPs?. *Trends in ecology & evolution*, *27*(10), 578-584. https://doi-org.ezproxy.library.wur.nl/10.1016/j.tree.2012.07.001

Jung, C., & Pörtl, D. (2019). How old are (pet) dog breeds?. *Pet Behaviour Science*, (7), 29-37. https://doi-org.ezproxy.library.wur.nl/10.21071/pbs.v0i7.11494

Keightley, P. D., Caballero, A., & García-Dorado, A. (1998). Population genetics: surviving under mutation pressure. *Current Biology*, *8*(7), R235-R237. https://doi-org.ezproxy.library.wur.nl/10.1016/s0960-9822(98)70148-4

Kettunen, A., Daverdin, M., Helfjord, T., & Berg, P. (2017). Cross-breeding is inevitable to conserve the highly inbred population of puffin hunter: The Norwegian Lundehund. *PLoS One*, *12*(1), e0170039. https://doi-org.ezproxy.library.wur.nl/10.1371/journal.pone.0170039

Leroy, G. (2011). Genetic diversity, inbreeding and breeding practices in dogs: results from pedigree analyses. *The Veterinary Journal*, *189*(2), 177-182. https://doi-org.ezproxy.library.wur.nl/10.1016/j.tvjl.2011.06.016

Leroy, G., & Rognon, X. (2012). Assessing the impact of breeding strategies on inherited disorders and genetic diversity in dogs. *The Veterinary Journal*, *194*(3), 343-348. https://doi-org.ezproxy.library.wur.nl/10.1016/j.tvjl.2012.06.025

Lewis, T. W., Abhayaratne, B. M., & Blott, S. C. (2015). Trends in genetic diversity for all Kennel Club registered pedigree dog breeds. *Canine genetics and epidemiology*, *2*(1), 1-10. https://doi-org.ezproxy.library.wur.nl/10.1186/s40575-015-0027-4

Lewis, T., & Windig, J. J. (2017). Dog breeds: towards genomic management of populations with a high incidence of genetic defects. *Genomic management of animal genetic diversity*, 179-205. https://doi-org.ezproxy.library.wur.nl/10.3920/978-90-8686-850-6_7

Lynch, M., Conery, J., & Burger, R. (1995). Mutation accumulation and the extinction of small populations. *The American Naturalist*, *146*(4), 489-518. https://doi-org.ezproxy.library.wur.nl/10.1086/285812

Markiesjes Welzijn Vereniging. (2021, 04 19). *Verenigingsfokreglement Markiesjes Welzijn Vereniging*. Retrieved from Markiesjes Welzijn Vereniging: https://markiesjeswelzijn.nl/fokken1/fokreglement

Markiesjesvereniging. (2021, 07 03). *Fokreglement*. Retrieved from Markiesjesvereniging: https://www.markiesjesvereniging.nl/rasbegeleiding/fokreglement.html

Masel, J. (2011). Genetic drift. *Current Biology*, *21*(20), R837-R838. https://doi-org.ezproxy.library.wur.nl/10.1016/j.cub.2011.08.007

Mathur, S., & DeWoody, J. A. (2021). Genetic load has potential in large populations but is realized in small inbred populations. *Evolutionary applications*, *14*(6), 1540-1557. https://doi-org.ezproxy.library.wur.nl/10.1111/eva.13216

Melis, C., Pertoldi, C., Ludington, W. B., Beuchat, C., Qvigstad, G., & Stronen, A. V. (2022). Genetic Rescue of the Highly Inbred Norwegian Lundehund. *Genes*, *13*(1), 163. https://doi-org.ezproxy.library.wur.nl/10.3390/genes13010163

Mooney, J. A., Yohannes, A., & Lohmueller, K. E. (2021). The impact of identity by descent on fitness and disease in dogs. *Proceedings of the National Academy of Sciences*, *118*(16), e2019116118. https://doi-org.ezproxy.library.wur.nl/10.1073/pnas.2019116118

NVSWH. (2002, 06 23). *Verenigingsfokreglement van de Nederlandse vereniging van Saarlooswolfhonden.* Retrieved from Saarlooswolfhonden.nl: https://saarlooswolfhonden.nl/wp-content/uploads/2021/12/VFRNVSWH20160611_20201201.pdf

NVSW. (2016, 06). *Verenigingsfokreglement.* Retrieved from NVSW.nl: https://www.nvsw.nl/uploads/VFR%20Wetterhoun%202022%20versie%20AV%202018.pdf

NVSW. (2019, 06). *Verenigingsfokreglement voor het ras: Stabijhoun.* Retrieved from NVSW.nl: https://www.nvsw.nl/uploads/VFR%20Stabijhoun%202022%20%20versie%20%20AV%202019.pdf

NVSW. (2022). *Outcross fokbeleid*. Retrieved from NVSW: https://www.nvsw.nl/wetterhoun/outcross-fokbeleid

Odom, A. (2022). Understanding Genetic Drift With Conceptual and Mathematical Models Breadcrumb. *The Science Teacher*, *89*(4).

Oldenbroek, K. (2016). *Het fokken in kleine populaties* [Powerpoint-slides]. CGN, Wageningen University and Research. Retrieved from https://edepot.wur.nl/409846

Oldenbroek, J. K., & Windig, J. J. (2012). *Het fokken van rashonden: omgaan met verwantschap en inteelt*. Raad van Beheer op Kynologische Gebied in Nederland.

Parker, H. G., Dreger, D. L., Rimbault, M., Davis, B. W., Mullen, A. B., Carpintero-Ramirez, G., & Ostrander, E. A. (2017). Genomic analyses reveal the influence of geographic origin, migration, and hybridization on modern dog breed development. *Cell reports*, *19*(4), 697-708. https://doi-org.ezproxy.library.wur.nl/10.1016/j.celrep.2017.03.079

Pavard, S., & Coste, C. F. (2021). Evolutionary demographic models reveal the strength of purifying selection on susceptibility alleles to late-onset diseases. *Nature Ecology & Evolution*, *5*(3), 392-400. https://doi-org.ezproxy.library.wur.nl/10.1038/s41559-020-01355-2

Rapaport, F., Boisson, B., Gregor, A., Béziat, V., Boisson-Dupuis, S., Bustamante, J., ... & Patin, E. (2021). Negative selection on human genes underlying inborn errors depends on disease outcome and both the mode and mechanism of inheritance. *Proceedings of the National Academy of Sciences*, *118*(3), e2001248118. https://doi-org.ezproxy.library.wur.nl/10.1073/pnas.2001248118

Rosenfeld, J. S. (2014). 50/500 or 100/1000? Reconciling short-and long-term recovery targets and MVPs. *Biological Conservation*, (176), 287-288. https://doi-org.ezproxy.library.wur.nl/10.1016/j.biocon.2014.05.005

Traill, L. W., Brook, B. W., Frankham, R. R., & Bradshaw, C. J. (2010). Pragmatic population viability targets in a rapidly changing world. *Biological Conservation*, *143*(1), 28-34. https://doi-org.ezproxy.library.wur.nl/10.1016/j.biocon.2009.09.001

Vereniging De Drentsche Patrijshond. (2021, 09 20). *Verenigingsfokreglement Drentsche Patrijshond.* Retrieved from Drentschepatrijshond.nl: https://drentschepatrijshond.com/fokkerij/vereningingsfokreglement/

Vereniging de Nederlandse Schapendoes. (2017). *Fokbeleid.* Retrieved from Schapendoes.nl: https://www.schapendoes.nl/language/nl/fokken/fokbeleid

Vereniging het Nederlandse Kooikerhondje. (2022, 07 01). *Verenigingsfokreglement (VFR).* Retrieved from Vereniging het Nederlandse Kooikerhondje: https://www.kooikerhondje.nl/download/verenigingsfokreglement/

Vereniging voor de Hollandse herder. (2020, 10 01). *Verenigingsfokreglement.* Retrieved from Vereniging voor de Hollandse herder: https://verenigingHollandseherder.nl/wp-content/uploads/2022/04/VFR-van-de-VVHH-def-versie-1-10-2020.pdf

Wang, S., Laloë, D., Missant, F. M., Malm, S., Lewis, T., Verrier, E., ... & Leroy, G. (2018). Breeding policies and management of pedigree dogs in 15 national kennel clubs. *The Veterinary Journal*, *234*,

130-135.
https://doi-org.ezproxy.library.wur.nl/10.1016/j.tvjl.2018.02.018

Wijnrocx, K., François, L., Stinckens, A., Janssens, S., & Buys, N. (2016). Half of 23 Belgian dog breeds has a compromised genetic diversity, as revealed by genealogical and molecular data analysis. *Journal of Animal Breeding and Genetics*, *133*(5), 375-383.
https://doi-org.ezproxy.library.wur.nl/10.1111/jbg.12203

Windig, J. J., & Doekes, H. P. (2018). Limits to genetic rescue by outcross in pedigree dogs. *Journal of Animal Breeding and Genetics*, *135*(3), 238-248.
https://doi-org.ezproxy.library.wur.nl/10.1111/jbg.12330

Windig, J. J., & Hulsegge, I. (2021). Retriever and pointer: Software to evaluate inbreeding and genetic management in captive populations. *Animals*, *11*(5), 1332.
https://doi-org.ezproxy.library.wur.nl/10.3390/ani11051332

Windig, J. J., & Oldenbroek, K. (2015). Genetic management of Dutch golden retriever dogs with a simulation tool. *Journal of Animal Breeding and Genetics*, *132*(6), 428-440.
https://doi-org.ezproxy.library.wur.nl/10.1111/jbg.12149

Worboys, M., Strange, J. M., & Pemberton, N. (2018). *The invention of the modern dog: breed and blood in Victorian Britain*. Johns Hopkins University Press.
https://doi-org.ezproxy.library.wur.nl/10.1007/s11016-020-00576-2

Wright, L. I., Tregenza, T., & Hosken, D. J. (2008). Inbreeding, inbreeding depression and extinction. *Conservation Genetics*, *9*(4), 833-843.
https://doi-org.ezproxy.library.wur.nl/10.1007/s10592-007-9405-0

Yordy, J., Kraus, C., Hayward, J. J., White, M. E., Shannon, L. M., Creevy, K. E., ... & Boyko, A. R. (2020). Body size, inbreeding, and lifespan in domestic dogs. *Conservation Genetics*, *21*(1), 137-148.
https://doi-org.ezproxy.library.wur.nl/10.1007/s10592-019-01240-x

APPENDIX

A. SMALLPOP REFERENCE POPULATION FOR MULTIPLE ALLELES

Alleles **(s=strength, y=age of onset)**	**Smallpop**			**Disease allele frequencies**		
	Fixed	*Segregating*	*Eliminated*	*Highest (segregating)*	*Lowest (segregating)*	*Average over all runs*
1-2 (s=0.01, y=0)	2, 3	32, 30	16, 17	0.992	0.008	0.338
3-6 (s=0.01, y=1)	5, 3, 5, 6	34, 37, 30, 35	11, 10, 15, 19	0.977	0.030	0.369
7-10 (s=0.01, y=4)	2, 6, 5, 6	34, 30, 31, 34	14, 14, 14, 10	0.992	0.008	0.382
11-14 (s=0.01, y=8)	5, 8, 5, 5	30, 29, 33, 38	15, 13, 12, 7	0.992	0.039	0.430
15-16 (s=0.05, y=0)	0, 0	30, 27	20, 23	0.952	0.008	0.189
17-20 (s=0.05, y=1)	1, 0, 0, 1	29, 29, 36, 38	20, 21, 14, 11	0.992	0.008	0.276
21-24 (s=0.05, y=4)	1, 2, 6, 5	29, 29, 37,	20, 19, 17, 14	0.992	0.008	0.303
25-28 (s=0.05, y=8)	4, 5, 4, 2	36, 30, 16, 32	10, 15, 10, 16	0.977	0.008	0.389
29-30 (s=0.1, y=0)	0, 0	24, 24	26, 26	0.656	0.008	0.124
31-34 (s=0.1, y=1)	0, 0, 0, 0	13, 30, 30, 17	27, 20, 20, 23	0.879	0.030	0.193
35-38 (s=0.1, y=4)	4, 0, 1, 1	34, 42, 34, 38	12, 8, 15, 11	0.970	0.030	0.363
39-42 (s=0.1, y=8)	5, 2, 6, 4	31, 36, 35, 34	14, 12, 9, 12	0.993	0.023	0.394
43 (s=0.25, y=0)	0	17	33	0.608	0.008	0.067
44-45 (s=0.25, y=1)	0, 0	20, 17	30, 33	0.708	0.015	0.09
46-47 (s=0.25, y=4)	0, 0	17, 24	33, 26	0.894	0.015	0.132
48-49 (s=0.25, y=8)	7, 4	32, 34	11, 12	0.970	0.053	0.433
50 (s=0.5, y=0)	0	2	48	0.115	0.083	0.004
51-52 (s=0.5, y=1)	0, 0	7, 8	43, 42	0.523	0.030	0.028
53-54 (s=0.5, y=4)	0, 0	21, 23	29, 27	0.773	0.008	0.143
55-56 (s=0.5, y=8)	5, 7	32, 32	13, 11	0.992	0.062	0.427
57 (s=1, y=0)	0	4	46	0.313	0.046	0.013
58 (s=1, y=1)	0	4	46	0.356	0.061	0.017
59 (s=1, y=4)	0	7	43	0.455	0.061	0.036
60 (s=1, y=8)	3	37	10	0.917	0.023	0.403

B. REFERENCE POPULATION FOR MULTIPLE ALLELES

Alleles **(s=strength, y=age of onset)**	**Midpop**			**Disease allele frequencies**		
	Fixed	*Segregating*	*Eliminated*	*Highest (segregating)*	*Lowest (segregating)*	*Average over all runs*
1-2 (s=0.01, y=0)	0, 0	50, 50	0, 0	0.788	0.061	0.322
3-6 (s=0.01, y=1)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.857	0.004	0.357
7-10 (s=0.01, y=4)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.880	0.084	0.403
11-14 (s=0.01, y=8)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.909	0.026	0.421
15-16 (s=0.05, y=0)	0, 0	50, 49	0, 1	0.296	0.002	0.172
17-20 (s=0.05, y=1)	0, 0, 0, 0	50, 50, 50, 49	0, 0, 0, 1	0.446	0.005	0.190
21-24 (s=0.05, y=4)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.766	0.022	0.361
25-28 (s=0.05, y=8)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.877	0.058	0.427
29-30 (s=0.1, y=0)	0, 0	49, 49	1, 1	0.246	0.001	0.109
31-34 (s=0.1, y=1)	0, 0, 0, 0	50, 50, 50, 48	0, 0, 0, 2	0.299	0.002	0.130
35-38 (s=0.1, y=4)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.171	0.048	0.285
39-42 (s=0.1, y=8)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.930	0.062	0.420
43 (s=0.25, y=0)	0	37	13	0.163	0.001	0.044
44-45 (s=0.25, y=1)	0, 0	40, 43	10, 7	0.190	0.001	0.055
46-47 (s=0.25, y=4)	0, 0	45, 44	5, 6	0.404	0.002	0.115
48-49 (s=0.25, y=8)	0, 0	50, 50	0, 0	0.767	0.040	0.402
50 (s=0.5, y=0)	0	25	25	0.139	0.007	0.028
51-52 (s=0.5, y=1)	0, 0	20, 25	30, 25	0.152	0.001	0.017
53-54 (s=0.5, y=4)	0, 0	48, 46	2, 4	0.217	0.002	0.085
55-56 (s=0.5, y=8)	0, 0	50, 50	0, 0	0.910	0.087	0.383
57 (s=1, y=0)	0	22	28	0.064	0.001	0.010
58 (s=1, y=1)	0	21	29	0.118	0.001	0.014
59 (s=1, y=4)	0	22	28	0.192	0.001	0.030
60 (s=1, y=8)	0	50	0	0.792	0.112	0.421

C. SMALLPOP BREEDING STRATEGY RESULTS FOR LEAST RELATED

Alleles (s=strength, y=age of onset)	**Smallpop**			**Disease allele frequencies**		
	Fixed	*Segregating*	*Eliminated*	*Highest (segregating)*	*Lowest (segregating)*	*Average over all runs*
1-2 (s=0.01, y=0)	4, 3	35, 38	11, 9	0.947	0.015	0.375
3-6 (s=0.01, y=1)	2, 3, 2, 4	38, 34, 38, 34	10, 13, 10, 12	0.997	0.023	0.370
7-10 (s=0.01, y=4)	3, 2, 4, 2	42, 36, 42, 37	5, 12, 4, 11	0.977	0.023	0.407
11-14 (s=0.01, y=8)	1, 2, 5, 5	40, 38, 35, 38	9, 10, 10, 7	0.977	0.023	0.414
15-16 (s=0.05, y=0)	0, 0	33, 35	17, 15	0.892	0.015	0.254
17-20 (s=0.05, y=1)	0, 0, 0, 1	34, 30, 40, 37	16, 20, 10, 12	0.954	0.015	0.274
21-24 (s=0.05, y=4)	4, 4, 3, 2	38, 35, 37, 42	8, 11, 10, 6	0.977	0.015	0.417
25-28 (s=0.05, y=8)	2, 4, 6, 0	38, 35, 38, 41	10, 11, 6, 9	0.977	0.023	0.413
29-30 (s=0.1, y=0)	0, 0	27, 32	23, 18	0.631	0.015	0.161
31-34 (s=0.1, y=1)	1, 0, 0, 0	32, 34, 34, 34	17, 16, 16, 16	0.894	0.023	0.210
35-38 (s=0.1, y=4)	2, 2, 2, 1	34, 37, 33, 41	14, 11, 15, 8	0.962	0.008	0.331
39-42 (s=0.1, y=8)	4, 4, 7, 6	39, 33, 34, 37	7, 13, 9, 7	0.922	0.008	0.425
43 (s=0.25, y=0)	0	15	35	0.462	0.115	0.063
44-45 (s=0.25, y=1)	0, 0	14, 12	36, 38	0.848	0.008	0.054
46-47 (s=0.25, y=4)	0, 0	23, 19	27, 31	0.800	0.008	0.113
48-49 (s=0.25, y=8)	6, 8	29, 34	15, 8	0.970	0.023	0.448
50 (s=0.5, y=0)	0	4	46	0.131	0.054	0.007
51-52 (s=0.5, y=1)	0, 0	6, 6	44, 44	0.333	0.008	0.019
53-54 (s=0.5, y=4)	0, 0	16, 15	34, 35	0.585	0.023	0.081
55-56 (s=0.5, y=8)	4, 6	39, 38	7, 6	0.992	0.008	0.470
57 (s=1, y=0)	0	2	48	0.212	0.115	0.007
58 (s=1, y=1)	0	2	48	0.205	0.144	0.007
59 (s=1, y=4)	0	7	43	0.386	0.061	0.031
60 (s=1, y=8)	6	36	8	0.992	0.008	0.481

D. MIDPOP BREEDING STRATEGY RESULTS FOR LEAST RELATED

Alleles (s=strength, y=age of onset)	**Midpop**			**Disease allele frequencies**		
	Fixed	*Segregating*	*Eliminated*	*Highest (segregating)*	*Lowest (segregating)*	*Average over all runs*
1-2 (s=0.01, y=0)	0, 0	50, 50	0, 0	0.753	0.019	0.330
3-6 (s=0.01, y=1)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.766	0.024	0.355
7-10 (s=0.01, y=4)	0, 0, 0, 0	49, 50, 50, 50	1, 0, 0, 0	0.924	0.010	0.410
11-14 (s=0.01, y=8)	0, 0, 0, 0	50, 50, 50, 49	0, 0, 0, 1	0.871	0.016	0.427
15-16 (s=0.05, y=0)	0, 0	50, 49	0, 1	0.429	0.022	0.164
17-20 (s=0.05, y=1)	0, 0, 0, 0	50, 49, 50, 50	0, 1, 0, 0	0.670	0.009	0.206
21-24 (s=0.05, y=4)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.884	0.003	0.311
25-28 (s=0.05, y=8)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.932	0.019	0.448
29-30 (s=0.1, y=0)	0, 0	49, 48	1, 2	0.312	0.001	0.101
31-34 (s=0.1, y=1)	0, 0, 0, 0	48, 50, 49, 49	2, 0, 1, 1	0.411	0.006	0.142
35-38 (s=0.1, y=4)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.770	0.010	0.282
39-42 (s=0.1, y=8)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.913	0.037	0.435
43 (s=0.25, y=0)	0	40	10	0.245	0.001	0.050
44-45 (s=0.25, y=1)	0, 0	41, 37	9, 13	0.184	0.002	0.057
46-47 (s=0.25, y=4)	0, 0	43, 46	7, 4	0.307	0.003	0.096
48-49 (s=0.25, y=8)	0, 0	50, 50	0, 0	0.798	0.036	0.422
50 (s=0.5, y=0)	0	21	29	0.176	0.001	0.023
51-52 (s=0.5, y=1)	0, 0	26, 24	24, 26	0.191	0.001	0.029
53-54 (s=0.5, y=4)	0, 0	41, 39	9, 11	0.339	0.002	0.072
55-56 (s=0.5, y=8)	0, 0	50, 50	0, 0	0.843	0.006	0.390
57 (s=1, y=0)	0	18	32	0.104	0.001	0.017
58 (s=1, y=1)	0	22	28	0.146	0.003	0.024
59 (s=1, y=4)	0	25	25	0.233	0.003	0.037
60 (s=1, y=8)	0	50	0	0.642	0.031	0.359

E. SMALLPOP BREEDING STRATEGY RESULTS FOR MEAN KINSHIP

Alleles (s=strength, y=age of onset)	**Smallpop**			**Disease allele frequencies**		
	Fixed	*Segregating*	*Eliminated*	*Highest (segregating)*	*Lowest (segregating)*	*Average over all runs*
1-2 (s=0.01, y=0)	1, 2	42, 43	7, 5	0.932	0.008	0.385
3-6 (s=0.01, y=1)	1, 1, 1, 3	37, 40, 41, 42	12, 9, 8, 5	0.977	0.008	0.367
7-10 (s=0.01, y=4)	4, 2, 0, 2	40, 44, 41, 45	6, 4, 9, 3	0.955	0.023	0.431
11-14 (s=0.01, y=8)	4, 5, 3, 4	40, 42, 41, 42	6, 3, 6, 4	0.992	0.015	0.477
15-16 (s=0.05, y=0)	0, 1	38, 35	12, 15	0.750	0.015	0.207
17-20 (s=0.05, y=1)	0, 1, 0, 0	37, 37, 37, 41	13, 13, 13, 9	0.894	0.015	0.267
21-24 (s=0.05, y=4)	1, 0, 0, 2	43, 45, 41, 42	6, 5, 9, 6	0.977	0.015	0.359
25-28 (s=0.05, y=8)	4, 0, 3, 2	39, 42, 40, 44	7, 8, 7, 4	0.985	0.023	0.432
29-30 (s=0.1, y=0)	0, 0	35, 32	15, 18	0.606	0.008	0.136
31-34 (s=0.1, y=1)	0, 0, 0, 0	38, 37, 36, 36	12, 13, 14, 14	0.685	0.008	0.185
35-38 (s=0.1, y=4)	1, 1, 2, 1	43, 43, 38, 40	6, 6, 10, 9	0.955	0.008	0.310
39-42 (s=0.1, y=8)	2, 2, 2, 2	45, 39, 42, 43	3, 9, 6, 5	0.985	0.008	0.426
43 (s=0.25, y=0)	0	19	31	0.288	0.008	0.047
44-45 (s=0.25, y=1)	0, 0	22, 16	28, 34	0.400	0.008	0.054
46-47 (s=0.25, y=4)	0, 0	30, 22	20, 28	0.682	0.008	0.131
48-49 (s=0.25, y=8)	3, 2	40, 44	7, 4	0.985	0.077	0.436
50 (s=0.5, y=0)	0	7	43	0.318	0.015	0.017
51-52 (s=0.5, y=1)	0, 0	6, 4	44, 46	0.242	0.015	0.011
53-54 (s=0.5, y=4)	0, 0	24, 32	26, 18	0.685	0.008	0.135
55-56 (s=0.5, y=8)	5, 4	41, 42	4, 4	0.985	0.008	0.459
57 (s=1, y=0)	0	7	43	0.315	0.046	0.018
58 (s=1, y=1)	0	7	43	0.371	0.015	0.024
59 (s=1, y=4)	0	8	42	0.318	0.008	0.020
60 (s=1, y=8)	2	42	6	0.985	0.061	0.432

F. MIDPOP BREEDING STRATEGY RESULTS FOR MEAN KINSHIP

Alleles (s=strength, y=age of onset)	**Midpop**			**Disease allele frequencies**		
	Fixed	*Segregating*	*Eliminated*	*Highest (segregating)*	*Lowest (segregating)*	*Average over all runs*
1-2 (s=0.01, y=0)	0, 0	50, 50	0, 0	0.660	0.020	0.332
3-6 (s=0.01, y=1)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.795	0.030	0.336
7-10 (s=0.01, y=4)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.814	0.071	0.401
11-14 (s=0.01, y=8)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.805	0.071	0.409
15-16 (s=0.05, y=0)	0, 0	50, 50	0, 0	0.457	0.046	0.180
17-20 (s=0.05, y=1)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.512	0.008	0.200
21-24 (s=0.05, y=4)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.761	0.039	0.360
25-28 (s=0.05, y=8)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.795	0.098	0.428
29-30 (s=0.1, y=0)	0, 0	50, 50	0, 0	0.441	0.004	0.111
31-34 (s=0.1, y=1)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.318	0.001	0.122
35-38 (s=0.1, y=4)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.734	0.058	0.316
39-42 (s=0.1, y=8)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.853	0.046	0.404
43 (s=0.25, y=0)	0	43	7	0.171	0.003	0.043
44-45 (s=0.25, y=1)	0, 0	46, 47	4, 3	0.196	0.001	0.053
46-47 (s=0.25, y=4)	0, 0	50, 46	0, 4	0.301	0.002	0.096
48-49 (s=0.25, y=8)	0, 0	50, 50	0, 0	0.851	0.119	0.399
50 (s=0.5, y=0)	0	27	23	0.109	0.002	0.026
51-52 (s=0.5, y=1)	0, 0	32, 32	18, 18	0.175	0.001	0.021
53-54 (s=0.5, y=4)	0, 0	48, 47	2, 3	0.266	0.005	0.084
55-56 (s=0.5, y=8)	0, 0	50, 50	0, 0	0.824	0.070	0.417
57 (s=1, y=0)	0	33	17	0.098	0.002	0.021
58 (s=1, y=1)	0	28	22	0.166	0.002	0.022
59 (s=1, y=4)	0	38	12	0.144	0.001	0.031
60 (s=1, y=8)	0	50	0	0.770	0.056	0.381

G. SMALLPOP BREEDING STRATEGY RESULTS FOR BREEDING CAP

Alleles (s=strength, y=age of onset)	**Smallpop**			**Disease allele frequencies**		
	Fixed	*Segregating*	*Eliminated*	*Highest (segregating)*	*Lowest (segregating)*	*Average over all runs*
1-2 (s=0.01, y=0)	0, 2	40, 41	10, 7	0.015	0.972	0.389
3-6 (s=0.01, y=1)	1, 2, 2, 0	39, 38, 39, 38	10, 10, 9, 12	0.008	0.974	0.345
7-10 (s=0.01, y=4)	2, 4, 2, 2	41, 42, 43, 41	7, 4, 5, 7	0.008	0.910	0.397
11-14 (s=0.01, y=8)	3, 3, 5, 3	38, 39, 38, 42	9, 8, 7, 5	0.030	0.974	0.433
15-16 (s=0.05, y=0)	0, 0	39, 38	11, 12	0.015	0.700	0.207
17-20 (s=0.05, y=1)	0, 0, 0, 0	40, 35, 33, 40	10, 15, 17, 10	0.023	0.970	0.218
21-24 (s=0.05, y=4)	2, 0, 2, 2	42, 41, 41, 40	6, 9, 7, 8	0.023	0.972	0.393
25-28 (s=0.05, y=8)	3, 2, 1, 1	38, 41, 40, 45	9, 7, 9, 4	0.023	0.972	0.381
29-30 (s=0.1, y=0)	0, 0	31, 34	19, 16	0.008	0.510	0.115
31-34 (s=0.1, y=1)	0, 0, 0, 0	31, 35, 35, 33	19, 15, 15, 17	0.016	0.839	0.161
35-38 (s=0.1, y=4)	1, 1, 1, 4	40, 44, 41, 40	9, 5, 8, 6	0.016	0.985	0.371
39-42 (s=0.1, y=8)	2, 3, 3, 0	37, 41, 39, 40	11, 6, 8, 10	0.038	0.985	0.355
43 (s=0.25, y=0)	0	17	33	0.032	0.451	0.052
44-45 (s=0.25, y=1)	0, 0	16, 13	34, 37	0.016	0.455	0.049
46-47 (s=0.25, y=4)	0, 0	25, 28	25, 22	0.015	0.774	0.158
48-49 (s=0.25, y=8)	3, 3	42, 42	5, 5	0.016	0.985	0.411
50 (s=0.5, y=0)	0	9	41	0.047	0.177	0.019
51-52 (s=0.5, y=1)	0, 0	1, 7	49, 43	0.048	0.455	0.013
53-54 (s=0.5, y=4)	0, 0	24, 21	26, 29	0.008	0.515	0.087
55-56 (s=0.5, y=8)	4, 1	39, 43	7, 6	0.065	0.984	0.436
57 (s=1, y=0)	0	5	45	0.030	0.127	0.008
58 (s=1, y=1)	0	7	43	0.016	0.177	0.009
59 (s=1, y=4)	0	9	41	0.069	0.455	0.031
60 (s=1, y=8)	4	43	3	0.045	0.973	0.472

H. MIDPOP BREEDING STRATEGY RESULTS FOR BREEDING CAP

Alleles (s=strength, y=age of onset)	**Midpop**			**Disease allele frequencies**		
	Fixed	*Segregating*	*Eliminated*	*Highest (segregating)*	*Lowest (segregating)*	*Average over all runs*
1-2 (s=0.01, y=0)	0, 0	50, 50	0, 0	0.701	0.0735	0.310
3-6 (s=0.01, y=1)	0, 0, 0, 0	50, 50, 50, 50	0, 0	0.712	0.040	0.343
7-10 (s=0.01, y=4)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.828	0.048	0.415
11-14 (s=0.01, y=8)	0, 0, 0, 0	49, 50, 50, 50	1, 0, 0, 0	0.879	0.044	0.423
15-16 (s=0.05, y=0)	0, 0	50, 50	0, 0	0.356	0.020	0.146
17-20 (s=0.05, y=1)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.441	0.005	0.186
21-24 (s=0.05, y=4)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.803	0.012	0.357
25-28 (s=0.05, y=8)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.921	0.047	0.421
29-30 (s=0.1, y=0)	0, 0	50, 50	0, 0	0.278	0.008	0.092
31-34 (s=0.1, y=1)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.328	0.008	0.113
35-38 (s=0.1, y=4)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.801	0.049	0.318
39-42 (s=0.1, y=8)	0, 0, 0, 0	50, 50, 50, 50	0, 0, 0, 0	0.875	0.030	0.442
43 (s=0.25, y=0)	0	46	4	0.108	0.003	0.036
44-45 (s=0.25, y=1)	0, 0	44, 38	6, 12	0.137	0.001	0.037
46-47 (s=0.25, y=4)	0, 0	50, 47	0, 3	0.390	0.006	0.126
48-49 (s=0.25, y=8)	0, 0	50, 50	0, 0	0.906	0.078	0.429
50 (s=0.5, y=0)	0	28	22	0.078	0.001	0.013
51-52 (s=0.5, y=1)	0, 0	28, 28	22, 22	0.098	0.001	0.019
53-54 (s=0.5, y=4)	0, 0	48, 48	2, 2	0.330	0.002	0.089
55-56 (s=0.5, y=8)	0, 0	50, 50	0, 0	0.815	0.059	0.433
57 (s=1, y=0)	0	24	26	0.100	0.003	0.014
58 (s=1, y=1)	0	22	28	0.084	0.004	0.012
59 (s=1, y=4)	0	38	12	0.224	0.001	0.047
60 (s=1, y=8)	0	50	0	0.794	0.148	0.427

I. MIDPOP BREEDING STRATEGY RESULTS FOR EXCLUDE CARRIERS

Alleles **(s=strength, y=age of onset)**	**Midpop**			**Disease allele frequencies**		
	Fixed	*Segregating*	*Eliminated*	*Highest (segregating)*	*Lowest (segregating)*	~~*Average over all runs*~~ *
1-2 (s=0.01, y=0)		all **		0.959	0.442	
3-6 (s=0.01, y=1)		all		0.598	0.041	
7-10 (s=0.01, y=4)		all		0.959	0.041	
11-14 (s=0.01, y=8)		3	1	0.959	0.402	
15-16 (s=0.05, y=0)		all		0.402		
17-20 (s=0.05, y=1)		2	2	0.959	0.402	
21-24 (s=0.05, y=4)		3	1	0.959	0.402	
25-28 (s=0.05, y=8)	1	3		0.598	0.402	
29-30 (s=0.1, y=0)		all		0.598	0.442	
31-34 (s=0.1, y=1)		all		0.217	0.027	
35-38 (s=0.1, y=4)		all		0.783	0.217	
39-42 (s=0.1, y=8)	1	3		0.783	0.190	
43 (s=0.25, y=0)			all			
44-45 (s=0.25, y=1)			all			
46-47 (s=0.25, y=4)			all			
48-49 (s=0.25, y=8)			all			
50 (s=0.5, y=0)			all			
51-52 (s=0.5, y=1)			all			
53-54 (s=0.5, y=4)			all			
55-56 (s=0.5, y=8)			all			
57 (s=1, y=0)			all			
58 (s=1, y=1)			all			
59 (s=1, y=4)			all			
60 (s=1, y=8)			all			

* Only one single run survived this simulation

** 'all' referring to all alleles with a particular strength and age of onset. When a number is given (1, 2, 3) it refers to 1, 2, or 3 alleles with a particular strength and age of onset.

J. TABLES WITH TOTAL NUMBER OF FIXED AND ELIMINATED ALLELES (SMALLPOP)

Table A: Alleles with age of onset of (y=0) years old, showing the number of times an allele was fixed or eliminated in a population (out of 50 times). For example, allele 1-0010 was fixed 2 times in 50 runs in the reference population.

	Reference		Least related		Mean kinship		Limited sires	
	Fixed	*Eliminated*	*Fixed*	*Eliminated*	*Fixed*	*Eliminated*	*Fixed*	*Eliminated*
1-0010	2	16	4	11	1	7	0	10
15-0050	0	20	0	17	0	12	0	11
29-010	0	26	0	23	0	15	0	19
43-0250	0	33	0	35	0	31	0	33
50-050	0	48	0	46	0	43	0	41
57-10	0	46	0	48	0	43	0	45
TOTAL	2	189	4	180	1	151	0	159

Table B: Alleles with age of onset of (y=1) years old, showing the number of times an allele was fixed or eliminated in a population (out of 50 times). For example, allele 3-0011 was fixed 5 times in 50 runs in the reference population.

	Reference		Least related		Mean kinship		Limited sires	
	Fixed	*Eliminated*	*Fixed*	*Eliminated*	*Fixed*	*Eliminated*	*Fixed*	*Eliminated*
3-0011	5	11	2	10	1	12	1	10
17-0051	1	20	0	16	0	13	0	10
31-011	0	27	1	17	0	12	0	19
44-0251	0	30	0	36	0	28	0	34
51-051	0	43	0	44	0	44	0	49
58-11	0	46	0	48	0	43	0	43
TOTAL	6	177	3	171	1	152	1	165

Table C: Alleles with age of onset of (y=4) years old, showing the number of times an allele was fixed or eliminated in a population (out of 50 times). For example, allele 7-0014 was fixed 2 times in 50 runs in the reference population.

	Reference		Least related		Mean kinship		Limited sires	
	Fixed	*Eliminated*	*Fixed*	*Eliminated*	*Fixed*	*Eliminated*	*Fixed*	*Eliminated*
7-0014	2	14	3	5	4	6	2	7
21-0054	1	20	4	8	1	6	2	6
35-014	4	12	2	14	1	6	1	9
46-0254	0	33	0	27	0	20	0	25
53-054	0	29	0	34	0	26	0	26
59-14	0	43	0	43	0	42	0	41
TOTAL	7	151	9	131	6	106	5	114

Table D: Alleles with age of onset of (y=8) years old, showing the number of times an allele was fixed or eliminated in a population (out of 50 times). For example, allele 11-0018 was fixed 5 times in 50 runs in the reference population.

	Reference		**Least related**		**Mean kinship**		**Limited sires**	
	Fixed	*Eliminated*	*Fixed*	*Eliminated*	*Fixed*	*Eliminated*	*Fixed*	*Eliminated*
11-0018	5	15	1	9	4	6	3	9
25-0058	4	10	2	10	4	7	3	9
39-018	5	14	4	7	2	3	2	11
48-0258	7	11	6	15	3	7	3	5
55-058	5	12	4	7	5	4	4	7
60-18	3	10	6	8	2	6	4	3
TOTAL	29	72	23	56	20	33	19	44

K. TABLES WITH TOTAL NUMBER OF FIXED AND ELIMINATED ALLELES (MIDPOP)

Table E: Alleles with age of onset of (y=0) years old, showing the number of times an allele was fixed or eliminated in a population (out of 50 times). For example, allele 1-0010 was fixed 0 times in 50 runs in the reference population.

	Reference		**Least related**		**Mean kinship**		**Limited sires**	
	Fixed	*Eliminated*	*Fixed*	*Eliminated*	*Fixed*	*Eliminated*	*Fixed*	*Eliminated*
1-0010	0	0	0	0	0	0	0	0
15-0050	0	0	0	0	0	0	0	0
29-010	0	1	0	1	0	0	0	0
43-0250	0	13	0	10	0	7	0	4
50-050	0	25	0	29	0	23	0	22
57-10	0	28	0	32	0	17	0	27
TOTAL	0	67	0	72	0	47	0	53

Table F: Alleles with age of onset of (y=1) years old, showing the number of times an allele was fixed or eliminated in a population (out of 50 times). For example, allele 3-0011 was fixed 0 times in 50 runs in the reference population.

	Reference		**Least related**		**Mean kinship**		**Limited sires**	
	Fixed	*Eliminated*	*Fixed*	*Eliminated*	*Fixed*	*Eliminated*	*Fixed*	*Eliminated*
3-0011	0	0	0	0	0	0	0	0
17-0051	0	0	0	0	0	0	0	0
31-011	0	0	0	2	0	0	0	0
44-0251	0	10	0	9	0	4	0	6
51-051	0	30	0	24	0	18	0	22
58-11	0	29	0	28	0	22	0	28
TOTAL	0	69	0	63	0	44	0	56

Table G: Alleles with age of onset of (y=4) years old, showing the number of times an allele was fixed or eliminated in a population (out of 50 times). For example, allele 7-0014 was fixed 0 times in 50 runs in the reference population.

	Reference		**Least related**		**Mean kinship**		**Limited sires**	
	Fixed	*Eliminated*	*Fixed*	*Eliminated*	*Fixed*	*Eliminated*	*Fixed*	*Eliminated*
7-0014	0	0	0	1	0	0	0	0
21-0054	0	0	0	0	0	0	0	0
35-014	0	0	0	0	0	0	0	0
46-0254	0	5	0	7	0	0	0	0
53-054	0	2	0	9	0	2	0	2
59-14	0	28	0	25	0	12	0	12
TOTAL	0	35	0	42	0	14	0	14

Table H: Alleles with age of onset of (y=8) years old, showing the number of times an allele was fixed or eliminated in a population (out of 50 times). For example, allele 11-0018 was fixed 0 times in 50 runs in the reference population.

	Reference		**Least related**		**Mean kinship**		**Limited sires**	
	Fixed	*Eliminated*	*Fixed*	*Eliminated*	*Fixed*	*Eliminated*	*Fixed*	*Eliminated*
11-0018	0	0	0	0	0	0	0	0
25-0058	0	0	0	0	0	0	0	0
39-018	0	0	0	0	0	0	0	0
48-0258	0	0	0	0	0	0	0	0
55-058	0	0	0	0	0	0	0	0
60-18	0	0	0	0	0	0	0	0
TOTAL	0	0	0	0	0	0	0	0

L. DATA MANAGEMENT PLAN

Data management plan belonging to the MSc thesis performed at the Animal Breeding and Genomics group by Roxy Bergsma, completed in August 2022.

The data and scripts can be found through Jack Windig (jack.windig@wur.nl).

AGE OF ONSET OF DISEASE SIMULATIONS

These files can be found in the folder: **Single alleles Age of onset simulations**

File names	Created in (month, year)	Remarks
AOO Midpop 1.05.1 .xlsx	13-2-22	Excel sheet
AOO Midpop 1.05.1 .txt	8-2-22	Pointer output file
AOO Midpop 1.05.2 .txt	8-2-22	Pointer output file
AOO Midpop 1.05.2 .xlsx	13-2-22	Excel sheet
AOO Midpop 1.05.3 .txt	8-2-22	Pointer output file
AOO Midpop 1.05.3 .xlsx	13-2-22	Excel sheet
AOO Midpop 1.05.4 .txt	8-2-22	Pointer output file
AOO Midpop 1.05.4 .xlsx	13-2-22	Excel sheet
AOO Midpop 1.05.5 .txt	8-2-22	Pointer output file
AOO Midpop 1.05.5 .xlsx	13-2-22	Excel sheet
AOO Midpop 1.05.6 .txt	8-2-22	Pointer output file
AOO Midpop 1.05.6 .xlsx	13-2-22	Excel sheet
AOO Midpop 1.05.7 .txt	8-2-22	Pointer output file
AOO Midpop 1.05.7 .xlsx	13-2-22	Excel sheet
AOO Midpop 1.05.8 .txt	8-2-22	Pointer output file
AOO Midpop 1.05.8 .xlsx	13-2-22	Excel sheet
AOO Smalpop 1.05.1 .xlsx	13-2-22	Excel sheet
AOO Smalpop 1.05.1 .txt	25-1-22	Pointer output file
AOO Smalpop 1.05.2 .txt	7-2-22	Pointer output file
AOO Smalpop 1.05.2 .xlsx	13-2-22	Excel sheet
AOO Smalpop 1.05.3 .txt	25-1-22	Pointer output file
AOO Smalpop 1.05.3 .xlsx	13-2-22	Excel sheet
AOO Smalpop 1.05.4 .txt	7-2-22	Pointer output file
AOO Smalpop 1.05.4 .xlsx	13-2-22	Excel sheet
AOO Smalpop 1.05.5 .txt	7-2-22	Pointer output file
AOO Smalpop 1.05.5 .xlsx	13-2-22	Excel sheet
AOO Smalpop 1.05.6 .txt	25-1-22	Pointer output file
AOO Smalpop 1.05.6 .xlsx	13-2-22	Excel sheet
AOO Smalpop 1.05.7 .txt	7-2-22	Pointer output file
AOO Smalpop 1.05.7 .xlsx	13-2-22	Excel sheet

AOO Smalpop 1.05.8 .txt	7-2-22	Pointer output file
AOO Smalpop 1.05.8 .xlsx	13-2-22	Excel sheet

STRENGTH OF DISEASE SIMULATIONS

These files can be found in: **Single alleles Strength of disease simulations**

File names	**Created in (month, year)**	**Remarks**
Midpop str 001.50.4 .xlsx	20-2-22	Excel sheet
Midpop str 001.05.4 .txt	20-2-22	Pointer output file
Midpop str 005.05.4 .txt	20-2-22	Pointer output file
Midpop str 005.50.4 .xlsx	20-2-22	Excel sheet
Midpop str 01.05.4 .txt	20-2-22	Pointer output file
Midpop str 01.50.4 .xlsx	20-2-22	Excel sheet
Midpop str 025.05.4 .txt	20-2-22	Pointer output file
Midpop str 025.50.4 .xlsx	20-2-22	Excel sheet
Midpop str 05.05.4 .txt	20-2-22	Pointer output file
Midpop str 05.50.4 .xlsx	20-2-22	Excel sheet
Midpop str 1.05.4 .txt	20-2-22	Pointer output file
Midpop str 1.50.4 .xlsx	20-2-22	Excel sheet
Smalpop str 001.50.4 .xlsx	20-2-22	Excel sheet
Smalpop str 001.05.4 .txt	20-2-22	Pointer output file
Smalpop str 005.05.4 .txt	20-2-22	Pointer output file
Smalpop str 005.50.4 .xlsx	20-2-22	Excel sheet
Smalpop str 01.05.4 .txt	20-2-22	Pointer output file
Smalpop str 01.50.4 .xlsx	20-2-22	Excel sheet
Smalpop str 025.05.4 .txt	20-2-22	Pointer output file
Smalpop str 025.50.4 .xlsx	20-2-22	Excel sheet
Smalpop str 05.05.4 .txt	20-2-22	Pointer output file
Smalpop str 05.50.4 .xlsx	20-2-22	Excel sheet
Smalpop str 1.05.4 .txt	20-2-22	Pointer output file
Smalpop str 1.50.4 .xlsx	20-2-22	Excel sheet

REFERENCE SIMULATIONS

These files can be found in the folder: **Multiple alleles Reference simulations**

File names	**Created in (month, year)**	**Remarks**
Midpop multi allele mix .xlsx	28-2-22	Excel sheet
Genmansim input midpop allele mix .txt	28-2-22	Pointer input file

Midpop multi allele mix .txt	28-2-22	Pointer output file
Smallpop multi allele mix .xlsx	28-2-22	Excel sheet
Genmansim input smallpop allele mix .txt	28-2-22	Pointer input file
Smallpop multi allele mix .txt	28-2-22	Pointer output file

BREEDING STRATEGY SIMULATIONS

These files can be found in the folder: **Multiple alleles Breeding strategy simulations**

File names	**Created in (month, year)**	**Remarks**
Midpopcarriersgone ALL (exclude carriers) .xlsx	31-3-22	Excel file
Midpopcarriersgone ALL (exclude carriers) .txt	21-3-22	Pointer output file
Genmansim input Midpop carriers gone ALL .txt	21-3-22	Pointer input file
Midpopcarriersgone SEVERE (exclude carriers) .xlsx	31-3-22	Excel file
Midpopcarriersgone SEVERE (exclude carriers) .txt	21-3-22	Pointer output file
Genmansim input Midpop carriers gone ONLY SEVERE .txt	21-3-22	Pointer input file
Midpoplimitedsires (breeding cap) .xlsx	31-3-22	Excel file
Midpoplimitedsires (breeding cap) .txt	31-3-22	Pointer output file
Genmansim input Midpop limited sires .txt	31-3-22	Pointer input file
MidpopMKpopulationminus3gen (mean kinship) .xlsx	31-3-22	Excel file
MidpopMKpopulationminus3gen (mean kinship) .txt	31-3-22	Pointer output file
Genmansim input Midpop MK within population minus 3gen .txt	31-3-22	Pointer input file
MidpopMK3gen (least related) .xlsx	31-3-22	Excel file
MidpopMK3gen (least related) .txt	21-3-22	Pointer output file
Genmansim input Midpop MK 3gen .txt	21-3-22	Pointer input file
Smallpopcarriersgone ALL (exclude carriers) .xlsx	31-3-22	Excel file
Smallpopcarriersgone ALL (exclude carriers) .txt	21-3-22	Pointer output file
Genmansim input Smallpopcarriers	21-3-22	Pointer input file

gone ALL .txt		
Smallpopcarriersgone SEVERE (exclude carriers) .xlsx	31-3-22	Excel file
Smallpopcarriersgone SEVERE (exclude carriers) .txt	21-3-22	Pointer output file
Genmansim input Smallpopcarriers gone ONLY SEVERE .txt	21-3-22	Pointer input file
Smallpoplimitedsires (breeding cap) .xlsx	31-3-22	Excel file
Smallpoplimitedsires (breeding cap) .txt	31-3-22	Pointer output file
Genmansim input Smallpoplimited sires .txt	31-3-22	Pointer input file
SmallpopMKpopulationminus3gen (mean kinship) .xlsx	31-3-22	Excel file
SmallpopMKpopulationminus3gen (mean kinship) .txt	31-3-22	Pointer output file
Genmansim input SmallpopMK within population minus 3gen .txt	31-3-22	Pointer input file
SmallpopMK3gen (least related) .xlsx	31-3-22	Excel file
SmallpopMK3gen (least related) .txt	31-3-22	Pointer output file
Genmansim input SmallpopMK 3gen .txt	31-3-22	Pointer input file