Arch. Tierz., Dummerstorf **47** (2004) 6, 595-604

Institute for the Biology of Farm Animals Dummerstorf, Germany

GERHARD DIETL, MARTINA LANGHAMMER and ULLA RENNE

Model simulations for genetic random drift in the outbred strain Fzt:DU

Dedicated to Prof. Dr. habil. Peter Glodek on the occasion of his 70th birthday

Abstract

The mouse outbred stock Fzt:DU has been bred in the Research Institute for the Biology of Farm Animals Dummerstorf, Germany for about 30 years. This paper describes the history and the development of some traits in this stock over 128 generations. It has been used as base population for several lines long-term selected for fertility, growth, fitness and behaviour and has been bred with an average number of 200 breeding pairs per generation using a rotational mating scheme. A simulation study was employed to investigate the effect of genetic random drift on the allele frequencies. The change of the drift variance, the probabilities for allele losses and the development of the effective population size over generations are represented. The effective population size was relatively high, compared to other mouse experiments worldwide, however the genetic variability of the Fzt:DU population is substantially reduced due the high number of generations of isolated reproduction. After 120 generations, the variance effective population size is reduced to approximately 3 animals.

Key Words: Long-term outbred mouse stock, simulation, random drift, effective population size, allele loss

Zusammenfassung

Titel der Arbeit: **Modellsimulationen zur genetischen Zufallsdrift des Auszuchtstammes Fzt:DU**

Der Mäuseauszuchtstamm Fzt:DU wird seit ca. 30 Jahren im Forschungsinstitut für die Biologie landwirtschaftlicher Nutztiere Dummerstorf mit 200 Tierpaaren/Generation nach einem Rotationsprinzip gezüchtet und bildete die Basispopulation der auf Fruchtbarkeit, Wachstum, Fitness und Verhalten langzeit-selektierten Mauslinien, die in unserem Labor entwickelt wurden. Die Geschichte des Stammes und die Entwicklung ausgewählter Eigenschaften im Verlaufe von 128 Generationen Zucht werden hier dargestellt. Mittels einer Simulationsstudie wird der Einfluss der genetischen Zufallsdrift auf die Allelfrequenz untersucht. Der Verlauf der Driftvarianz, die Wahrscheinlichkeiten des Allelverlustes sowie die Entwicklung der effektiven Populationsgröße über die Generationen werden gezeigt.

Trotz der im weltweiten Vergleich zu anderen Labormausexperimenten hohen effektiven Populationsgröße ist der Stamm wegen der Vielzahl von Generationen isolierter Reproduktion in seiner genetischen Vielfalt erheblich reduziert. Nach 120 Generationen beträgt die varianzeffektive Populationsgröße nur noch etwa 3 Tiere.

Schlüsselwörter: Langfristig ausgezüchteter Mausstamm, Simulation, Zufallsdrift, effektive Populationsgröße, Allelverlust

Introduction

The outbred laboratory mouse strain Fzt:DU that has been developed in the early seventies in the Research Institute for the Biology of Farm Animals Dummerstorf, Germany, and it was used as a base population for a series of long-term selection experiments. Special selection lines with high fertility (SCHÜLER et al., 1990 a, b), growth (BÜNGER et al., 1998, 2001), behaviour (RENNE et al., 1987; RENNE and

BÜNGER, 1990; RENNE and LANGHAMMER, 2000) and endurance fitness (RENNE and LANGHAMMER, 1999; FALKENBERG et al., 2000) were derived from this strain. The outbred strain was developed by systematic crossbreeding of four inbred and four outbred lines for three generations (Table 1).
Until today the Fzt:DU strain has been reproduced by random mating, but avoiding full sib mating. SCHÜLER (1985) described in detail the used crossbreeding plan from which a similar allele contribution from all founder lines in the outbred strain can be expected.
However, as each population with finite size, the strain was and still is subjected to the genetic random drift. Therefore it is of particular interest to analyse the changes of frequencies of those alleles that were unique in the different founder lines and to describe the development of the effective population size over time.
As the exact pedigree was not recorded a computer simulation program was used based on information of the total number of mating pairs per generation, the average litter size, it's variance, the probability of successful mating and the survival rate until breeding age.

Material and methods

The founder lines used to develop the strain Fzt:DU are given in Table 1. The crossbreeding structure, the number of mating pairs (1 male : 1 female) in each combination is shown in Table 2. In all following generations 200 pairs were mated randomly avoiding full-sib mating. For some generations the development of fertility and growth traits is presented in Tables 3 and 4.

Table 1
The founder lines (Die Gründerlinien)

line	designation	Place of origin
1	NMRI orig	Bayer AG Wuppertal (Dr. Meister)
2	Han: NMRI	″
3	Han: CFW	″
4	CBA/Bln	Institut für Krebsforschung Berlin Buch der DAW (Dr. Horn)
5	AB/Bln	″
6	C57BL/Bln	″
7	XV11/Bln	″
8	Han: CF1	Bayer AG Wuppertal (Dr. Meister)

The first simulation step generated the eight founder populations with 15 litters, each and with a sex ratio of 1:1 and a given number of alleles at the considered gene loci and their frequencies. The allele frequencies were assumed to be equal for all founder populations. The offspring of the founder populations (generation 1 = crossing generation 1) were generated according the mating plan (Table 2, 1st column) by simulation. 120 generations per simulation were generated with 10 000 replicates in each of the simulations.
For the simulation of the data for the next generation a mating ratio of 1:1 was assumed and following probabilities: 0.8 - mating produces a litter and 0.95 - chance to reach breeding age. Pseudo-random numbers (μ=10; s=2.5) were used to simulate the number of born pups per litter from each successful pregnancy.

Since equal allele frequencies were assumed for all founder populations, the obtained results, the drift variances, the effective population sizes and the probabilities of allelic losses or allele fixations were averaged over all founder populations. To describe the effective population size a variance effective population size was used:
Ne = {allele frequency (1-allele frequency)} / (2•drift variance)

Table 2
Breeding program for crossing the 8 original lines (Zuchtprogramm zur Kreuzung der 8 Originallinien)

Number of crossing generation		
1 (500 pairs)	2 (500 pairs)	3 (500 pairs)
(1 x 2)→ 9	(9 x 11)→ 25	(25 x 29)→ 41
(2 x 3)→ 10	(10 x 12)→ 26	(27 x 31)→ 41
(3 x 4)→ 11	(11 x 13)→ 27	(33 x 37)→ 41
(4 x 5)→ 12	(12 x 14)→ 28	(35 x 39)→ 41
(5 x 6)→ 13	(13 x 15)→ 29	(26 x 30)→ 41
(6 x 7)→ 14	(14 x 16)→ 30	(28 x 32)→ 41
(7 x 8)→ 15	(15 x 9)→ 31	(34 x 38)→ 41
(8 x 1)→ 16	(16 x 10)→ 32	(36 x 40)→ 41
(1 x 4)→ 17	(17 x 19)→ 33	
(2 x 5)→ 18	(18 x 20)→ 34	
(3 x 6)→ 19	(19 x 21)→ 35	
(4 x 7)→ 20	(20 x 22→ 36	
(5 x 8)→ 21	(21 x 23)→ 37	
(6 x 1)→ 22	(22 x 24)→ 38	
(7 x 2)→ 23	(23 x 17)→ 39	
(8 x 3)→ 24	(24 x 18)→ 40	

Results

All stocks were transferred by hysterectomy into a new semi-barrier facility before generation 18. The two facilities were too different and traits in the old facility were measured only very rarely in this stock, so that they were left out here. The phenotypic description of the outbred stock Fzt:DU in the new facility up to generation 128 for fertility (Table 3) and growth traits (Table 4) showed no substantial trend in any of the observed traits, although considerable inbreeding had accumulated. The average litter sizes (LS: pups) and weights (LW: g) at birth, 10d and 21d (weaning) were 11.5 pups, 19.2g; 10.4 pups, 56.1g; 10.2 pups and 100.3g, respectively. Average body weights (g) at weaning (BW21) and at 42d (BW42) in males and females were 11.1, 10.8 (BW21) and 29.9, 24.7 (BW42), respectively. Deviations from the overall mean were probably of environmental origin or due to some unintended and weak selection with a preference for bigger litters and more healthy looking animals, which might be somewhat bigger than the average. The low between generation variation indicates a stable and good standardisation of the environment over all these > 100 generations (>25 years).

The development of drift variance and the probabilities of allelic losses over time are shown in Figures 1-7. It is of note that the allele frequencies always refer to one individual founder line. For example an allele frequency of 10% (Fig. 1) indicates that this founder line specific allele occurred at a frequency of 10% in the original founder line, but in the synthetic strain (Fzt:DU) it would have an expected frequency of 1.25%

(=10%/number of founder lines). The given drift variances describe therefore the variability of frequencies of founder-line-specific alleles in the total population.

Table 3
Simple means ($\bar{X}$) and their phenotypic standard deviations (s) for litter size and litter weight (Mittelwerte und phänotypische Standardabweichungen für Wurfgröße und Wurfmasse)

		LSO		LWO		LS10		LW10		LS21		LW21	
gen.	n	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
18	32	10.2	2.2	16.4	3.3	9.5	2.2	44.8	7.6	9.2	2.2	92.2	17.0
19-43	467	11.4	2.5	17.5	3.5	10.3	2.4	53.2	11.3	10.0	2.4	95.5	21.7
44-54	2082	11.4	2.4	18.0	3.5	10.6	2.3	56.5	9.6	10.5	2.3	101.2	21.1
55-64	1900	11.2	2.6	18.2	3.9	10.1	2.7	51.9	14.4	9.9	3.8	92.2	29.0
65-74	1922	11.4	2.6	19.0	4.0	10.7	2.4	56.6	11.4	10.2	2,6	95.8	24.7
75-84	1791	11.4	2.6	19.0	3.9	10.0	2.4	52.5	11.1	9.7	2.5	90.5	22.9
85-94	1758	11.4	2.7	18.6	4.1	10.4	2.4	54.4	10.3	10.1	2.4	95.4	22.1
95-104	1763	11.9	2.8	20.6	4.7	10.8	2.6	60.5	10.5	10.6	2.5	116.7	24.2
105-114	1816	11.6	2.7	20.1	4.2	10.4	2.5	58.8	10.7	10.2	2.5	111.4	22.6
115-124	1845	11.6	2.8	20.2	4.8	10.3	2.6	57.4	11.5	10.0	2.6	107.3	25.2
125-128	736	11.6	2.6	20.4	4.4	10.6	2.4	59.2	11.2	10.4	2.4	107.6	26.9
$b_{18\text{-}128}$		0.005		0.036		0.003		0.035		0.006		0.061	

LSx, LWx: Litter size and litter weight at age x (values are pooled over generations)
$b_{18\text{-}128}$:regression coefficients over generations 18 till 128

Table 4
Simple means ($\bar{X}$) and their phenotypic standard deviations (s) for body weights of males and females (Mittelwerte und phänotypische Standardabweichungen für Körpermassen männlicher und weiblicher Tiere)

		BW21 m		BW21 f		BW42 m		BW42 f	
gen.	n	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
44-54	440	10.4	2.0	10.2	1.9	28.0	3.4	21.5	2.7
55-64	345	10.7	2.1	9.8	2.0	28.9	3.3	22.2	2.7
65-74	238	10.2	4.0	9.6	2.2	29.2	3.9	23.2	2.8
75-84	455	9.7	2.4	9.5	2.1	28.5	3.4	23.8	2.4
85-94	675	10.7	2.4	9.8	2.3	29.9	3.8	22.9	3.1
95-104	691	11.8	2.6	11.3	2.4	31.2	3.3	25.7	2.5
105-114	1693	11.7	2.5	11.4	2.4	30.0	3.3	25.2	2.5
115-124	2290	11.3	2.7	11.0	2.4	30.4	3.5	25.6	2.6
125-128	911	11.0	2.7	10.8	2.6	29.9	3.9	25.2	2.9

BWx: body weight at age x; m=male; f=female (values are pooled over generations)

There is a linear increase of the drift variance over the time the population was maintained. Its increment depends very much on the allele frequencies (Figures 1 and 2). For alleles at a low frequency e.g. 1% the drift variance will increase after 120 generations to c4% only, however for alleles with higher frequencies, e.g. 70% the drift variance can amount to 130% after 120 generations. The case where allele frequencies are 100% refers mainly to alleles, which had been fixed in the four inbred founder lines, which should be quite frequent in the inbred founder populations.

Standard deviations for the estimates for rare alleles (frequencies 0.125 to 1.25%) are given in Table 5; they increase with increasing allele frequencies.

Table 5
Standard deviations [SD%] and coefficients of variation [CV%] of frequencies with rare and very rare alleles in the total population (Standardabweichungen und Variationskoeffizienten für die Frequenz seltener und sehr seltener Allele in der Gesamtpopulation)

generation	allele frequencies 0.125%		0.625%		1.25%	
	SD	CV	SD	CV	SD	CV
5	0.30	240	0.66	106	0.93	74
10	0.43	344	0.96	154	1.33	106
30	0.74	592	1.66	266	2.33	186
60	1.08	864	2.29	366	3.29	263
90	1.34	1072	2.79	446	3.98	318
120	1.58	1264	3.22	515	4.60	368

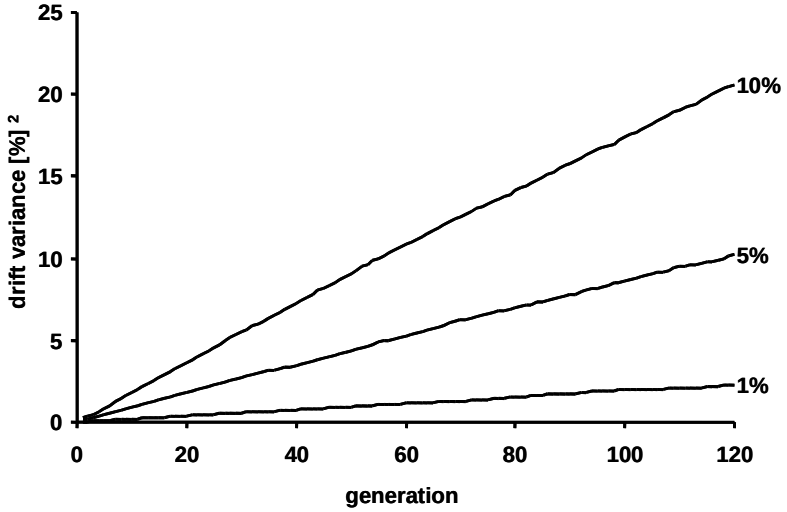

Fig. 1: Drift variance at allele frequencies of 1, 5 and 10 percent, respectively (Driftvarianz bei einer Allelfrequenz von 1, 5 und 10%)

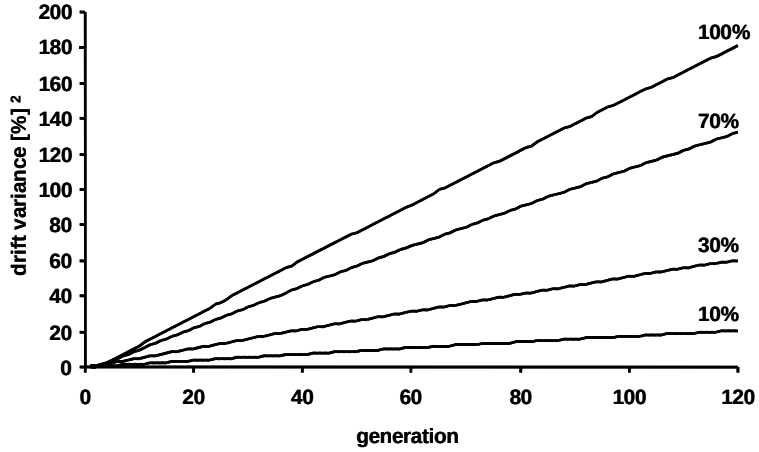

Fig. 2: Drift variance at allele frequencies from 10 to 100 percent (Driftvarianz bei einer Allelfrequenz von 10-100%)

Figures 3 to 7 show likelihood's for the simultaneous loss of alleles in the total population during the experiment in dependence of their initial frequencies and their initial existence in 1 to 8 of the founder lines. As expected, the likelihood's of the loss

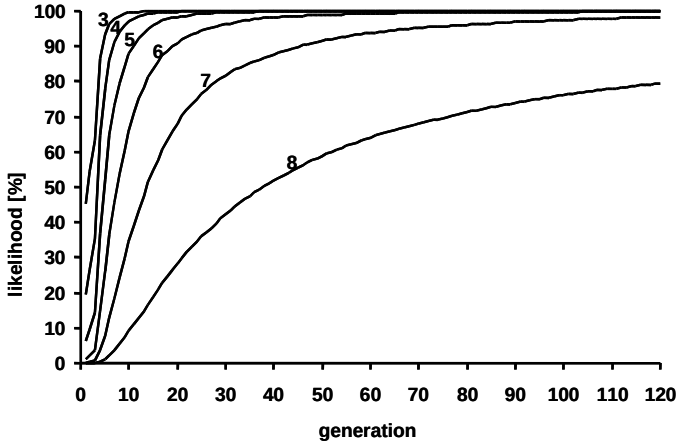

Fig. 3: Likelihood for the simultaneous loss of x allele origins, x=3 to 8, with an allele frequency of 2 percent (Wahrscheinlichkeit für den gleichzeitigen Verlust von x Originalallelen, x=3-8, mit einer Allelfrequenz von 2%)

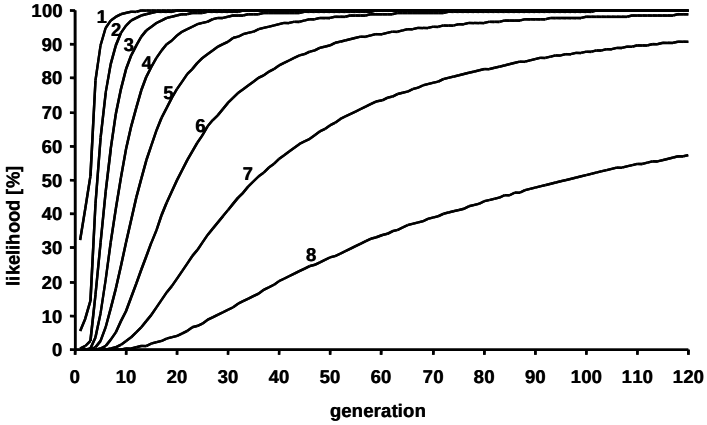

Fig. 4: Likelihood for the simultaneous loss of x allele origins, x=1 to 8, with an allele frequency of 5 percent (Wahrscheinlichkeit für den gleichzeitigen Verlust von x Originalallelen, x=1-8, mit einer Allelfrequenz von 5%)

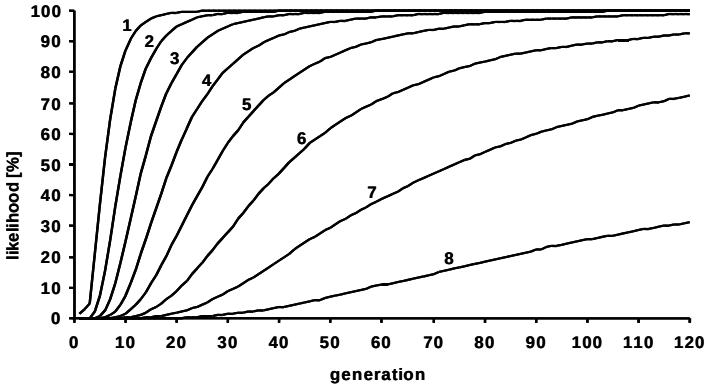

Fig. 5: Likelihood for the simultaneous loss of x allele origins, x=1 to 8, with an allele frequency of 10 percent (Wahrscheinlichkeit für den gleichzeitigen Verlust von x Originalallelen, x=1-8, mit einer Allelfrequenz von 10%)

of rare alleles were higher than those from alleles with high frequencies and the loss of rare alleles was lower when it existed initially in more founder lines. The allelic loss started early and the likelihood's increased strongly.

In Figure 3 the presentation of the probabilities for the cases "at least one founder line" and "at least 2 founder lines" were not included, as they likelihod reached very early

almost 100 percent. Although the loss of alleles of all 8 origins reached a likelihood of 80% in generation 120, the probabilities of fixation were in all examined cases very close to 0. They are therefore not presented.

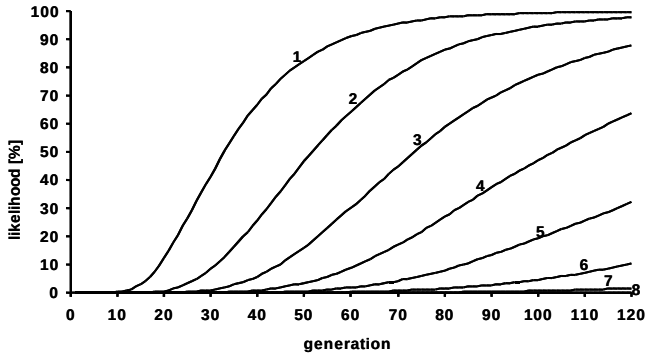

Fig. 6: Likelihood for the simultaneous loss of x allele origins, x=1 to 8, with an allele frequency of 50 percent (Wahrscheinlichkeit für den gleichzeitigen Verlust von x Originalallelen, x=1-8, mit einer Allelfrequenz von 50%)

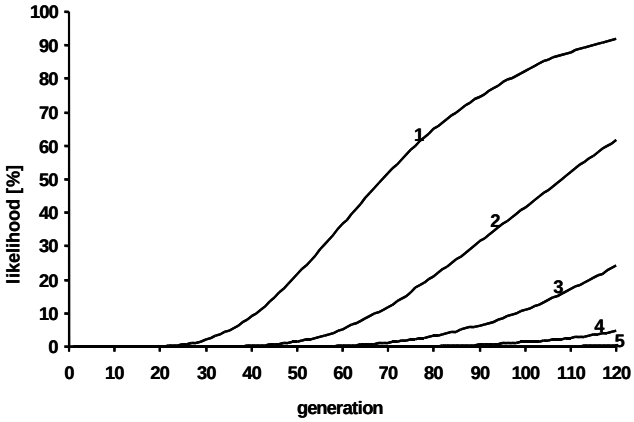

Fig. 7: Likelihood for the simultaneous loss of x allele origins, x=1 to 5, with an allele frequency of 100 percent (Wahrscheinlichkeit für den gleichzeitigen Verlust von x Originalallelen, x=1-5, mit einer Allelfrequenz von 100%)

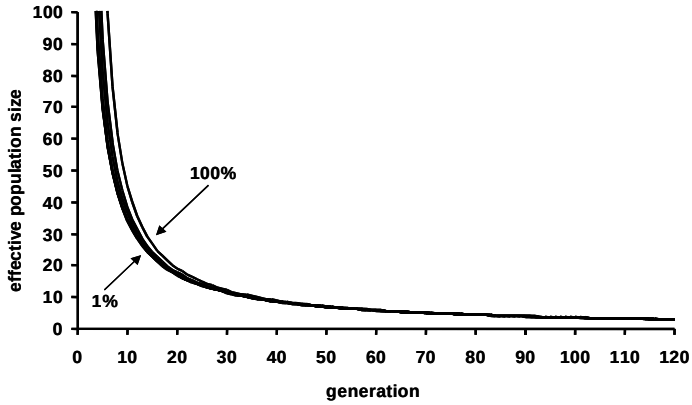

Fig. 8: Variance effective population size at allele frequencies from 1 to 100 percent (Varianzeffektive Populationsgröße bei Allelfrequenzen von 1 bis 100 %)

The development of the variance effective population size (Fig. 8) is reciprocally proportional to the drift variance and decreases mostly in the first 20 generations when

it drops from 100 to below 20 and approaches later on asymptotically 0, with a value of 3 in generation 120.

Discussion

Drift variance: It has been shown that the drift variance increases approximately linearly over generations. This refers to allele frequencies up to 10% but also for higher frequencies up to 100%, indicating that the change in the allele frequencies at 120 generations seems to be far from restricting factors and far from the theoretical maximum of 1250 [%]². The variance of allele frequency changes, the drift variance, is $p\ (1-p)/(2N)$, with p the initial allele frequency and N the number of individuals in the population. This function reaches a maximum with minimal N and p=0.5. Using the initial allele frequency in percent means the units for the maximum are %2 and it can be calculated from 50% (100%-50%) with N approaching a minimum.

From the development shown it can be expected that the drift variance of the strain Fzt:DU will still increase approximately in a linear way in the next 50 generations and the increase will remain stronger for higher frequencies than for lower frequencies.

The increase from generation 119 to generation 120 for an allele frequency of 100% will be 0.90 [%] ², whereas it would be with an allele frequency of 10% 0.11 [%] ², only. Although this experiment has been based on a relative high number of breeding pairs, compared to other 'outbred' mouse populations, the drift variances are high, with coefficients of variation of mostly much over 100% (Table 5). This suggests very skewed distributions at allele frequencies close to 0 or 1, making extremely high alterations of allele frequencies more likely.

Allelic loss: A high loss of alleles with low frequencies in the founder lines certainly occurred in the first 10 to 20 generations of the strain Fzt:DU (Figures 3 and 4). But also in later generations there was a high likelihood for substantial loss of alleles, that initially occurred with an allele frequency of approximately 10 percent (Figure 5).

At 120 generations, it has to be expected that also alleles with higher initial frequencies in the various founder lines disappeared from the gene pool of the strain (figure 6). The probability for an allele loss for alleles for which the founder lines had been originally homozygous amounts to 93 percent after 120 generations. The simultaneous loss of initially homozygous alleles of 2 founder lines is likely with 61.7 percent and correspondingly that of 3 fixed in the founder lines reaches a likelihood of 24.4 percent (figure 7). Therefore it has to be assumed that not all line-specific alleles are still completely present in the synthetic strain Fzt:DU. This applies also to all inbred founder lines, which had a high degree of homozygosity.

Fixation of alleles: The results show clearly that the fixation of an allele of one single origin in the population is extremely unlikely after 120 generations of random mating. Knowing that this likelihood will show a sigmoid development over generations, it can be predicted that there will be no substantial changes in the next lets say 100 generations.

Variance effective population size: The development of the variance effective population size (Figure 8) shows a substantial loss of genetic variability that is unavoidable in small closed populations. After 120 generations, the variance effective population size has been reduced from over 100 to approximately 3 animals. Such development is essentially unavoidable and not much dependent on the way the strain

was created and its subsequent reproduction was managed (mating ratio of 1:1; avoiding of full-sib mating).
This process will continue even if the loss of genetic variability is now in an asymptotic stage.

Conclusions

The genetic random drift leads to a very substantial population dynamics in small populations. This seems especially important if a strain was created by crossing several heterogeneous founder lines.
The outbred strain Fzt:DU has and had in all generations a unique genetic structure. This refers also to points when the various long-term selection lines were initiated at different times. This implies these long-term selection lines had by chance different and non reproducible ‘genetic scenarios’ at the outset and might be not fully comparable from that point of view, which could be of importance for recent marker based linkage studies.
The outbred strain Fzt:DU has a considerably reduced genetic variability due to high number of generations of isolated reproduction. This process will continue, although with much lower pace, but it will become genetically slowly but steadily close to an inbred line.
New selection lines should better be derived from a new heterogeneous outbred population, unless there is a special scientific purpose justifying this choice.
The further maintenance of the Fzt:DU strain needs to be justified as it is of restricted use only.
The question arises in how far intra-family selection could have delayed the development.
One could also develop a maintenance strategy of such base populations incorporating the repeated introgression of genes from the original founder lines at certain times.
Although the chosen approach ignores the effect of *de novo* mutations on the genetic variability it seem impressive how little apparent genetic variability is left after more than 100 generations in a closed population of even 200 breeding pairs.

References

BÜNGER, L.; RENNE, U.; DIETL, G.; KUHLA, S. :
Long-term selection for protein amount over 70 generations in mice. Genetical Research **72** (1998), 93-109

BÜNGER, L.; RENNE, U.; BUIS, R. C.:
Body weight limits in mice - Long-term selection and single genes. Encyclopedia of Genetics (REEVE, E.C.R., ed.), Fitzroy Dearborn Publishers, (2001), 337-360

FALKENBERG, H.; LANGHAMMER, M.; RENNE, U.:
Comparison of biochemical blood traits after long-term selection on high or low locomotory activity in mice. Arch. Tierz., Dummerstorf **43** (2000), 513-522

RENNE, U.; BÜNGER, L.:
Ergebnisse einer Langzeitselektion auf die Open-field-Aktivität von Labormäusen. 8. Leipziger Tierzuchtsymposium Populationsgenetische Grundlagen und ihre Umsetzung in der Tierzucht, Vortragssammelband (1990)

RENNE, U.; BÜNGER, L.; SCHÜLER, L.:
Experimentelle Untersuchungen an Labormäusen zur Selektion auf physische und "psychische" Belastbarkeit - direkte Selektionserfolge. Arch. Tierz., Berlin **30** (1987), 321-328

RENNE, U.; LANGHAMMER, M.:
Selection response after long-term selection for high and low running activity in mice with special consideration of a selection limit. Proceedings of the 50th Annual Meeting of the EAAP, Zürich 22-26 August (1999), GPh5.12:58

RENNE, U.; LANGHAMMER, M.:
Effects of long-term selection for exploratory behaviour of mice in an unknown surrounding. Arch. Tierz., Dummerstorf **43** (2000) Special Issue, 231

SCHÜLER, L.:
Der Mäuseauszuchtstamm Fzt:DU und seine Anwendung als Modell in der Tierzuchtforschung. Arch. Tierz., Berlin **28** (1985) 4, 357-363

SCHÜLER, L.; RENNE, U.; BÜNGER, L.:
Selection for litter weight on the 21st day after long-term selection for first litter performance in laboratory mice. J. Anim. Breed. Genet. **107** (1990a), 161-168

SCHÜLER, L.; RENNE, U.; BÜNGER, L.:
Kreuzungswirkungen bei Merkmalen der Fruchtbarkeit und des Wachstums nach Langzeitselektion auf Erstwurfleistung bei der Labormaus. J. Anim. Breed. Genet. **107** (1990b), 241-248

Received: 2004-08-03

Accepted: 2004-09-22

Author's address
Dr. habil. GERHARD DIETL, Dr. MARTINA LANGHAMMER, Dr. ULLA RENNE*
Forschungsinstitut für die Biologie landwirtschaftlicher Nutztiere
Forschungsbereich Genetik und Biometrie
Wilhelm-Stahl-Allee 2
D-18196 Dummerstorf
Germany

*Corresponding autor