Investigations into genetic variability in Holstein horse breed using pedigree data

L. Roos [a,*], D. Hinrichs [a], T. Nissen [b], J. Krieter [a]

[a] Institute of Animal Breeding and Husbandry, Christian-Albrechts-University, Hermann-Rodewald-Straße 6, D-24118 Kiel, Germany
[b] Verband der Züchter des Holsteiner Pferdes e.V., Abteilung Zucht, D-24106 Kiel, Germany

ARTICLE INFO

ABSTRACT

A pedigree data set including 129,923 Holstein warmblood horses was analyzed to determine genetic variability, coefficients of inbreeding, the age of inbreeding and the genetic contributions of founder animals and foreign breeds. The reference population contained all horses which had been born between 1990 and 2010. The average Pedigree Completeness Index (PEC) for the reference population was determined as 0.88 and the average complete generation equivalent (GE) was computed at 5.62. The mean coefficient of inbreeding for the reference population (inbred and non-inbred horses) was 2.27%. Most of the inbreeding was defined as "new" inbreeding, which had evolved during recent generations. The effective population size and the effective number of founders were calculated to be 55.31 and 50.2 effective individuals respectively. The most influential foreign breed was the English Thoroughbred with a contribution of 25.98%, followed by Anglo Normans (16.38%) and Anglo Arabians (3.27%). At 2.75%, Hanoverian warmblood horses were determined to be the most important German horse breed. The stallions Cor de la bryere, Ladykiller xx and Cottage son xx were found to be the most important male ancestors. The mare Warthburg was defined as the most affecting female. It was possible to detect the occurrence of the loss of genetic diversity within the Holstein horse breed, related to unequal founder contributions caused by the intensive use of particular sire lines. However, a slight increase in the effective population size and a stagnation of inbreeding during the last generation might show the impact of more open access given to foreign stallions in the recent past.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Based on its success in international riding competitions, the Holstein horse breed has become one of the most popular breeds, especially in show jumping.

The official breeding association was founded in 1935 and today's complete breeding population includes 7693 registered mares and 225 licensed stallions within twelve breeding districts. In the 19th century, the Holstein horse breed was influenced by the Yorkshire Coach horse and by Thoroughbreds (Löwe, 1988).

Due to rising mechanization, the breeding goal has shifted from medium- weight draft or riding horse for agricultural and cavalry use (before 1950) to a large framed, athletic and expressive sport horse with a preferential aptitude for show jumping.

This process of refinement has been driven by an increased use of English Thoroughbred and Anglo-Norman stallions.

Together with the Trakehner Horse breed, the Holstein horse is the unique German sport horse breed working with closed studbooks.

---

* Corresponding author.
E-mail address: lroos@tierzucht.uni-kiel.de (L. Roos).

Accordingly, the studbook for mares is strictly closed and the use of stallions from foreign breeds in terms of breeding trails is minimized. Due to the increased use of artificial insemination and against the background of the intensive use of certain sires, an increase in terms of the rate of inbreeding and the contributions of fewer ancestors is probable. A refreshment of previous knowledge is needed concerning the composition of the Holstein gene pool.

There has not been any investigation concerning genetic composition of Holstein horse breed. However, Hamann and Distl (2008) and Teegen et al. (2009) did some research on the population structure of the Hanoverian and Trakehner breed respectively.

Therefore, the aim of this study was to point out the updated levels of inbreeding, the proportion of foreign blood and to specify the genetic contributions of outstanding founders to the current structure of breeding stock. Additionally, the study applied some alternative concepts regarding the evolvement of inbreeding.

2. Material and methods

2.1. Pedigree data

Pedigree data used for this study was provided by the Association of Holstein Horse Breeders (Kiel/Germany) with support of the "Landeskontrollverband Schleswig-Holstein" which is assigned to administer the pedigree data base. In 2010 the whole pedigree data set contained 131,272 animals.

After revision and verification, a data-set of 129,923 animals with 55,796 males and 74,127 females was included in the analysis. Approximately 1% of undetermined data was excluded.

The reference population applied in this study included all horses born between 1990 and 2010 ($n=78,677$, with known parents). The first recorded ancestor was born in the year 1869. Choosing a reference population consisting of all animals born in a period of two generation intervals, the intension was to depict inbreeding situation for the actual breeding stock completely as possible. Even if some of the animals died or probably not used anymore, evaluating a shorter period of time would exclude reproductive individuals and their progeny from the analysis (e.g. competing mares, resuming their breeding use after several years).

2.2. Data analysis

The following parameters of population structure were exploited in the analysis for all horses within the reference population, based on the whole pedigree data-set: The average coefficient of inbreeding, the effective number of founders, the effective number of ancestors and the effective number of founder genomes. Additionally, the generation intervals were determined for the four pathways sire to sire, sire to dam, dam to sire and dam to dam. Therefore, the average age of the parents at the time of birth of their first reproductive offspring was used.

To identify the amount of pedigree completeness and to quantify the possibility to ascertain inbreeding, the pedigree completeness index (PEC) (MacCluer et al., 1983) was computed as follows:

$$PEC_{\mathrm{animal}} = \frac{4C_{sire}C_{dam}}{C_{sire}C_{dam}},$$

where C_{sire} and C_{dam} are the amount of pedigree information contributed by the two parental lines.

To specify the number of entire generations, the complete generation equivalents (GE) were calculated thus for each individual j:

$$G_j = \sum n_i/2^g,$$

where n_i is the number of known ancestors in generation i and g is the number of known generations for individual j.

In case of individuals with two unknown parents, animals were considered as nonrelated founders. The contribution of founders could have been different, because some of them had been used with a greater intensity than others.

Due to this fact, the amount of founders provided less information about the genetic diversity of a population.

To overcome this problem, Lacy (1989) introduced the effective number of founders (f_e), defined as the number of equally contributing founders expected to produce the same genetic diversity as the population under study. The more equal the contributions of the founders the greater is the effective number of founders.

In case of an equal contribution of all founders, the effective and actual number of founders is the same.

Boichard et al. (1997) developed another characteristic factor to clarify genetic diversity with regard to the loss of allelic diversity. The so called effective number of ancestors (f_a) also embraced for the contributions of all ancestors and was defined as the minimum number of ancestors explaining the complete genetic diversity of the current population. The computation of this parameter was predicated on the marginal contributions of the 1000 most important ancestors.

Bottlenecks or a frequent use of special sires and their offspring are known as reasons for the loss of allelic variability. To identify ancestors which influenced the genetic composition of the population more than others, it was necessary to look at the difference between the number of effective founders and effective ancestors.

A larger amount of effective founders in proportion to the number of effective ancestors referred to ancestors which assisted the formation of the population to a greater extent than others (Boichard et al., 1997). The underlying fact was that the contributions of ancestors did not matter for the generations when they are marginalized.

The effective number of founder genomes (f_g) was defined by Lacy (1989) as "that number of equally contributing founders with no random loss of founder alleles in descendants that would be expected to produce the same genetic diversity as in the population under study". The Gene drop procedure introduced by Boichard et al. (1997) was used to compute the number of effective founder genomes. Because the use of breeding animals is