# Inbreeding in a Family Tree and in a Population

A.N. Volobuev[1,*]

[1]Samara State Medical University, Department of Medical Physics, Samara, Russia

**Abstract**

On the basis of Hardy – Weinberg's law the problem of inbreeding in a family tree and a population was investigated. With use of an inbreeding factor are received the discrete equation for a family tree and differential equation for a population. The numerical solution of the differential equation for a population was found and analyzed at various values of the inbreeding factor. Migration of inbred population is investigated in view of natural selection. It was shown that velocity of migration falls with increase of the inbreeding factor. Interrelation of the recessive allele frequency at woman for a migrating population with inbreeding factor and standard parameter of selection was found.

## Introduction

Problem of inbreeding is one of the major problems of genetics. Since times of the human occurrence the population size was very small. People lived separate tribes. Therefore, consanguineous mating or inbreeding has been submitted enough. In process of the population size increase the intensity of such mating was reduced. But till now the consanguineous mating are present at small communities.

Inbreeding are widely submitted at animals. Very much frequently the man uses consanguineous mating of animals at artificial selection to receive useful properties of animals: speed of run, force, fleshiness, a plenty of milk, a wool, etc.

Let's consider principles of inbreeding in a family trees and a population.

### Inbreeding in a Family Tree

There are family trees and populations where the principle panmictic mating is essentially broken. Usually it is geographically or socially relatively isolated small communities. In these communities the inbreeding i.e. consanguineous mating of various orders (a parent - a child, sibs, cousins, uncle - niece, etc.), can be wide-spread.

The problem of inbreeding influence on the population is the important problem of genetics [1]. Now this problem yet has not received the final solution.

In most cases the inbreeding in a family tree take into account on three generations. It is the most distant relatives who are taken into account at consanguineous mating there is second cousins.

The inbreeding is taken into account with the help of inbreeding factor.

The inbreeding factor is the probability two alleles are identical by origin.

Let's consider the method of factor inbreeding factor calculation in a family tree offered by Wright [1, 2].

For the inbreeding factor finding it is necessary:

- to find for the marrying descendants all ways in a family tree which through the common ancestor connect these descendants.

- to count up number of steps (alternations of generations) on each way.

Then the inbreeding factor is equal:

$$F_t = \frac{1}{2}\sum_{i=1}^{r} 2^{-m_i}$$

......(1)

where $r$ there is number of ways, $m_1, m_2, ... m_i$ - number of steps on each way.

Once again we shall note that the number of addends in the sum (1) is equal to number of ways which through the common ancestors connect marrying.

Let's consider concrete cases of family trees, fig. 1. In this figure the square means a male individual the circle – female individual. Dotted lines and identical digits specify possible ways in a family tree which through the common ancestor (the man or the woman) connect of the descendants consanguineous marrying. On some sites of ways, fig. 1c, there are two digits 1 and 2. It specifies that such sites are the common for the various ways.

The marriage a father - daughter is incest, fig. 1a. The descendant (woman) is connected with the ancestor (man) there 1 way with one step, i.e. $m = 1$. Hence, there is inbreeding factor.

$F_t = 1/2.2^{-1} = 1/4$

There is marriage a brother - sister (sibs), fig. 1b. Digits specify the number of the way. The numbers of identical digits on the way from the marrying man to the woman is equal the number of steps.

Let each descendant who has married connects with partner two ways with two steps on the each way, i.e. $m = 2$. Hence, there is inbreeding factor.

$F_t = 1/2.(2^{-2} + 2^{-2}) = 1/4$

The marriage between cousins is shown on fig. 1c.

Each descendant who has married connects with partner the 2 ways with 4 steps on the each way, i.e. $m = 4$. Hence, there is inbreeding factor.

$F_t = 1/2.(2^{-4} + 2^{-4}) = 1/16$

The formula (1) is easy for understanding from the following logic.

Obviously there 1/2 genes of the child is the common with each parent, 1/4 - with the grandfather
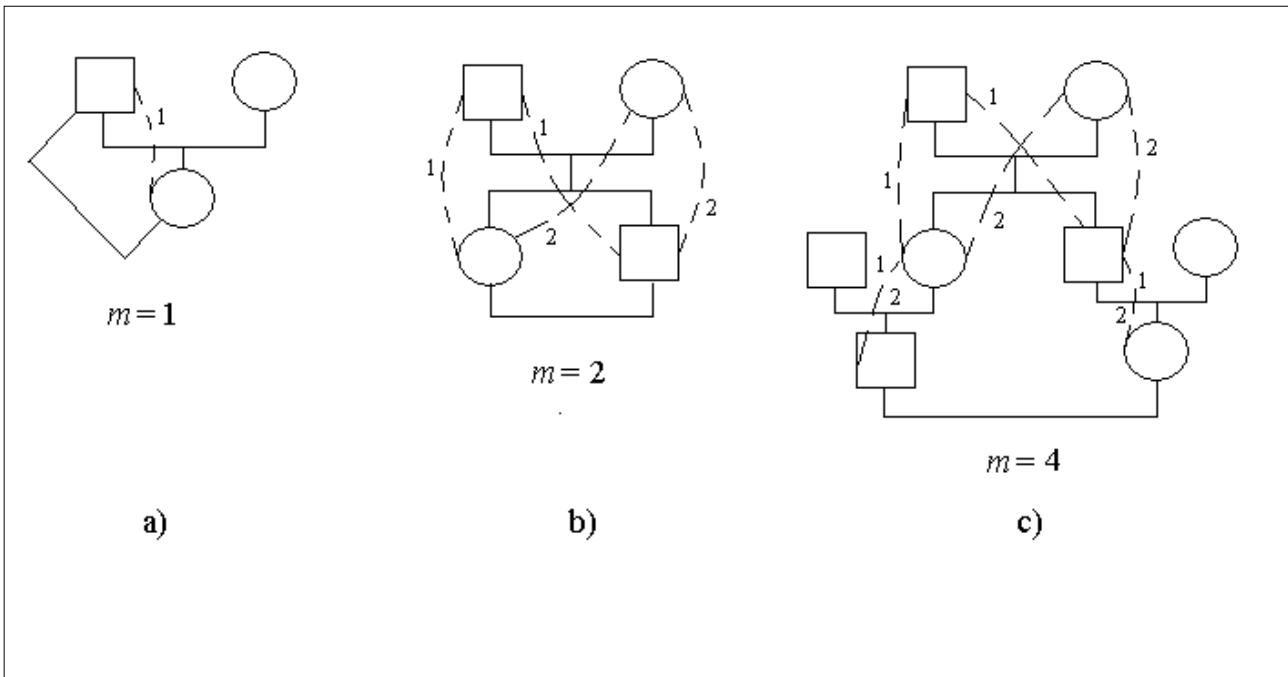
Figure 1. Some kinds of inbreeding in a family tree

and grandmother, 1/8 - with the great-grandfather and great-grandmother, etc. If from the woman to any ancestor leads the way with $\eta$ steps the share of the common genes with this ancestor is equal $2^{-\eta}$. If from the man to any ancestor leads the way with $\mu$ steps the share of the common genes is equal $2^{\mu}$. At the marriage of the man and woman with the common ancestor the share of the common genes with this ancestor will be $2^{\eta}$. $2^{\mu} = 2^{m}$ where m = $\eta + \mu$ there is total number of steps through the common ancestor from the woman up to the man marrying. The size $1/2\ 2^{-m}$ shows probability that the gene casually chosen at the woman will be identical by origin to the gene casually chosen at the man. The same size gives the contribution to the inbreeding factor $F_t$ one way through the common ancestor from the woman up to the man marrying.

For example, if the common ancestor will be away from marrying on $\eta = \mu = 5$ generations the probability them to have the common gene is equal. $1/2\ 2^{-(5+5)} = 1/2048$ This size at the analysis of mating can be neglected. (Table 1)

The inbreeding factor $F_t$ is the probability of two alleles in posterity are identical by origin i.e. occur from same gene of ancestor [1]. Or, that too most, the inbreeding factor is a probability of the allele origin in posterity is consanguineous. Hence, the probability of

the non-consanguineous origins of identical alleles in homologous chromosomes is equal 1 - $F_t$.

Let's find total probability of the recessive homozygote $aa$ occurrence (event $aa$) at random mating for an autosomal genes [3]:

$$P(aa) = P(H_1)P\left(\frac{aa}{H_1}\right) + P(H_2)P\left(\frac{aa}{H_2}\right) \quad \ldots\ldots(2)$$

This formula follows from the following logic of mating: the origin of identical alleles in homologous autosomes is non-consanguineous (the probability of this hypothesis is $P(H_1) = 1 - F_t$ )also arises homozygote $aa$ (conditional probability of this event is $P(aa/H_1) = q^2$) or origin of identical alleles is consanguineous (the probability of this hypothesis is $P(H_2) = F_t$ ) and arises homozygote $aa$ (conditional probability of this event is $P(aa/H_2) = q$). The probability $q$ of the homozygote $aa$ occurrence in the latter case is determined by that if allele $a$ is in one autosome then at consanguineous origin of homozygote same allele necessarily should be and in the homologous autosome.

Hence, the total probability of the recessive homozygote occurrence is equal:

$$P(aa) = (1 - F_t)q^2 + F_t q = q^2 + F_t pq \quad \ldots\ldots\ldots(3)$$

where it is taken into account $p = 1 - q$ there is the dominant allele $a$ frequency.

Similarly there find total probability of the dominant homozygote $AA$ occurrence at random mating:

$$P(AA) = (1 - F_t)p^2 + F_t p = p^2 + F_t pq \quad \ldots\ldots\ldots (4)$$

Thus, according to (3) and (4) total frequency of homozygotes (dominant and recessive) in inbred family tree increases on $2F_t$ $pq$. Hence, frequency of heterozygotes decreases for the same size and to become equal $2pq - 2pqF_t = 2pq(1-F_t)$ since the sum of the all zygote frequencies is equal 1.

In view of inbreeding factor and also (3) and (4) the Hardy - Weinberg law in the following generation of inbred family tree for an autosomal genome needs to be written down the distributions of zygote genotypes as [4]:

$$(AA)(p_n^2 + F_t p_n q_n) : (Aa)[2 p_n q_n(1 - F_t)] : (aa)\left(q_n^2 + F_t p_n q_n\right) \quad \ldots\ldots\ldots(5)$$

Where $p_n$ is the dominant alleles $A$ frequency in generation $n$, $q_n$ - the recessive alleles $a$ frequency, so $p_n + q_n = 1$.

The ratio (5) refers to also Wright's ratio [5]. Wright's ratio shows that in an inbred family tree after the consanguineous mating of individuals are observed the following effects: there is increase in the autosomal homozygotes fraction in a family tree on $F_t p_n q_n$ and reduction of the autosomal heterozygotes fraction in $(1-F_t)$ ones in comparison with a panmictic population.

But already through one generation the Hardy – Weinberg balance is restored that specifies the small importance of the separate consanguineous mating for a family tree:

$$q_{n+1} = \frac{1}{2} 2 p_n q_n(1 - F_t) + \left(q_n^2 + F_t p_n q_n\right) = q_n \quad \ldots\ldots\ldots(6)$$

Obviously, for a population as whole the separate act of the consanguineous mating will have still smaller importance than for a family tree.

The same effect of variation of zygote genotypes frequency is observed and for sex-linked genes.

Distribution of the genotypes linked to the $X$-chromosome in panmictic family tree is described by Hardy – Weinberg law as [1]:

$$(AA) p_f p_m : (Aa)(p_m q_f + p_f q_m) : (aa) q_m q_f \quad \ldots\ldots\ldots(7)$$

The frequency of a dominant alleles $A$ we shall designate at the men $p_m$ and at the women $p_f$. For recessive alleles $a$ it is accordingly $q_m$ and $q_f$.

At mating there is a ratio of genotypes at the women according to product $(p_f + q_f)(p_m + q_m)$.

Let's find the total probability of the recessive homozygote $aa$ occurrence at random mating for the daughter's genes linked to the $X$-chromosome under the formula (2). The logic of events is similar to the case for autosomes except that the conditional probability of a genotype $aa$ occurrence from the non-consanguineous alleles according to (7) is equal $P(aa/H_1) = q_{mn} q_{fn}$. For the consanguineous alleles the conditional probability of a genotype $aa$ occurrence is equal $P(aa/H_2) = q_{f(n+1)}$. At daughters, i.e. in generation $n+1$ the probability of a homozygote $aa$ occurrence is determined by that if allele $a$ is in one $X$-chromosome then at consanguineous origin of homozygote same allele necessarily should be and in the homologous $X$-chromosome.

Hence, the total probability of the recessive homozygote occurrence is equal:

$$P(aa) = P(H_1)P\left(aa/H_1\right) + P(H_2)P\left(aa/H_2\right) =$$
$$= (1 - F_t)q_{mn}q_{fn} + F_t q_{f(n+1)} = (1 - F_t)q_{mn}q_{fn} + F_t \frac{1}{2}\left(q_{fn} + q_{mn}\right) =$$
$$= q_{mn}q_{fn} + \frac{F_t}{2}\left(q_{fn} + q_{mn} - 2q_{mn}q_{fn}\right) = q_{mn}q_{fn} + \frac{F_t}{2}\left(p_{mn}q_{fn} + p_{fn}q_{mn}\right) \quad \ldots\ldots(8)$$

where $p_{mn}$ is the dominant alleles $A$ frequency in generation $n$ at men and $p_{fn}$ - at women. For the recessive alleles $a$ accordingly $q_{mn}$ and $q_{fn}$. Besides it is used $q_{mn} = 1 - q_{mn}$ and $p_{fn} = 1 - q_{fn}$.

Similarly finding the total probability of the dominant homozygote occurrence with use:

$$p_{f(n+1)} = \frac{1}{2}\left(p_{fn} + p_{mn}\right)$$

$$P(AA) = P(H_1)P\left(AA/H_1\right) + P(H_2)P\left(AA/H_2\right) =$$

$$= (1 - F_t)p_{mn}p_{fn} + F_t p_{f(n+1)} = (1 - F_t)p_{mn}p_{fn} + F_t\left[\frac{1}{2}\left(p_{fn} + p_{mn}\right)\right] =$$

$$p_{mn}p_{fn} + \frac{F_t}{2}\left(p_{fn} + p_{mn} - 2p_{mn}p_{fn}\right) = p_{mn}p_{fn} + \frac{F_t}{2}\left(p_{mn}q_{fn} + p_{fn}q_{mn}\right) \quad \ldots\ldots(9)$$

and a heterozygote P(Aa) = 1− P(aa) - P(AA)  we shall find distribution of the genotypes in *X*-chromosomes of the inbred family tree at mating:

$$(AA)\left[p_{fn}p_{mn} + (p_{mn}q_{fn} + p_{fn}q_{mn})\frac{F_t}{2}\right] : (Aa)[(p_{mn}q_{fn} + p_{fn}q_{mn})(1 - F_t)] :$$

$$(aa)\left[q_{mn}q_{fn} + (p_{mn}q_{fn} + p_{fn}q_{mn})\frac{F_t}{2}\right] \quad \ldots\ldots(10)$$

It is simultaneously demonstrated that as well as in autosomes the inbred additives in dominant and recessive homozygote frequencies in the sex-linked chromosomes are identical.

Other way of the ratio (10) finding will consist in the following.

Let's find first the total probability of a heterozygote *Aa* occurrence, i.e. P(Aa)  at the opportunity of the random consanguineous mating. This probability follows from the following logic of mating: the origin of alleles is non-consanguineous (the probability of this hypothesis is P(H$_1$) = 1 - $F_t$ ) also arises heterozygote *Aa* (conditional probability of this event is P(Aa/H$_1$) = $p_{mn}q_{fn} + p_{fn}q_{mn}$, see the formula (7)) or origin of alleles is consanguineous (the probability of this hypothesis is P(H$_2$) = $F_t$) and arises heterozygote *Aa* (conditional probability of this event is equal to zero P(Aa/H$_2$) = 0 ). Last statement is defined by that at consanguineous origin similar alleles in homologous *X*-chromosomes cannot the heterozygote *Aa* arise consisting different alleles.

Hence, using the formula of total probability we find total probability of the heterozygote *Aa* occurrence:

$$P(Aa) = P(H_1)P\left(Aa/H_1\right) + P(H_2)P\left(Aa/H_2\right) =$$

$$= (1 - F_t)\left(p_{mn}q_{fn} + p_{fn}q_{mn}\right) + F_t \cdot 0 = (1 - F_t)\left(p_{mn}q_{fn} + p_{fn}q_{mn}\right) \quad \ldots\ldots(11)$$

Reduction of a heterozygotes *Aa* frequency occurs as has been proved earlier due to identical increase in dominant *AA* and recessive *aa* homozygotes frequency. For example, the increase in recessive *aa* homozygotes frequency is equal:

$$\Delta P(aa) = -\frac{1}{2}\Delta P(Aa) = -\frac{1}{2}\left[-F_t\left(p_{mn}q_{fn} + p_{fn}q_{mn}\right)\right] =$$

$$= \frac{F_t}{2}\left(p_{mn}q_{fn} + p_{fn}q_{mn}\right) \quad \ldots\ldots(12)$$

where reduction of the heterozygotes *Aa* frequency relative of panmictic family tree according to (7) and (11) is equal. ΔP(Aa) = - F$_t$ ($p_{mn}q_{fn}$ + $p_{fn}q_{mn}$) Using (11) and (12) to update (7) for cases of inbreeding we find the distribution of genotypes (10) in *X*-chromosomes in the daughters inbred family tree.

In following generation the frequency of the recessive alleles *a* at women is equal:

$$q_{f(n+1)} = \frac{1}{2}(p_{mn}q_{fn} + p_{fn}q_{mn})(1 - F_t) + q_{mn}q_{fn} + (p_{mn}q_{fn} + p_{fn}q_{mn})\frac{F_t}{2} =$$

$$= \frac{1}{2}\left[(1 - q_{mn})q_{fn} + (1 - q_{fn})q_{mn}\right](1 - F_t + F_t) + q_{mn}q_{fn} = \frac{1}{2}\left[q_{fn} + q_{mn}\right] \qquad \text{........(13)}$$

That fully complies with same frequency at absence of inbreeding, i.e. there is restoration of the Hardy – Weinberg balance or is more exact approach to the balance.

The ratio (7) and (13) show as a whole the allele frequencies in a family tree for the account of inbreeding do not vary. There is only an alleles redistribution. A part of alleles leave the heterozygotes and pass to the homozygotes.

*Genetic Load in a Family Tree*

In the inbred populations so-called the genetic load frequently connected to harmful mutations only recessive alleles passed to offspring is accumulates.

It is connected to increase in the recessive homozygotes *aa* frequency in the inbred populations or family tree, see (3) or (8), in comparison with panmictic. The increase of inbreeding can increase frequency of recessive homozygotes for mutant alleles.

The concept of the genetic load in the population was used for the first time the known geneticist Nobel laureate H. J. Muller [6]. However, actually the theory of the genetic load has been developed only for a family tree.

The consanguineous mating in a population increase homozygotes frequency of harmful mutant alleles. Therefore, it is supposed that inbreeding is one of principal causes of occurrence and accumulation of the genetic load.

The genetic load can be expressed in lethal equivalents. The lethal equivalent is, for example, one lethal mutation resulting in death of an individual in all cases or two lethal mutations each of which resulting in death of the individuals in 50% cases, etc.

At research of the genetic load in a family tree the action of selection we shall not take into account to not complicate the analysis. It is allowable since selection submits to other laws rather than occurrence of mutations.

According to the genetic load theory the formula (10) needs to be transformed [1]:

$$(AA)\left[p_{fn}p_{mn} + (p_{mn}q_{fn} + p_{fn}q_{mn})\frac{F_t}{2}\right]:$$
$$:(Aa)\left[(p_{mn}q_{fn} + p_{fn}q_{mn})(1 - F_t)(1 - sh)\right]:$$
$$:(aa)\left[q_{mn}q_{fn} + (p_{mn}q_{fn} + p_{fn}q_{mn})(1 - s)\frac{F_t}{2}\right], \qquad \text{........(14)}$$

where *s* there is probability of the recessive homozygote destruction as consequence of the inbreeding, *sh* - probability of the heterozygote destruction as consequence of the inbreeding, *h* - so-called dominant degree of the mutation which have occurred owing to inbreeding. If h = 1 there are heterozygotes and recessive homozygotes perish with the identical frequency, if h = 0 the heterozygotes are resistant to inbreeding. Dominant homozygotes are assumed absolutely resistant to mutations. It is mutational model of process.

Let's notice that the ratio (14) concerns to a family tree as separate usually small making population.

Using (14) we shall find the recessive allele *a* frequency at women in the following generation:

$$q_{f(n+1)} = \frac{1}{2}(p_{mn}q_{fn} + p_{fn}q_{mn})(1 - F_t)(1 - sh) +$$

$$+ q_{mn}q_{fn} + (p_{mn}q_{fn} + p_{fn}q_{mn})(1 - s)\frac{F_t}{2} =$$

$$= (p_{mn}q_{fn} + p_{fn}q_{mn})\left[\frac{1}{2}(1 - F_t)(1 - sh) + (1 - s)\frac{F_t}{2}\right] + q_{mn}q_{fn} =$$

$$= \frac{1}{2}\left[(1 - q_{mn})q_{fn} + (1 - q_{fn})q_{mn}\right]\left[1 - sh - F_t + F_t sh + F_t - F_t s\right] + q_{mn}q_{fn} =$$

$$= \left(\frac{1}{2}(q_{fn} + q_{mn}) - q_{mn}q_{fn}\right)(1 - sh + F_t sh - F_t s) + q_{mn}q_{fn} =$$

$$= \frac{1}{2}(q_{fn} + q_{mn}) - \left(\frac{1}{2}(q_{fn} + q_{mn}) - q_{mn}q_{fn}\right)(sh - F_t sh + F_t s) =$$

$$= \frac{1}{2}(q_{fn} + q_{mn}) - \frac{1}{2}(q_{fn} + q_{mn})(sh - F_t sh + F_t s) + q_{mn}q_{fn}(sh - F_t sh + F_t s) =$$

$$= \frac{1}{2}(q_{fn} + q_{mn})(1 - sh + F_t sh - F_t s) + q_{mn}q_{fn}(sh - F_t sh + F_t s) \quad \ldots\ldots(15)$$

In the following generation there is no restoration of the Hardy - Weinberg balance, i.e. accumulation of the genetic load in a family tree, and hence and in a population is observed.

Let's enter the designation $F = sh - F_t sh + F_t s$ - factor of inbreeding for the population. We shall note that if the heterozygotes are resistant to mutations owing to inbreeding $h = 0$ , and recessive homozygotes are perish with probability $s = 1$ the inbreeding factors of the populations and family tree are equal $F = F_t$. Thus the formula (15) can be written down as:

$$q_{f(n+1)} = \frac{1}{2}(q_{fn} + q_{mn})(1 - F) + q_{mn}q_{fn}F \quad \ldots\ldots(16)$$

Lack of the formula (16) is absence normalization on the general fraction of genes less than 1 the rest after lethal mutations. However the standard way of normalization results in too complex formulas which do not allow proceed to the analysis of the population.

For updating the formula (16) it is used the following logic. In the condition of balance of the family tree with inbreeding the formula (13) for calculation of the recessive allele $a$ fraction is correct. The formula (13) is absolutely correct since is normalized on the general portion of genes equal 1.

Taking into account (13), and also F<<1 and small size of the second addend in the right part (16) it is possible to assume that for normalization (16) it is necessary to divide the right part into value 1 - F .

Therefore, the corrected formula (16) looks like:

$$q_{f(n+1)} = \frac{1}{2}(q_{fn} + q_{mn}) + q_{mn}q_{fn}\frac{F}{1 - F} \quad \ldots\ldots(17)$$

Accuracy of the formula (17) can be estimated as follows. Using the formula similar (13) for frequency of dominant allele $A$ in generations of women and summing it up with the formula (13) for frequency of recessive allele $a$ we find $p_{f(n+1)} + q_{f(n+1)} = 1/2(p_{fn} + q_{fn} + p_{mn} + q_{mn}) = 1$ that it is possible to assume a condition of normalization. If instead of the formula (13) to lead similar transformations with the formula (16) the condition of normalization is not carried out even in the main (first) order on allele frequencies. For the formula (17) condition of normalization in the main

order is carried out. The mistake of normalization is O = ($p_{mn}p_{fn}$ + $q_{mn}q_{fn}$) F/(1 - F)  i.e. concerns to higher (second) order of smallness on allele frequencies.

The formula (17) allows find the recessive allele *a* fraction at women in the following generation of the family tree as the sum of equilibrium addend and gradually increasing owing to inbreeding the genetic load.

*Inbreeding in a Population*

Let's carry out the analysis of the inbreeding action on a population.

At research of inbreeding in populations the action of selection we shall not take into account (it will be taken into account later) to not complicate the analysis. It is allowable since selection submits to other laws rather than occurrence of mutations owing to inbreeding.

Using a standard technique of transition from a family tree to a population [7] in the basis of the analysis we shall put the formula (17) which allows to calculate fraction of the recessive allele *a* in the *X*-chromosome of women in the following generation as the sum of the equilibrium term and gradually increasing genetic load owing to inbreeding.

Taking into account that the man inherits the *X*-chromosome from mother i.e. $q_{mn}$ = $q_{f(n-1)}$ it is possible to transform (17) to kind:

$$q_{fn} - 2q_{f(n+1)} + q_{f(n-1)} = -2\frac{F}{1-F}q_{f(n-1)}q_{fn} \qquad ......(18)$$

where *F* there is inbreeding factor for a population.

The differential equation for change of frequency $q_{fn}$ of recessive allele *a* in the inbred population we shall write down in general form:

$$\frac{d^2 q_{fn}}{dn^2} + \eta\frac{dq_{fn}}{dn} = \mu q_{fn}^2 \ , \qquad .......(19)$$

where the independent variable *n* there is in this case time of the population life normalized on average on family trees of the population time of one alternation generation (approximately 25 years) i.e. actually continuous dimensionless time, η  and μ  - constants.

Let's pass in (19) to finite-differential form:

$$\frac{q_{f(n+1)} - 2q_{fn} + q_{f(n-1)}}{\Delta n^2} + \eta\frac{q_{fn} - q_{f(n-1)}}{\Delta n} = \mu q_{fn}^2 \qquad ........(20)$$

Let's collect the like terms:

$$q_{f(n+1)} - (2 - \eta\Delta n)q_{fn} + (1 - \eta\Delta n)q_{f(n-1)} = \mu\Delta n^2 q_{fn}^2 \qquad ........(21)$$

Let's multiply (21) on  -2  besides we shall replace   $q_{fn}^2 = q_{f(n-1)}q_{fn}$

that is allowable for continuous in time of the generation alternation scale i.e. for a population:

$$-2q_{f(n+1)} + 2(2 - \eta\Delta n)q_{fn} - 2(1 - \eta\Delta n)q_{f(n-1)} = -2\mu\Delta n^2 q_{f(n-1)}q_{fn} \qquad .........(22)$$

Let's identify (22) and (18). In result we have system of three algebraic equations:

$$2(2 - \eta \Delta n) = 1 \quad \text{......(23)}$$

$$-2(1 - \eta \Delta n) = 1 \quad \text{.........(24)}$$

$$\mu \Delta n^2 = \frac{F}{1-F} \quad \text{.........(25)}$$

The solution of the given system looks like:

$$\eta = \frac{3}{2\Delta n}, \quad \text{.........(26)}$$

that similarly the equilibrium panmictic population, [7].

$$\mu = \frac{F}{\Delta n^2 (1-F)} \quad \text{.......(27)}$$

Hence, the differential equation (19) can be copied as:

$$\frac{d^2 q_f}{dn^2} + \frac{3}{2\Delta n} \frac{dq_f}{dn} = \frac{F}{\Delta n^2 (1-F)} q_f^2 \quad \text{.......(28)}$$

The number of generation for the population does not play a role - the continuous time scale therefore, an index $n$ is omitted.

The found nonlinear differential equation (28) is the Hardy-Weinberg law for a population with the account of inbreeding.

At F = 0 (inbreeding is absent) the equation (28) it will be transformed to the kind that corresponds to the equilibrium panmictic population:

We use in (28) new independent variable . t = n/Δn√F/(1-F) If to accept [7] Δn = 3/2ln 2

we shall find

$$t = \frac{2\ln 2n}{3} \sqrt{\frac{F}{1-F}} \approx 0.462 n \sqrt{\frac{F}{1-F}}$$

For the new variable $t$ the equation (28) depends only on one constant parameter ε:

$$\frac{d^2 q_f}{dt^2} + \varepsilon \frac{dq_f}{dt} = q_f^2 \quad \text{.......(29)}$$

where it is designated *ε = 3/2 √ (1-F) / F* . At variation of the inbreeding factor 0 ≤ F ≤ 1 the size is ∞ ≥ ε ≥ 0.

The nonlinear differential equation (29) can be solved only numerically.

Initial conditions first of all are necessary for the solution of the differential equation (29): initial frequency $q_{f0}$ of the recessive allele *a* and initial speed of increase in this frequency ($dq_f$/ dt)₀ or ($dq_f$/dn)₀.

Initial frequency $q_{f0}$ depends on time of the beginning of the population analysis. Therefore, without reduction of the analysis generality at t = 0 or n = 0 we shall accept $q_{f0}$ = 0.6

Various variants of numerical calculations under the formula (29) show that the result of calculation at small initial speed of the frequency increase $q_f$ practically does not depend on this speed. Therefore, was accepted. ($dq_f$/

dt)$_0$ = ($dq_f$/dn)$_0$ = 0 The step of the variable $t$ equal $\Delta t$ = 0.001 and numbers of steps 20000 at calculation were used.

On fig. 2 the calculation of the population dynamics of recessive allele $a$ on the basis of the equation (29) at various inbreeding factors $F$ is shown. On the abscissa axis the size n $\approx$ t/0.462$\sqrt{(1-F)/F}$ is plotted.

From graphs it is visible that at presence of inbreeding in the population there is continuous increase of the recessive allele $a$ frequency $q_f$ in which can arise lethal adverse (at association in homozygotes) mutations i.e. there is the negative genetic load in the population. The growth of the recessive allele frequency received in calculation is observed due to growth of the recessive homozygotes $aa$ frequency. Thus the frequency of a heterozygotes $Aa$ at inbreeding falls that leads to the general falling of the intact recessive allele $a$ frequency $q_f$ on all zygotes in a population ($AA$, $Aa$, $aa$) due to adverse mutations only the recessive homozygotes $aa$.

Traditionally [8] the genetic load is estimated on the basis of the analysis of deadborns or children's death rate with the help of two parameters $A$ and $B$ included in the approximate equation for the fraction of the survived zygotes:

S = exp (A+BF)      ……..(30)

Unfortunately the data on experimental sizes of parameters $A$ and $B$ have very much variation of the values. In [1] cited the data on the population of France B = 1.5 - 2.5 and B/A = 15.06 - 21.41.

More detailed data are available on the population of Japan. In different prefectures the ratio B/A = 15.2-(-5). Negative value B/A means that children's death rate in consanguineous marriages was lower than in non consanguineous marriages. Average sizes on the country A = 0.1036, B = 0.67 and B/A = 6.7

It is possible to assume that there is some threshold of a recessive homozygous genes allowable frequency compatible with a survival of individual.

Taking into account values $A$ and $B$, and also inbreeding factor for Japan F = 0.004 [1] under the formula (30) we shall find the fraction of the survived zygotes S = 0.9. Fraction of the perish zygotes there is

1-S = 0.1. The elementary estimation under the formula 1-S = q$^2_f$ = 0.1 gives the fraction of lethal recessive alleles at association theirs in homozygous

$q_f \approx 0.316$ [4]. Hence the fraction of not mutant genes equal to threshold frequency of the recessive alleles is $q_{f\,max}$ = 1-$q_f$=0.684.

On fig. 2 dotted line shows the threshold frequency. For Japan this frequency is attained approximately in 167 generation.

The received result shows that accumulation of the genetic load it is process very slow. Achievement of the threshold frequency for Japan to be carried out only in 167 generations or 4175 years at time of the one generation alternation T = 25 years.

In [1] it is noticed that from the beginning of the population development with inbreeding factor F = 0.003 - 0.005 up to achievement of an equilibrium frequency which is established at action of selection should pass about 4500 years. In calculation the formula of linear increase of recessive alleles frequency was used at mutation speed $\mu$ = 10$^{-5}$. Increase of the mutant alleles number was compensated by selection against recessive homozygotes with parameter of selection 0.5.

Therefore, most likely that limiting values of the genetic load in a human population of many countries especially taking into account small sizes of inbreeding factor (Canada F = 45.10$^{-5}$, France F = 23.10$^{-5}$, Germany F = 19.10$^{-5}$, Italy F = 16.10$^{-5}$ [1] etc.) even at presence of selection till now are not achieved.

In our opinion with help of the offered method it is possible to examine not only harmful for the population genetic load related with inbreeding but also useful and indifferent genetic load (or attributes).

At early stages of mankind development the role of inbreeding was more in connection with smaller number of people and isolation of separate groups. Therefore, accumulation of the genetic load was more intensively (the inbreeding factor was much more). Primeval people after the going out from Africa and moving in northern Europe began to live in conditions of smaller light exposure including in the ultra-violet part of spectrum. In result the skin with originally dark pigmentation started to have white color. It promoted occurrence of vitamin D which deficiency results in

rickets [1]. Process of dark pigmentation reduction of skin can be examined as accumulation of the recessive useful genetic load. Natural selection fixed given phenotypic attribute at people in region.

To the indifferent genetic load in population it is possible to ascribe occurrence such recessive attribute as blond hair, for example, in northern part of Europe where apparently at early stages the inbreeding has been especially widespread. Color of hair does not play a role in survival of population. Fixed of this attribute was carried out by the sexual selection i.e. marriage preferences of people in region.

However it is traditionally accepted to examine accumulation of the genetic load in the population are determined by harmful mutations. Some fears of H. J. Muller are connected to this process about danger of the future biological degeneration of mankind [6].

*Migration of Inbred Populations*

As it was already specified above in primitive populations, for example, leaving of Africa and moving aside northern Europe influence of inbreeding was important. Therefore, we shall consider influence of inbreeding on a migrating population.

The way of the inbreeding account for a motionless population has been analyzed in paragraph 3, equation (28). Adding in the equation (28) for allele frequency $q_f$ motionless inbred populations the item reflecting movement of a population [9] we shall find the general equation for a condition of genome in a moving inbred population:

$$\frac{d^2 q_f}{dn^2} + \frac{3}{2\Delta n}\frac{dq_f}{dn} = D\frac{d^2 q_f}{dX^2} + \frac{F}{\Delta n^2(1-F)}q_f^2 \qquad \ldots\ldots(31)$$

The found equation is the nonlinear differential equation of the second order with square-law nonlinearity. $q_f = f(\varsigma)$ We shall search the solution of the equation (31) as a traveling wave $\varsigma = kX - \omega n$ (where In this case the equation (31) will be transformed to the kind:

$$\frac{d^2 q_f}{d\varsigma^2}\left(\omega^2 - Dk^2\right) - \frac{3}{2\Delta n}\frac{dq_f}{d\varsigma}\omega = \frac{F}{\Delta n^2(1-F)}q_f^2 \qquad \ldots..(32)$$

Let's lower the order of the equation (32) having designated.

$$\frac{dq_f}{d\varsigma} = Z\frac{\frac{3}{2\Delta n}\omega}{\omega^2 - Dk^2}$$

We shall receive Abel's differential equation:

$$Z\left[\frac{dZ}{dq_f} - 1\right] = \frac{\left(\omega^2 - Dk^2\right)}{\left(\frac{3}{2\Delta n}\right)^2\omega^2}\frac{F}{\Delta n^2(1-F)}q_f^2 = \frac{4}{9}\left(1 - \frac{D}{V^2}\right)\frac{F}{(1-F)}q_f^2$$

$$\ldots..(33)$$

Where $\omega/k = V$ there is speed of the population movement. Let's note that the received differential equation does not depend on size $\Delta n$.

The analytical solution of the equation (33) does not exist. However the greatest interest has dependence of the population movement speed $V$ on inbreeding factor $F$. This dependence can be found without the solving of the equation (33).

Let's consider the right part of the equation (33) which defines the inbreeding in a population. The inbreeding factor $F$ is present both at numerator, and in the denominator. Factor $F$ serves in numerator for establishment of all right part existence in the equation (33). At F = 0 the right part of the equation (33) disappears and the population to become panmictic i.e. in this case $q_f = p_{f0} = const$. If F ≠ 0 the recessive allele frequency $q_f$ grows, fig. 2, i.e. inbreeding factor in numerator is responsible for dynamics of the function $q_f$ growth. Therefore, the inbreeding factor in numerator of the equation (33) right part should be kept at any transformations.

Taking into account the given position we shall transform the right part (33) as follows:

$$Z\left[\frac{dZ}{dq_f} - 1\right] = \frac{4}{9}\left(\frac{1}{1-F} - \frac{D}{\left(V\sqrt{1-F}\right)^2}\right)Fq_f^2 \qquad \ldots\ldots.(34)$$

All parameters connected to the wave phase are present only at the second addend in brackets of the right part (34). Therefore, it is possible to conclude that for the account of the inbreeding factor influence on the moving population speed it is necessary to multiply the speed of the population wave in $\sqrt{(1-F)}$.

Using result received in [9] for the genome speed $V$ of moving population and itself population we can transform it to the kind:

$$V = \left( \frac{\sqrt{Dk^2 + \left(\frac{1}{2}\ln 2\right)^2} - \frac{1}{2}\ln 2}{kT} \right)\sqrt{1-F} =$$

$$= \left( \sqrt{\frac{D^*}{T} + \left(\frac{\lambda}{4\pi T}\ln 2\right)^2} - \frac{\lambda}{4\pi T}\ln 2 \right)\sqrt{1-F} \quad ........(35)$$

On fig. 3 the dependence of the moving population speed on inbreeding factor is shown at the following parameters: $\lambda$ = 0.4 km, D* = 10km$^2$/year, T≈25 years.

The received result, fig. 3, can be interpreted as follows. At increase of inbreeding in the moving population its speed of movement decreases. Biologically it can be connected by that in the population with consanguineous mating the individuals become more weakened and during moving by less dynamical.

At F = 1 according to (35) the population stops. The reason of this stop there is fast and full degeneration of the population. If to address to the equation (28) for the motionless population (or (31) at D = 0) at F → 1 the right part of the equation which determines the inbreeding quickly grows aspiring to infinity. This increase can be compensated only to increase in velocity of frequency $q_f$ growth in the left part of the equation (28). But this speed is limited by the biologo-reproductive opportunities of the population. Therefore, in a reality achievement F = 1 does not occur because the population quickly degenerates.

During migration of populations owing to small number of individuals the role of consanguineous mating was great. The found differential equation (31) of the moving inbred populations allows to analyse reduction in speed $V$ of a population movement depending on inbreeding factor.

*Migration of Inbred Populations at Presence of Selection*

The population is in process of migration usually enough long time. During migration of a population there is an alternation of generations and natural selection operates.

Using (31) and result received in [10], it is possible to write down the differential equation migrating inbred populations in which natural selection operates.

$$\frac{d^2 q_f}{dn^2} + \frac{3}{2\Delta n}\frac{dq_f}{dn} = D\frac{d^2 q_f}{dX^2} + \frac{F}{\Delta n^2 (1-F)}q_f^2 - \frac{s}{2\Delta n^2}\left(q_f^2 - q_f^3\right)$$

$$..........(36)$$

The received nonlinear equation is difficult for the analysis. However, taking into account cubic nonlinearity of the differential equation of the second order it is possible to assume presence of solutions as solitary population waves [11]. Apparently, the migrating population represents such solitary wave.

However, the equation (36) allows receive very important analytical result. For this purpose it is used wave substitution ς = 1/Δn$^2$ (kX - ωn) where $k$ there is wave number, ω - dimensionless cyclic frequency of the population waves. In result we shall find:

$$\frac{d^2 q_f}{d\varsigma^2}\left(\omega^2 - Dk^2\right) - \frac{3}{2}\frac{dq_f}{d\varsigma}\omega = \left(\frac{F}{1-F} - \frac{s}{2}\right)q_f^2 + \frac{s}{2}q_f^3$$

$$........(37)$$

Let's find the particular solution of the equation (37). We shall equate the right part of this equation to zero.

$$\left(\frac{F}{1-F} - \frac{s}{2}\right)q_f^2 + \frac{s}{2}q_f^3 = 0$$

In result we have:

$$q_f = \left(1 - \frac{2F}{(1-F)s}\right) = const \quad .......(38)$$

Obviously, value of frequency $q_f$ (or $p_f = 1 - q_f = 2F/(1-F)s$ ) in the formula (38) satisfies to the equations (36) and (37).

The received expression (38) can be interpreted as follows. If in inbred populations which began to migrate there is some frequency of a recessive allele in *X*-chromosomes at women for preservation of this frequency should be interrelation (38) between inbreeding factor *F*, parameter of selection *s* and this frequency $q_f$. Formula (38) is correct and for motionless populations.

We assume the frequency of the recessive allele is equal $q_f$ = 0.6. If to use inbreeding factor F = 19.10$^{-5}$, for example, for Germany [1] that for a migrating population this frequency did not vary the parameter of selection must be s = 9.6.10$^{-4}$ .

Table 1. inbreeding factors for the some consanguineous marriages are shown in addition.

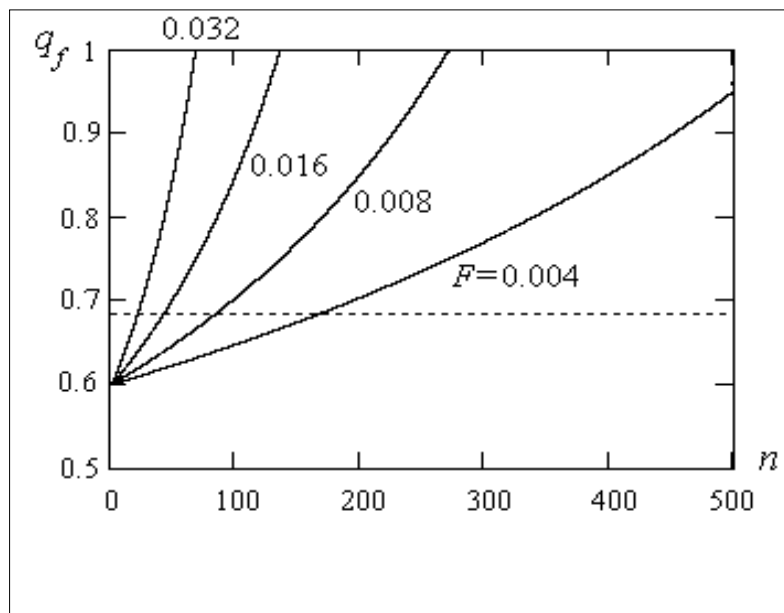| The kind of consanguineous marriage | Inbreeding factor $F_t$ |
|---|---|
| Second cousins | 1/32 |
| Cousin uncle - niece | 1/32 |
| Third cousins | 1/64 |



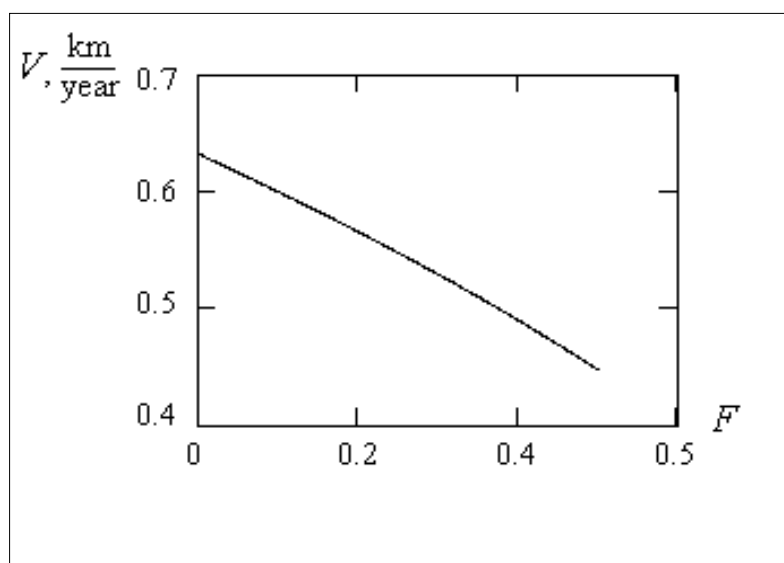Figure 2. Population dynamics of recessive allele $a$ at various inbreeding factors $F$



Figure 3. Dependence of velocity of migrating population on inbreeding factor

The problem of interaction of populations is very important. Obviously, it is connected to a finding of the multiwave solutions of the equation (36) that is extremely complex mathematical problem. Population waves can cooperate with each other in the various ways.

*Population Waves can have Character of a Solitons [11].*

Solitons - it is stable nonlinear solitary waves which at interaction with other local perturbations or with each other show particle like properties. For example, they restore the initial form, i.e. interactions similarly to absolutely elastic particles, for example, to absolutely elastic spheres. As well as spheres the solitons, for example, can be reflected from a barrier. The marked properties of solitons are caused they has the strict balance of the nonlinear processes resulting in increase of a steepness of a wave forward front and dispersive processes lead to destruction a solitary wave, i.e. to reduction of a front steepness. The balance of these processes leads to stability of soliton, i.e. invariability of its form.

We shall note that the third degree of nonlinearity of the differential equation (36) is sufficient for occurrence of a soliton [11]. At interaction of the solitons the population waves can change directions on opposite on type impacting spheres. I.e. populations try to avoid the destruction. If populations belong to poorly cooperating individuals (elephants and birds) waves of populations at interaction pass through each other practically not influencing one on another.

As a whole, the further analysis of interaction of populations has no genetic-mathematical character and passes to area of the social phenomena.

## Conclusion

Inbreeding in a family tree and a population it is possible to take into account with the help of an inbreeding factor which is probability that two alleles at various individuals are identical by origin.

The Hardy – Weinberg's law has a various kind for inbred a family tree and populations. In a family tree the inbreeding factor is included into the discrete equation, and populations in the differential equation. The numerical solution of this differential equation is found.

Found the distribution of genotypes in *X*-chromosomes shows that as well as in autosomes at inbreeding the fraction of homozygotes increased, and the fraction of heterozygotes is reduced. At the greater predisposition to mutations of the recessive homozygotes which frequency increases at inbreeding the accumulation of a genetic load is possible.

It is shown that speed of a migrating population at increase of inbreeding factor is reduced.

Interrelation of the recessive allele frequency at woman for a migrating population with inbreeding factor and standard parameter of selection was found.

## References

1. Vogel F., Motulsky A. Human genetics. Volumes 1, 2, 3, Springer-Verlag, Berlin, 1990. 1068 p.

2. Wright S. Evolution in Mendelian population. Genetics. 19, 1931. Pp. 97-159.

3. Ventcel E.S. Probability theory. Moscow, Science, 1969. 576 p.

4. Ayala F., Kiger, J. Jr. Modern genetics. Volume 3, The Benjamin/Cummings Publishing Company, Inc. California, 1984. 335 p.

5. Li, C.C. First Course in Population Genetics. The Boxwood Press Pacific Grove, California, 1976. 556 p.

6. Muller H.J. Our load of mutation. Am. J. Hum. Genet. 1950. Pp. 111-176.

7. Volobuev A.N. Genetic-mathematical modelling of mutational processes in a population. Journal of Genetic Engineering. Vol. 1. Issue 1. 2019. Pp. 1-9.

8. Morton N.E., Crow J.F., Muller H.J. "An estimate of the mutational damage in man from data on consanguineous marriages. Proc. Natl. Acad. Sci. USA, 42, 1956. Pp. 855-863.

9. Volobuev A.N. Migration of a Population. Journal of Genetic Engineering. Vol. 1. Issue 1. 2019-20. Pp. 15-20.

10. Volobuev A.N. Natural Selection in a Population is a Problem of Nonlinear Genetics. Journal of Genetic Engineering. Vol. 1. Issue 1. 2019. Pp. 10-14.

11. Dodd R., Eilbek G., Ghibbon G., Morris G. Solitons and Nonlinear Wave Equations. Moscow, World, 1988. 696 p.