Powering the breeding of gene-modified animals

Beating Murphy s law: Powering the breedings of gene-modified animals

Better planning of laboratory animal crosses: Beating Murphy

Mendel´s laws and prospective “powering”

Group size planning of breedings

Philippe Bugnon, Achim Tresch, Frank Brand, Thorsten Buch

MGI

# Introduction

The analysis of gene-modified rodents constitutes one pillar of modern biomedical research. According to there exist around XX different gene-modified strains, excluding so far unpublished strains and those established in industry. The total number of rodents bred in Switzerland per year is XXXX, in the EU it is YYY. Today not only single mutants are bred and analyzed: rather combinations of multiple alleles of different genes have become mainstay in research. Obtaining these complex genotypes and breeding of these animals for obtaining the required number of animals for a designed experiment is based on Mendel´s laws of inheritance, thus segregation and independent assortment (assuming the genetic loci are found on separate chromosomes). In daily laboratory life a new breeding is planned with help of the Punnet square. The possible allele combinations possibly found in a haploid germ cell, be it sperm or oocyte, are written at two orthogonal sides of the Punnet square, with the sperms at the top and the oocytes at the side. The different sperm cells and oocytes defined by their alleles are combined in the center of the square to form the possible allele combinations of the next generations. Further, the Punnet square allows the determination of the frequency of appearance for each new diploid genotype of the next generation. But unfortunately, as every biomedical researcher using gene-modified rodents can attest to, knowing the frequency of the expected genotypes does not correlate directly to the actual breeding outcomes. With our usually rather small number of parallel breedings and thus offspring, we operate in the realm of stochastic events. Yet, this fact is usually not taken into account when breeding of gene-modified rodents is planned – leading to unnecessary delays in breeding success. In practice bad planning of breedings often leads to researchers collecting animals for an experiment over time – leading to a cohort of animals with grossly varying ages, a possible confounder in many experiments. We show here, how inclusion of the probability of occurrence of Mendelian genotypes in small breeding cohorts can be incorporated into everyday breeding planning.

# Results

## The probability of obtaining a genotype in a certain litter.

While Mendel´s laws of segregation and independent assortment allow us to calculate the theoretical frequencies of genotypes in the next generation of a given breeding, this information alone is insufficient to predict the likelihood of appearance of these genotypes, especially in situations of small group sizes, hence where lowest number of surplus animals are supposed to be generated. As shown in Fig. 1A for a simple cross of two animals heterozygous for a gene deficiency (knockout, KO), the Mendelian frequencies of genotype appearance in the next generation are 0.25 (WT/WT) : 0.5 (WT/KO) : 0.25 (KO/KO). When we assume that a litter of a typical breeding in C57BL/6 background yields 6 pups, we can model such a breeding step by randomly picking 6 individuals from the (infinite) pool of individuals endowed with the three genotypes in the frequencies 1:2:1 (Fig. 1B, depicted in the colors blue, green, and pink, respectively). Now it becomes clear that the appearances of 1, 2, 3 or more individuals of the KO/KO genotype is governed by stochastic rules. The occurrence of at least 1(2, 3 etc.) animals of a respective genotype is described by a cumulative binomial function (Fig. 1C and D). The reason for a cumulative binomial distribution becomes clear when we need for our experiment only 1 KO/KO individual. In this situation we do not require a litter to have exactly one KO/KO individual; we can continue with our experimental plan also with litters containing 2, 3, 4 ,5 or even 6 KO/KO individuals. Thus, we can add up the probabilities for obtaining 1, 2, 3, 4, 5 and 6 KO/KO individuals. Obviously the probability of obtaining a single KO/KO individual is quite high, when we can take it from any litter containing at least one KO/KO individual. The outcome of this binomial calculation yields the probability of the litter containing a certain minimal number of KO/KO (or any other) genotype from a given breeding (Fig. 1C and D). While this is useful information it does not help us in planning for obtaining a certain number of KO/KO individuals. The probabilities for obtaining a given number of KO/KO individuals increase when we produce more pups by setting up more breedings (Fig. 1E and F). The red lines indicate the required animals to be born to obtain 3 pups of the respective genotype frequency with a likelihood of 90%. Taken together, we can calculate the probability of a genotype appearing in a single or several litter(s) by combining Mendel´s rules and the cumulative binomial function.

## Predicting Breeding Outcome for One Group Cases

While the knowledge of probabilities regarding a breeding outcome may help us in better understanding why many of our breedings fail, thus, why we often do not obtain sufficient animals for the next breeding step or our experiments, it would be even more useful to be able to do make a group size prediction for appropriate setup of breedings. Hence, if you require for your experiment 5 KO/KO individuals, would it not be sensible if you would plan your breedings in such a way, that you obtain these 5 animals with a probability of e.g. 90%? Using the above-described binomial function we can make exactly these predictions. They depend on the probability of a genotype appearing in the next generation according to Mendel and the required number of animals of a particular genotype. Fig. 2 shows the comparison of the required number of animals as calculated with the Mendel rules alone and including the 90% probability of breeding success. Obviously, as we always want to produce the least necessary number of offspring to achieve our target number of animals with the correct genotype, we have to generate additional animals. In Table 1 the required number of offspring of 90% success rate of 4 different Mendelian probabilities are given. When planning a set of breedings, it should be kept in mind, though, that litters come in quantiles of e.g. 6 (depending on strain and mutation), requiring the researcher to decide at times for breedings yielding less than the optimal number of target animals at 90% success probability – thus leading to lower success probability - or breedings yielding more than the optimal number of target animals at 90% success probability – thus leading to higher success probability. Animal numbers required for success probabilities other than 90% can be easily calculated using R (for the script see Material and Methods). Taken together, we can make “power” calculations for breeding outcomes and thus optimize the probability of obtaining the required number of animals for a given genotype.

## Predicting Breeding Outcome for Multiple Group Cases

In many experimental setups multiple groups of animals necessary for an experiment are produced by the same breeding pairs. Again, our example is the use of WT/WT and KO/KO animals from breedings of heterozygous parents. We want to obtain and use 3 WT/WT and 3 KO/KO animals. While this appears at first glance to be identical to the already described case, the use of two genotypes changes predictions as outcome for one genotype depends on outcome of the respective other genotype. This becomes clear in the most extreme (yet possible) outcome cases. When all born pups are WT/WT, then obviously none will be KO/KO (and vice versa). The influence that use of two genotypes from the same set of breedings has on probability of success for obtaining the desired number of animals is shown in Fig. 3A. We plotted the number of born animals versus the success probability of obtaining 3 individuals of each genotype (A and B, both occurring at a 0.25 Mendelian frequency). Below 10 animals success is close to zero, beyond 30 animals no reasonable gains in terms of success can be achieved. The red line indicates the number of pups required to obtain 3 animals of both required genotypes with a 90% probability. Again, we can use this information to “power” our breedings: In Fig. 3B we show the required number of animals born for obtaining the target group size (identical group size for both genotypes) for 4 different Mendelian outcomes with a success probability of 0.9 in comparison to the direct requirements from a Mendelian calculation. The fraction of additional animals that are required is considerably higher than for the single genotype case with 90% success probability (Fig. 2). The exact numbers of required animals are shown in table 2. This first assessment was performed with both required groups of identical size and both required genotypes appearing with identical Mendelian frequencies. For groups of identical or different sizes another graphical approach has to be used: Fig. 3C shows the correlation between the experimentally required number of offspring of two genotypes appearing with the Mendelian frequency of 0.25 (x- and y-axis) and the number of animals necessary to be born for a success probability of 0.9 (the isolines). Here also the required number of offspring can be obtained for situations in which the group sizes are unequal. The same predictions can be generated when the required genotypes have different Mendelian frequencies, as shown in Fig. 3D for Mendelian frequencies of 0.5 and 0.25.

# Discussion

Application of the 3R has become standard, frequently also a legal requirement, in experimental sciences using animals. In this context also practIces in general husbandry including breeding practices need to be scientifically optimized. Further, animals that are produced in husbandries but are not used for experiments are a new focal point of political attention. Many of these surplus animals are result of the very basic biology of mammalian genetics. We show here that for adequate planning of breeding outcomes even more surplus animals will have to be generated than when calculations are based on simple probabilities according to Mendel´s rules due to the stochastic nature of breedings with small colony sizes. Here we show the extent of these additional animals required for breeding success probabilities of 90%. We also show how the common practice of using different genotypes derived from the same breedings influences breeding success and thus requires even further additional animals. We provide tables that facilitate easy breeding planning for the practitioner. The R scripts can be easily adapted for other use cases, such as higher or lower success probabilities. Through simple modification of the respective frequency of occurance, they can also be applied for the planning of breedings where the outcomes are not following strict Mendelian frequencies. An example would be breedings yielding in some allele combinations embryonically lethal genotypes of partial penetrance (e.g. Li 2008 Dev. Cell, Kallapur 1999 Mol Reprod Dev ).



While our calculations take into account the stochastics of obtaining successful breeding outcomes in terms of required genotypes, the scientist who plans his breedings should still take into account the general breeding efficiency of a particular strain. This can be done through the production index (PI, Ayadi 2011, Festing and Peters 1999). A table of production indices of various strains can be found in White 2007.; it is recommended however to establish the index for each mouse line and facility independently. An example how to perform the required calculations can be found in table 3. We introduce an adjustement factor for use of the PI in a timed breeding. Assuming that in a continuous breeding setting females are distributed over the full 3 weeks of gestation, we made the assumption that in a timed mating all these pregnancies occurred together (hence the correction factor of 3). Another method to meet the requirement of obtaining a defined number of animals within a specific period of time was presented by Festing (Festing UFAW Handbook Laboratory Animals Breeding and Genetics 6th edition). It involves the littering percentages at days after start of breeding, the average effective litter size, and the percentage of effective fertility.

# Material and Methods

Table 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **90% Breeding Success** | | | | |
|  | **Genotype probability according to Mendel** | | | |
|  | **0.5** | **0.25** | **0.125** | **0.0625** |
| **Number of animals required for experiment** | **Number of animals required to be born** | | | |
| **1** | 4 | 9 | 18 | 36 |
| **2** | 7 | 15 | 30 | 61 |
| **3** | 9 | 20 | 41 | 84 |
| **4** | 12 | 25 | 52 | 106 |
| **5** | 14 | 30 | 62 | 126 |
| **6** | 17 | 35 | 73 | 147 |
| **7** | 19 | 40 | 82 | 167 |
| **8** | 21 | 45 | 92 | 186 |
| **9** | 24 | 50 | 102 | 206 |
| **10** | 26 | 55 | 111 | 225 |
| **11** | 28 | 59 | 121 | 244 |
| **12** | 31 | 64 | 130 | 263 |
| **13** | 33 | 69 | 140 | 282 |
| **14** | 35 | 73 | 149 | 301 |
| **15** | 37 | 78 | 158 | 319 |

Table 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **90% Breeding Success** | | | | |
|  | **Genotype probabilities according to Mendel** | | | |
|  | **0.5** | **0.25** | **0.125** | **0.0625** |
| **Number of animals required for experiment** | **Number of animals required to be born** | | | |
| **1** | 5 | 11 | 23 | 47 |
| **2** | 8 | 17 | 36 | 74 |
| **3** | 11 | 23 | 49 | 99 |
| **4** | 13 | 29 | 60 | 122 |
| **5** | 16 | 34 | 71 | 144 |
| **6** | 18 | 40 | 82 | 165 |
| **7** | 21 | 45 | 92 | 186 |
| **8** | 23 | 50 | 102 | 207 |
| **9** | 26 | 55 | 112 | 228 |
| **10** | 28 | 60 | 122 | 248 |
| **11** | 30 | 65 | 132 | 268 |
| **12** | 33 | 69 | 142 | 287 |
| **13** | 35 | 74 | 152 | 307 |
| **14** | 37 | 79 | 162 | 327 |
| **15** | 40 | 84 | 171 | 346 |

Table 3

|  |  |  |
| --- | --- | --- |
|  |  | Example |
| Targeted group size | A | 10 |
| Probability according to Mendel | B | 0.125 |
| Animals required to be born according to Mendel | C = A / B | 80 |
| Animals required to be born according to table 1 | D | 111 |
| PI (production index) | E | 0.5 |
| Adjustment factor for timed breeding (versus continuous) | F | 3 |
| Adjusted PI | G = F\*E | 1.5 |
| Age range for experiment | H | 2 weeks |
| Number of females required for obtaining experimental animals | (D/G)/H | 37 |

|  |  |  |
| --- | --- | --- |
|  |  | Example |
| Targeted group size | A | 10 |
| Probability according to Mendel | B | 0.125 |
| Animals required to be born according to Mendel | C = A / B | 80 |
| Animals required to be born according to table 1 | D | 111 |
| average effective litter size | F | 7 |
| percentage of effective fertility | G | 84% [Green and Witham, 1991](http://www.informatics.jax.org/silver/references.shtml" \l "Green6" \t "_blank) |
| No of litters needed | According to Table 3.12 Festing | 21 |
| Number of females required for obtaining experimental animals |  | ca. 30 females |

# Supplementary Material