Runs of homozygosity for autozygosity estimation and genomic analysis in production animals

Arnaldo Basso Rebelato[1] and Alexandre Rodrigues Caetano[2]

[1]Universidade de Brasília, Campus Universitário Darcy Ribeiro, CEP 70910-900 Brasília, DF, Brazil. E-mail: arnaldo.br@me.com
[2]Embrapa Recursos Genéticos e Biotecnologia, Parque Estação Biológica, Avenida W5 Norte (Final), CEP 70770-917 Brasília, DF, Brazil. E-mail: alexandre.caetano@embrapa.br

Abstract – Runs of homozygosity (ROHs) are long stretches of homozygous genomic segments, identifiable by molecular markers, which can provide genomic information for accurate estimates to characterize populations, determine evolutionary history and demographic information, estimate levels of consanguinity, and identify selection signatures in production animals. This review paper aims to perform a survey of the works on the efficiency of ROHs for these purposes. Factors such as genetic drift, natural or artificial selection, founder effect, and effective population size directly influence the size and distribution of ROHs along the genome. Individually, genome estimates of consanguinity based on ROHs can be obtained using the F_{ROH} index, which is generally considered more accurate than indexes based on other types of genomic or genealogical information. High frequencies of specific ROHs in a population can be used to identify selection signatures. The results of recent studies with ROHs in domestic animals have shown the efficiency of their use to characterize herds in a reliable and accessible way, using genomic information.

Index terms: autozygosity, F_{ROH}, genotyping, selection signatures, single nucleotide polymorphism, SNPs.

Introduction

The domestication process of production animals, followed, in the last decades, by intense directional selection using quantitative methods, has resulted in significant genetic increments in adaptation, type, and production (Randhawa et al., 2016). Consequently, drastic reductions were observed in the effective population size (Falconer & Mackay, 1996), as well as production losses due to high consanguinity rates (Reverter et al., 2017) in breeds under intense selection (Scraggs et al., 2014; Zavarez et al., 2015).

The balance between high rates of genetic gain and of diversity loss require breeding programs to take into account and control endogamous breeding rates in the studied populations (Peripolli et al., 2017). Traditionally, these processes use pedigree information, but random errors in the steps necessary to collect, takes notes of, register, and store information on genetic relationships may have serious negative consequences and cause unintended increases in autozygosity levels (Curik, et al., 2014; Hudson et al., 2015; Zavarez et al., 2015). Therefore, the use of genomic data to complement or correct existing genealogical data may positively

impact genetic gains obtained in the long run (Hudson et al., 2015; Marras et al., 2015; Zhang et al., 2015b).

The first studies with molecular markers to aid in animal breeding date from the end of the 1980s, and the obtained results, based on a few markers and high-cost techniques, were inconsistent regarding the efficiency and viability of using genomic data for animal production (Caetano, 2009). However, the association of technologies to fabricate DNA microarrays with molecular methods developed to genotype single nucleotide polymorphisms (SNPs) lead to the development of panels with thousands of markers, allowing an extensive evaluation of the genome in a fraction of the time and cost required by previous methodologies (Caetano, 2009; Silva et al., 2015).

Defined as long and uninterrupted stretches of the genome with genotypes in homozygosis (Lencz et al., 2007), the runs of homozygosity (ROHs) may be identified through the analysis of high-density SNP panels, preferentially with more than 50.000 SNPs. The directed analysis of data generated with these panels has allowed studies to identify and characterize ROHs in different species (Curik et al., 2014). In addition,

secondary analyses of ROHs may be used to: identify and monitor the consanguinity rate in production animals; map and identify recessive deleterious alleles; characterize the population demography, structure, and history; and estimate individual and population genomic relationships and autozygosity (Saura et al., 2015; Peripolli et al., 2017).

Currently, SNP genotyping using DNA chips is the most consistent and low-cost solution to generate high-definition data (Silva et al., 2015). This methodology is being used in experiments to identify and use of ROHs in production animal species, such as bovine (Zavarez et al., 2015; Zhang et al., 2015a; Mastrangelo et al., 2016; Szmatoła et al., 2016; Reverter et al., 2017), ovine (Al-Mamun et al., 2015; Beynon et al., 2015), equine (Metzger et al., 2015), swine (Gomez-Raya et al., 2015; Saura et al., 2015), poultry (Wolc et al., 2015), feline (Bertolini et al., 2016), and bees (Fuller et al., 2015). Table 1 presents a summary of the objectives and conclusions of recent studies related to different uses of ROHs in production animals.

This review paper presents information on ROHs, the methods used for their identification, and their different applications in studies that may positively impact animal production, mainly regarding genomic estimates for the characterization of a population in order to monitor autozygosity and its implications.

Runs of homozygosity

The term ROHs was used for the first time to denominate stretches with 100 or more consecutive SNPs, with absence of heterozygotes or missed calls, in one chromosome (Lencz et al., 2007). ROHs were initially identified in human chromosomes (Broman & Weber, 1999), because, at the time, this was the only species with a wide genome coverage, a requirement for the correct identification of ROHs (Purfield et al., 2012). Studies on the size and distribution of ROHs in other animal species were only carried out approximately eight years later (Gibson et al., 2006), when the first high-density SNP panels for production animals were developed.

The formation process of ROHs is shown in Figure 1, where individual A represents the common ancestor of parents D and E of individual F. The ROH fragment of individual F, as well as the homologous chromosome of the common ancestor A, which originated this ROH, is represented by the green color. The other colors indicate chromosome fragments that did not originate from the homologous chromosome of individual A (Gomez-Raya et al., 2015).

The correct identification of ROHs depends on the control of several factors, such as genotyping quality, minimum size of ROHs, and number of allowed heterozygotes, which may compromise estimates due to eventual genotyping errors (Ferenčakovićet al., 2013b). The type of chip used to obtain data also influences the identification of ROHs, since a wide genome coverage allows the identification of a greater number of runs. It should be noted that chips with densities greater than 50,000 SNPs are necessary to detect with precision ROHs shorter than 5.0 Mb (Purfield et al., 2012; Zhang et al., 2015a).

High-density SNP panels for production animals were only developed starting in 2008 (Caetano, 2009). Currently, there are high-density genotyping chips for several production species, which have allowed for studies to identify ROHs for these species (Curik et al., 2014). The obtained results have shown the potential of this approach to identify genomic regions of interest. The characterization of ROHs in different populations, breeds or lines of a species is important to obtain information on evolutionary history (Metzger et al., 2015; Sorbolini et al., 2015; Zavarez et al., 2015), demographics (Bosse et al., 2012), or related to consanguinity in a population (Marras et al., 2015).

Software used to identify ROHs

Currently, the software most used for the identification of ROHs are Plink v1.07 (Purcell et al., 2007) and SNP & Variation Suite (SVS) of Golden Helix (Curik et al., 2014). However, others cited in the literature, such as Beagle and Germline, were also used for this purpose (Howrigan et al., 2011).

The Plink uses a sliding window to identify both ROHs and consecutive stretches that contain a minimum number of homozygous SNPs, at a minimum pre-specified distance. With this method, the software carries out basic detections of homozygous stretches identified by the sliding window, and the user only needs to define the parameter "minimum size" of segments to be identified (Curik et al., 2014).

Table 1. Recent studies with runs of homozygosity (ROHs) for the detection of autozygosity, identification of selection signatures, and characterization of populations of production animals.

| Author	Year	SNP panel	Species	Objective of using ROHs	Conclusions
Gomez-Raya et al.	2015	35K	Swine	Obtain new autozygosity coefficients based on the size and distribution of ROHs and compare them with ROH frequency (F_{ROH}) and pedigree information (F_{PED})	New coefficients of autozygosity provided additional information on recent consanguinity
Howard et al.	2015	54K	Bovine	Identify regions of the genome that suffered diversity losses due to consanguinity and determine the relationship between additive effects and ROHs	Several regions are associated with ROH <4 Mb in the studied population and there is correlation between genetic additive effect and ROH <4 Mb, which is an indicative of the influence of geographical location on homozygosity
Kim et al.	2015b	54K	Bovine	Identify the association between ROHs and consanguinity coefficients (F), as well as analyze ROHs identified in populations of dairy cattle	ROHs reflect homozygosis strongly affected by recent artificial selection
Kim et al.	2015a	54K	Bovine	Identify selection signatures in Jersey cattle under selection since 1960 and compare them by the method for the identification of selection signatures via extended haplotypes	The analysis of the estimated ROHs allows to efficiently identify selection signatures
Marras et al.	2015	54K	Bovine	Compare F_{ROH} with the coefficient of genomic inbreeding matrices (F_{GRM}) and F_{PED}, in five different bovine breeds	ROH is a powerful tool to estimate the coefficient of consanguinity
Metzger et al.	2015	50K	Equine	Identify selection signatures in improved and non-improved horses	The size and frequency of ROHs vary according to population diversity and selection pressure
Saura et al.	2015	60K	Swine	Detect genomic regions responsible for the loss of genetic variability in reproductive characteristics of improved swine	Genomic estimates based on the presence or absence of ROHs are a viable alternative for the detection of variability losses due to consanguinity
Zavarez et al.	2015	777K	Bovine	Characterize levels of autozygosity, based on ROHs, in a population of Nellore cattle	The analysis of ROHs allows characterizing herds according to their endogamy levels, besides identifying genomic regions with possible selection signatures for the breed
Zhang et al.	2015a	54K	Bovine	Compare different estimators of coefficient of consanguinity, calculated from pedigree information, with data from 50K SNP chip and from complete sequencing and estimate based on ROH	F_{ROH} reflects direct homozygosity levels because it is not affected by allele frequencies, whereas F_{PED} is limited due to its dependence on correct pedigree information
Zhang et al.	2015b	Complete genome	Bovine	Study ROH patterns to verify the effect of selection pressure and demography in the increase in the frequency of deleterious and non-deleterious sequences within ROHs	ROH is efficient to detect functional variants in bovine populations and contributes for a better understanding of endogamy and selection effects in these populations
Mastrangelo et al.	2016	54K	Bovine	Quantify consanguinity from information on ROHs estimated in three different bovine breeds of economic importance	Consanguinity values in three bovine breeds, obtained via the detection and distribution of ROHs, evidenced the need of implementing conservation programs aiming to control consanguinity levels
Gurgul et al.	2016	54K	Bovine	Compare F_{ROH}, F_{GRM}, and F_{PED}, test the correlation between these coefficients, and analyze if ROH or GRM may be interesting to estimate recent consanguinity in Holstein cattle	Correlations between F_{ROH} and F_{PED} tend to increase with the increase of pedigree information and indicate that F_{ROH} is the most recommended to estimate recent consanguinity
Szmatoła et al.	2016	54K	Bovine	Characterize ROHs in four different bovine breeds and identify genomic regions with a high frequency of ROHs and that have been subjected to directional selection	The size and distribution of ROHs vary according to the studied breed, with different ROH patterns for native, conserved or commercial breeds, and to possible selection signatures
Reverter et al.	2017	729K 71K 19K	Bovine	Compare F_{GRM}, the coefficient of homozygosity (F_{HOM}), and F_{ROH} as estimators of consanguinity using data from three different panel densities in zebu cattle	In heterogeneous populations (crossed animals), endogamy measures that do not depend on allele frequency, such as F_{HOM} and F_{ROH}, are better recommended to estimate endogamy

Plink is available for free download on the Linux, MS-DOS, Apple Mac, and C/C++ source platforms (Plink…, 2017). Being free is the great advantage of this software compared with SVS of Golden Helix, whose higher cost, over U$ 1,000 per year (Golden Helix, 2017), is justified by its great capacity for data management, friendly design, quality of the produced material, and guaranteed support .

The SVS software does not use sliding windows, but considers all SNP in homozygosis as a possible starting point of a new ROH. Each SNP is classified as "in homozygosis", "in heterozygosis" or "missed calls", and provides a cluster of homozygous stretches with a number of SNPs in homozygosis greater than the one specified for each chromosome and individual. Then, a second algorithm groups all calculated stretches in clusters and provides a list with the minimum number of individuals that have these stretches in common. This more modern and complex method allows the user to define groups of parameters, such as minimum size

of ROHs in base pairs and SNP numbers, minimum density, maximum gap, and maximum number of allowed heterozygotes and of missed calls (Curik et al., 2014).

Factors affecting the formation of ROH patterns

The processes of natural and artificial selection may alter genotypic patterns and produce contrasting patterns in populations subjected to distinct selective pressures (Sorbolini et al., 2015). The selection of a small number of superior animals tends to reduce the observed phenotypic variability, besides leading to genome remodeling in production animals, generating ROH patterns (Kim et al., 2013), due to increased homozygosity in genomic regions close to the locus that controls traits of interest (Zhang et al., 2015b). Genomic regions that contain loci subjected to artificial
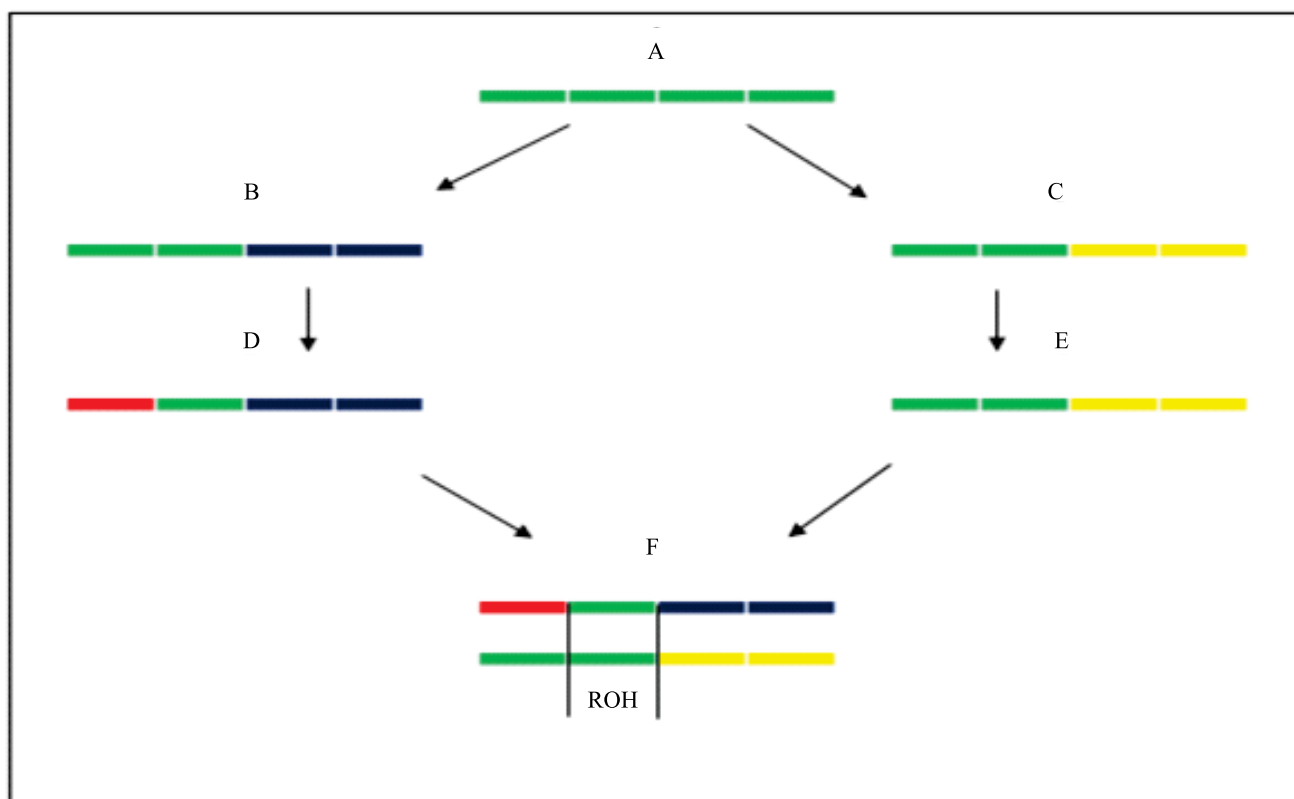
Figure 1. Representation of the formation of runs of homozygosity (ROHs). Individual F presents a ROH (in green) formed by the pairing of stretches in homozygosis of the homologous chromosome of a common ancestor A.

selection, in general, present, a greater concentration of ROHs (Metzger et al., 2015).

Demographic factors that lead to genetic drift and to natural or artificial selection pressures may also cause genomic modifications in a species (Ramey et al., 2013). The increase in selection intensity in breeding programs observed in the last decades, together with the use of a small number of animals as breeding stock, has contributed to reduce the effective size of populations of production animals. This contributes to an increase in consanguinity and genetic drift, as well as to a reduction in genetic variability (Peripolli et al., 2017), creating a tendency of long ROHs in the populations of animals subjected to these conditions.

By definition, ROHs are long and uninterrupted stretches of genotypes in homozygosis, as previously mentioned. These stretches are generated by inbreeding events (Ferenčaković et al., 2013a) and, therefore, their size varies according to the number of generations during which inbreeding occurred. Consequently, the size of the ROHs in a herd tends to decrease with each generation. Long stretches in homozygosis, i.e., long ROHs, indicate high consanguinity between individuals (Curik et al., 2014) due to recent inbreeding in the population.

The differences between patterns of ROHs suggest that artificial selection modifies autozygosity in the genome (Metzger et al., 2015; Szmatoła et al., 2016; Peripolli et al., 2017). Selection and/or drift events result in the formation of long ROHs (Pemberton et al., 2012; Howard et al., 2015), which, subsequently, suffer recombination and mutation effects, causing inherited ROHs to decrease in size in each successive generation (Curik et al., 2014).

Characterization of populations through ROHs

The length and frequency of ROHs are used to provide information about the ascendency of an individual or about the structure and history of the population of origin (Howrigan et al., 2011; Purfield et al., 2012). The distribution of ROHs along the genome is not random or uniform, but strongly dependent on local recombination and mutation rates (Bosse et al., 2012), as well as on other evolutionary forces (Ramey et al., 2013). Therefore, the formation and distribution of ROHs throughout the genome is a

result of the combination of genomic variables, such as recombination rate, and of signs of recent directional selection (Pemberton et al., 2012).

In production animals, ROHs longer than those found in human populations are expected due to the processes of artificial selection and to the reduced effective population size (Curik et al., 2014). Breeds and specialized lines of production animals are usually subjected to intense selection of allele clusters, which positively affects production, reproductive, and racial pattern characteristics. Therefore, it is common to observe high endogamy rates and, consequently, ROHs abundantly distributed along the genome and present in high frequencies in these populations (Zavarez et al., 2015; Zhang et al., 2015a).

The analysis of the size, position, and frequency of ROHs throughout the genome may provide information about genomic characteristics, recombination rates, and direction of selection, besides evidencing the relationship between distinct populations (Bosse et al., 2012; Metzger et al., 2015). Long ROHs usually indicate intensive selection pressures (Metzger et al., 2015) and recent consanguinity events (Al-Mamun et al., 2015), whereas short ROHs suggest genetic diversity loss due to bottleneck or founder effects (Al-Mamun et al., 2015) or to past consanguinity events in the population (Howrigan et al., 2011).

Metzger et al. (2015), while studying ROH patterns in different horse populations, observed variations in the number and frequency of runs between purebred and and crossbred animals. According to these authors, there may be a relationship between specific ROH patterns and the history of natural or artificial selection pressures. Saura et al. (2015) reported losses in genetic variability due to selective breeding by identifying the presence or not of ROHs in the genome of improved swine. Szmatoła et al. (2016) also identified distinct ROH patterns in three bovine populations: native, subjected to conservation processes, and commercial. The obtained results are indicative that, once the ROH pattern of the species, breeds, or populations is known, it is possible to determine the evolutionary history or the demographic information that characterize them.

Estimation of consanguinity through the analysis of the frequency of ROHs (F_{ROH})

Autozygocity occurs when mated individuals, with common ancestors, transmit chromosomes identical by descent (IBD) to their progeny. The evaluation of these segments allows calculating the inbreeding coefficient (F) used to estimate the level of autozygosity in an individual with one or more common ancestors. From this coefficient, it is possible to determine the probability of the alleles in a random region of the genome being IBD (Wright, 1922).

Traditionally, pedigree data are used to compose the relationship matrix "A", constructed with the expected proportion of IBD loci (VanRaden, 1992), calculated by the expression:

$$A_{ii} = \sum L^2_{ij}D_{ij}$$

where A_{ii} is the ith element of the diagonal of matrix A, which is equivalent to the inbreeding coefficient of the ith animal plus 1.

Genomic information may be used to estimate the autozygosity of a population (McQuillan et al., 2008) in a fast, reliable, and low-cost manner (Silva et al., 2015). In this case, the inbreeding matrix A, based on pedigree information, is substituted by the inbreeding genomic matrix "G", constructed from information from marker panels, usually high density ones. The inbreeding coefficients are obtained using the proportion of loci identical by state and carry more information than the traditional coefficient (Pértile et al., 2016). Four inbreeding coefficients may be estimated from genomic information: F_{UNI}, of united gametes; F_{HOM}, of homozygosity; F_{GRM}, of genomic inbreeding matrices; and F_{ROH}, of RHOs (Zhang et al., 2015a), described subsequently.

According to Wright (1922), F_{UNI} may be estimated from the correlation between united gametes, by the formula (Yang et al., 2011):

$$F_{UNI} = [X^2_i - (1 + 2p_i) x_i + 2p^2_i] / 2p_i (1 - p_i)$$

where p_i is the observed fraction of the first allele on locus i, and x_i is the number of copies of the reference allele.

F_{HOM} is estimated based on excessive homozygosity according to Wright (1922), and is obtained by the following formula (Yang et al., 2011):

$$F_{HOM} = (HOM_O - HOM_E) /(1 - HOM_E)$$
$$= 1 - x_i (2 - x_i) /2p_i (1 - p_i)$$

where HOM_O and HOM_E are the number of observed and expected genotypes in homozygosis, respectively; p_i is the observed fraction of the first allele on locus i; and x_i is the number of copies of the reference allele.

F_{GRM} is an estimate of the genetic relationship of an individual with itself, obtained from the main diagonal of the genomic relationship matrix (GRM), using genotyping data from the high-density SNP panels (VanRaden, 2008), according to the expression:

$$F_{GRM} = [x_i - E(x_i)]^2 /[2p_i (1 - p_i) - 1]$$
$$= (x_i - 2\hat{p}_i)^2 /[2p_i (1 - p_i) -1]$$

where p_i is the observed fraction of the first allele on locus i, and x_i is the number of copies of the reference allele.

F_{ROH} is calculated from the addition of estimated ROHs, separated according to minimum run sizes. This coefficient may be defined as the proportion of the autosomal genome in ROHs in relation to the autosomal genome covered by SNPs (McQuillan et al., 2008), as:

$$F_{ROH} = \sum L_{ROH} / L_{auto}$$

where $\sum L_{ROH}$ represents the total ROHs above a minimum specified size (L_{ROH}), identified in an individual; and L_{auto} is the total size of the autosomal genome covered by SNPs.

The autozygosity estimates obtained from the inbreeding coefficient (F), calculated using genomic information, are generally more precise than the one obtained from pedigree information (F_{PED}), since the latter does not take into account inbreeding in the founders of the herd, for which there is no pedigree information (in this case, F=0 is assumed). F_{PED} also does not consider sampling effects due to selection (Curik et al., 2014; Zhang et al., 2015a). The first studies with domestic animals in which F_{ROH} was used focused on comparing F_{ROH}, an estimate based on the length of ROHs, with F_{PED}, based on pedigree information (Curik et al., 2014; Sölkner et al., 2010). In this comparison, F_{ROH} was shown to be more efficient than F_{PED} and was recommended as an interesting alternative to correct pedigree errors (Hudson et al., 2015).

The first research reports using FROH estimates for domestic animals were made after the 2010s. Ferenčaković et al. (2011), considering the good results obtained in studies with ROHs in humans, used data from genotyped bovine and their pedigree to establish a correlation between the estimates of F_{ROH} and F_{PED}. The authors found a high correlation (0.68) between both estimates when using complete pedigree information ($F_{PED}T$). Positive correlations between these estimates were also reported by Marras et al. (2015), who found values between 0.66 and 0.70 according to breed, and by Mastrangelo et al. (2016), who obtained even more

expressive correlations between 0.83 and 0.95 using different dairy cattle breeds.

Kinship estimates based on ROHs (F_{ROH}) are usually superior to those identified by pedigree. Ferenčaković et al. (2011) found that the advantage of the index originating from ROHs, when calculated for segments >1 Mb ($F_{ROH} > 1$ Mb), is identifying events of past consanguinity and that were not identified by the F_{PED} estimates. Scraggs et al. (2014), while studying purebred cattle, observed a superior mean for the F_{ROH} genomic estimate than for F_{PED}, corroborating the assumption that F_{ROH} provides additional information about recent consanguinity, compared with F_{PED} (Gomez-Raya et al., 2015).

Considering different estimates obtained from genomic information (F_{UNI}, F_{HOM}, F_{GRM}, and F_{ROH}), individual consanguinity based on ROHs (F_{ROH}) is the most precise (Marras et al., 2015; Zhang et al., 2015b; Gurgul et al., 2016). This is because it is a direct homozygosity measure, calculated from molecular information, and is less susceptible to selection effects and errors caused by sample variations in the gamete-generating phase (Marras et al., 2015), which is not observed in other genomic estimates, influenced by allele frequency. Therefore, the F_{ROH} index is a more precise estimate than other consanguinity estimates obtained from pedigree information or even than other genomic estimates; therefore, it is a viable option for the correction of pedigree errors (Ferenčaković et al., 2011; Marras et al., 2015), and, according to Zhang et al. (2015a), is the most recommended to determine IBD.

ROHs and selection signatures

Selection signatures are a result of genotypic alterations in populations subjected to some form of selection pressure (Ramey et al., 2013; Sorbolini et al., 2015), and are characterized by an increase in allele frequency in one or more genes, or gene clusters, involved in the processes of adaption of a population to specific conditions – such as resistance to diseases, tolerance to cold/heat, or maternal ability – or of improvement for specific purposes – such as meat, milk, and wool production, among others. Therefore, methods that allow identifying these signatures may lead to the identification of the genes involved in the processes related to the productivity of production animals.

ROHs usually cover genomic regions large enough to contain genes or gene clusters, which may be under selection for generations. Therefore, the identification of ROHs may aid in visualizing and identifying haplotype patterns characteristic of breeds or species, fixed due to selection pressures. Starting in 2007, the first reports for the identification of common ROHs in humans affected by Alzheimer and schizophrenia were carried out, aiming to identify genes associated to the development of these diseases (Lencz et al., 2007; Nalls et al., 2009). In the following years, studies with production animals were performed, also identifying ROHs but for the identification of genes or gene clusters related to the productivity of the population and to the characterization of the animal breed or line (Metzger et al., 2015; Sorbolini et al., 2015; Zavarez et al., 2015).

More recently, research activities to identify selection signatures, based on the analysis of the frequency of ROHs, have been carried out for several species of production animals. Fuller et al. (2015) identified ROHs related to adaptive characteristics of commercial bee species, observing the effect of environmental temperature on honey production. In bovine, O'Brien et al. (2014), Somavilla et al. (2014), and Zavarez et al. (2015) identified ROHs associated to the adaptive potential and reproductive and productive characteristics of zebu breeds, whereas Kim et al. (2015a) identified more than 15 regions related to milk production in Jersey herds.

Concluding remarks

The identification and characterization of ROHs may aid in identifying genomic regions that affect traits of interest for the productivity of commercial herds of production animals. Due to the dependency on high-density SNP panels for the correct identification of ROHs, studies with production animals are recent, starting in 2008. However, with the development of DNA chips specific for production species, researches with ROHs have grown exponentially. ROHs may be used for the genomic characterization of herds and have been shown to be efficient for obtaining inbreeding estimates via the F_{ROH} index or for the identification of patterns characteristic of genomically studied breeds or species, known as selection signatures. The

potential of ROHs for aiding conventional production techniques is immense and, with the exponential increase of available genomic data due to new or developing genotyping technology, the tendency is that ROHs will also have other applications.

References

AL-MAMUN, H.A.; CLARK, S.A.; KWAN, P.; GONDRO, C. Genome-wide linkage disequilibrium and genetic diversity in five populations of Australian domestic sheep. **Genetics Selection Evolution**, v.47, art.90, 2015. DOI: 10.1186/s12711-015-0169-6.

BERTOLINI, F.; GANDOLFI, B.; KIM, E.S.; HAASE, B.; LYONS, L.A.; ROTHSCHILD, M.F. Evidence of selection signatures that shape the Persian cat breed. **Mammalian Genome**, v.27, p.144-155, 2016. DOI : 10.1007/s00335-016-9623-1.

BEYNON, S.E.; SLAVOV, G.T.; FARRÉ, M.; SUNDUIMIJID, B.; WADDAMS, K.; DAVIES, B.; HARESIGN, W.; KIJAS, J.; MACLEOD, I.M.; NEWBOLD, C.J.; DAVIES, L.; LARKIN, D.M. Population structure and history of the Welsh sheep breeds determined by whole genome genotyping. **BMC Genetics**, v.16, art.65, 2015. DOI: 10.1186/s12863-015-0216-x.

BOSSE, M.; MEGENS, H.-J.; MADSEN, O.; PAUDEL, Y.; FRANTZ, L.A.F.; SCHOOK, L.B.; CROOIJMANS, R.P.M.A; GROENEN, M.A.M. Regions of homozygosity in the porcine genome: consequence of demography and the recombination landscape. **PLoS Genetics**, v.8, e1003100, 2012. DOI: 10.1371/journal.pgen.1003100.

BROMAN, K.W.; WEBER, J.L. Long homozygous chromosomal segments in reference families from the Centre d'Étude du Polymorphisme Humain. **American Journal of Human Genetics**, v.65, p.1493-500, 1999. DOI: 10.1086/302661.

CAETANO, A.R. Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. **Revista Brasileira de Zootecnia**, v.38, p.64-71, 2009. Número Especial. DOI: 10.1590/S1516-35982009001300008.

CURIK, I.; FERENČAKOVIĆ, M.; SÖLKNER, J. Inbreeding and runs of homozygosity: a possible solution to an old problem. **Livestock Science**, v.166, p.26-34, 2014. DOI: 10.1016/j.livsci.2014.05.034.

FALCONER, D.S.; MACKAY, T.F.C. **Introduction to quantitative genetics**. 4th ed. Essex: Longman, 1996. 464p.

FERENČAKOVIĆ, M.; HAMZIC, E.; GREDLER, B.; CURIK, I.; SÖLKNER, J. Runs of homozygosity reveal genome-wide autozygosity in the Austrian Fleckvieh cattle. **Agriculturae Conspectus Scientificus**, v.76, p.325-328, 2011.

FERENČAKOVIĆ, M.; HAMZIĆ, E.; GREDLER, B.; SOLBERG, T.R.; KLEMETSDAL, G.; CURIK, I.; SÖLKNER, J. Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. **Journal of Animal Breeding and Genetics**, v.130, p.286-293, 2013a. DOI: 10.1111/jbg.12012.

FERENČAKOVIĆ, M.; SÖLKNER, J.; CURIK, I. Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. **Genetics Selection Evolution**, v.45, art.42, 2013b. DOI: 10.1186/1297-9686-45-42.

FULLER, Z.L.; NIÑO, E.L.; PATCH, H.M.; BEDOYA-REINA, O.C.; BAUMGARTEN, T.; MULI, E.; MUMOKI, F.; RATAN, A.; MCGRAW, J.; FRAZIER, M.; MASIGA, D.; SCHUSTER, S.; GROZINGER, C.M.; MILLER, W. Genome-wide analysis of signatures of selection in populations of African honey bees (*Apis mellifera*) using new web-based tools. **BMC Genomics**, v.16, art.518, 2015. DOI: 10.1186/s12864-015-1712-0.

GIBSON, J.; MORTON, N.E.; COLLINS, A. Extended tracts of homozygosity in outbred human populations. **Human Molecular Genetics**, v.15, p.789-795, 2006. DOI: 10.1093/hmg/ddi493.

GOLDEN HELIX. **SNP & Variation Suite**. Available at: <http://goldenhelix.com/forms/SNP_Variation/pricing.html>. Accessed on: July 4 2017.

GOMEZ-RAYA, L.; RODRÍGUEZ, C.; BARRAGÁN, C.; SILIÓ, L. Genomic inbreeding coefficients based on the distribution of the length of runs of homozygosity in a closed line of Iberian pigs. **Genetics Selection Evolution**, v.47, art.81, 2015. DOI: 10.1186/s12711-015-0153-1.

GURGUL, A.; SZMATOLA, T.; TOPOLSKI, P.; JASIELCZUK, I.; ŻUKOWSKI, K.; BUGNO-PONIEWIERSKA, M. The use of runs of homozygosity for estimation of recent inbreeding in Holstein cattle. **Journal of Applied Genetics**, v.57, p.527-530, 2016. DOI: 10.1007/s13353-016-0337-6.

HOWARD, J.T.; HAILE-MARIAM, M.; PRYCE, J.E.; MALTECCA, C. Investigation of regions impacting inbreeding depression and their association with the additive genetic effect for United States and Australia Jersey dairy cattle. **BMC Genomics**, v.16, art.813, 2015. DOI: 10.1186/s12864-015-2001-7.

HOWRIGAN, D.P.; SIMONSON, M.A.; KELLER, M.C. Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms. **BMC Genomics**, v.12, art.460, 2011. DOI: 10.1186/1471-2164-12-460.

HUDSON, N.J.; PORTO-NETO, L.; KIJAS, J.W.; REVERTER, A. Compression distance can discriminate animals by genetic profile, build relationship matrices and estimate breeding values. **Genetics Selection Evolution**, v.47, art.78, 2015. DOI: 10.1186/s12711-015-0158-9.

KIM, E.-S.; COLE, J.B.; HUSON, H.; WIGGANS, G.R.; VAN TASSELL, C.P.; CROOKER, B.A.; LIU, G.; DA, Y.; SONSTEGARD, T.S. Effect of artificial selection on runs of homozygosity in U.S. Holstein cattle. **PLoS ONE**, v.8, e80813, 2013. DOI: 10.1371/journal.pone.0080813.

KIM, E.-S.; SONSTEGARD, T.S.; ROTHSCHILD, M.F. Recent artificial selection in U.S. Jersey cattle impacts autozygosity levels of specific genomic regions. **BMC Genomics**, v.16, art.302, 2015a. DOI: 10.1186/s12864-015-1500-x.

KIM, E.-S.; SONSTEGARD, T.S.; VAN TASSELL, C.P.; WIGGANS, G.; ROTHSCHILD, M.F. The relationship between runs of homozygosity and inbreeding in Jersey cattle under selection. **PLoS ONE**, v.10, e0129967, 2015b. DOI: 10.1371/journal.pone.0129967.

LENCZ, T.; LAMBERT, C.; DEROSSE, P.; BURDICK, K.E.; MORGAN, T.V.; KANE, J.M.; KUCHERLAPATI, R.; MALHOTRA, A.K. Runs of homozygosity reveal highly penetrant recessive loci in schizophrenia. **Proceedings of the National Academy of Sciences of the United States of America**, v.104, p.19942-19947, 2007. DOI: 10.1073/pnas.0710021104.

MARRAS, G.; GASPA, G.; SORBOLINI, S.; DIMAURO, C.; AJMONE-MARSAN, P.; VALENTINI, A.; WILLIAMS, J.L.; MACCIOTTA, N.P.P. Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. **Animal Genetics**, v.46, p.110-121, 2015. DOI: 10.1111/age.12259.

MASTRANGELO, S.; TOLONE, M.; DI GERLANDO, R.; FONTANESI, L.; SARDINA, M.T.; PORTOLANO, B. Genomic inbreeding estimation in small populations: evaluation of runs of homozygosity in three local dairy cattle breeds. **Animal**, v.10, p.746-754, 2016. DOI: 10.1017/S1751731115002943.

MCQUILLAN, R.; LEUTENEGGER, A.-L.; ABDEL-RAHMAN, R.; FRANKLIN, C.S.; PERICIC, M.; BARAC-LAUC, L.; SMOLEJ-NARANCIC, N.; JANICIJEVIC, B.; POLASEK, O.; TENESA, A.; MACLEOD, A.K.; FARRINGTON, S.M.; RUDAN, P.; HAYWARD, C.; VITART, V.; RUDAN, I.; WILD, S.H.; DUNLOP, M.G.; WRIGHT, A.F.; CAMPBELL, H.; WILSON, J.F. Runs of homozygosity in European populations. **American Journal of Human Genetics**, v.83, p.359-372, 2008. DOI: 10.1016/j.ajhg.2008.08.007.

METZGER, J.; KARWATH, M.; TONDA, R.; BELTRAN, S.; ÁGUEDA, L.; GUT, M.; GUT, I.G.; DISTL, O. Runs of homozygosity reveal signatures of positive selection for reproduction traits in breed and non-breed horses. **BMC Genomics**, v.16, art.764, 2015. DOI: 10.1186/s12864-015-1977-3.

NALLS, M.A.; GUERREIRO, R.J.; SIMON-SANCHEZ, J.; BRAS, J.T.; TRAYNOR, B.J.; GIBBS, J.R.; LAUNER, L.; HARDY, J.; SINGLETON, A.B. Extended tracts of homozygosity identify novel candidate genes associated with late-onset Alzheimer's disease. **Neurogenetics**, v.10, p.183-190, 2009. DOI: 10.1007/s10048-009-0182-4.

O'BRIEN, A.M.P.; UTSUNOMIYA, Y.T.; MÉSZÁROS, G.; BICKHART, D.M.; LIU, G.E.; VAN TASSELL, C.P.; SONSTEGARD, T.S.; DA SILVA, M.V.; GARCIA, J.F.; SÖLKNER, J. Assessing signatures of selection through variation in linkage disequilibrium between taurine and indicine cattle. **Genetics Selection Evolution**, v.46, art.19, 2014. DOI: 10.1186/1297-9686-46-19.

PEMBERTON, T.J.; ABSHER, D.; FELDMAN, M.W.; MYERS, R.M.; ROSENBERG, N.A.; LI, J.Z. Genomic patterns of homozygosity in worldwide human populations. **American Journal of Human Genetics**, v.91, p.275-292, 2012. DOI: 10.1016/j.ajhg.2012.06.014.

PERIPOLLI, E.; MUNARI, D.P.; SILVA, M.V.G.B.; LIMA, A.L.F.; IRGANG, R.; BALDI, F. Runs of homozygosity: current knowledge and applications in livestock. **Animal Genetics**, v.48, p.255-271, 2017. DOI: 10.1111/age.12526.

PÉRTILE, S.F.N.; SILVA, F.F. e; SALVIAN, M.; MOURÃO, G.B. Seleção e associação genômica ampla para o melhoramento genético animal com uso do método ssGBLUP. **Pesquisa Agropecuaria Brasileira**, v.51, p.1729-1736, 2016. DOI: 10.1590/s0100-204x2016001000004.

PLINK…: whole genome association analysis toolset. Available at: <http://zzz.bwh.harvard.edu/plink/contact.shtml#cite>. Accessed on: July 4 2017.

PURCELL, S.; NEALE, B.; TODD-BROWN, K.; THOMAS, L.; FERREIRA, M.A.R.; BENDER, D.; MALLER, J.; SKLAR, P.; BAKKER, P.I.W. de; DALY, M.J.; SHAM, P.C. PLINK: a tool set for whole-genome association and population-based linkage analysis. **American Journal of Human Genetics**, v.81, p.559-575, 2007. DOI: 10.1086/519795.

PURFIELD, D.C.; BERRY, D.P.; MCPARLAND, S.; BRADLEY, D.G. Runs of homozygosity and population history in cattle. **BMC genetics**, v.13, art.70, 2012. DOI: 10.1186/1471-2156-13-70.

RAMEY, H.R.; DECKER, J.E.; MCKAY, S.D.; ROLF, M.M; SCHNABEL, R.D.; TAYLOR, J.F. Detection of selective sweeps in cattle using genome-wide SNP data. **BMC Genomics**, v.14, art.382, 2013. DOI: 10.1186/1471-2164-14-382.

RANDHAWA, I.A.S.; KHATKAR, M.S.; THOMSON, P.C.; RAADSMA, H.W. A meta-assembly of selection signatures in cattle. **PLoS ONE**, v.11, e0153013, 2016. DOI: 10.1371/journal.pone.0153013.

REVERTER, A.; PORTO-NETO, L.R.; FORTES, M.R.S.; KASARAPU, P.; CARA, M.A.R. de; BURROW, H.M.; LEHNERT, S.A. Genomic inbreeding depression for climatic adaptation of tropical beef cattle. **Journal of Animal Science**, v.95, p.3809-3821, 2017. DOI: 10.2527/jas2017.1643.

SAURA, M.; FERNÁNDEZ, A.; VARONA, L.; FERNÁNDEZ, A.I.; CARA, M.Á.R. de; BARRAGÁN, C.; VILLANUEVA, B. Detecting inbreeding depression for reproductive traits in Iberian pigs using genome-wide data. **Genetics Selection Evolution**, v.47, art.1, 2015. DOI: 10.1186/s12711-014-0081-5.

SCRAGGS, E.; ZANELLA, R.; WOJTOWICZ, A.; TAYLOR, J.F.; GASKINS, C.T.; REEVES, J.J.; AVILA, J.M. de; NEIBERGS, H.L. Estimation of inbreeding and effective population size of full-blood wagyu cattle registered with the American Wagyu Cattle Association. **Journal of Animal Breeding and Genetics**, v.131, p.3-10, 2014. DOI: 10.1111/jbg.12066.

SILVA, J.M. da; GIACHETTO, P.F.; SILVA, L.O.C. da; CINTRA, L.C.; PAIVA, S.R.; CAETANO, A.R.; YAMAGISHI, M.E.B. Genomic variants revealed by invariably missing genotypes in Nelore cattle. **PLoS ONE**, v.10, e0136035, 2015. DOI: 10.1371/journal.pone.0136035.

SÖLKNER, J.; FERENČAKOVIĆ, M.; GREDLER, B.; CURIK, I. Genomic metrics of individual autozygosity, applied to a cattle population. In: ANNUAL MEETING OF THE EUROPEAN ASSOCIATION OF ANIMAL PRODUCTION, 61., 2010, Heraklion, Greece. **Book of Abstracts**. Netherlands: Wageningen Academic, 2010. p.306. (Book of Abstracts nº 16).

SOMAVILLA, A.L.; SONSTEGARD, T.S.; HIGA, R.H.; ROSA, A.N.; SIQUEIRA, F.; SILVA, L.O.C.; TORRES JÚNIOR, R.A.A.; COUTINHO, L.L.; MUDADU, M.A.; ALENCAR, M.M.; REGITANO, L.C.A. A genome-wide scan for selection signatures in Nellore cattle. **Animal Genetics**, v.45, p.771-781, 2014. DOI: 10.1111/age.12210.

SORBOLINI, S.; MARRAS, G.; GASPA, G.; DIMAURO, C.; CELLESI, M.; VALENTINI, A.; MACCIOTTA, N.P.P. Detection of selection signatures in Piemontese and Marchigiana cattle, two breeds with similar production aptitudes but different selection histories. **Genetics Selection Evolution**, v.47, art.52, 2015. DOI: 10.1186/s12711-015-0128-2.

SZMATOŁA, T.; GURGUL, A.; ROPKA-MOLIK, K.; JASIELCZUK, I.; ZĄBEK, T.; BUGNO-PONIEWIERSKA, M. Characteristics of runs of homozygosity in selected cattle breeds maintained in Poland. **Livestock Science**, v.188, p.72-80, 2016. DOI: 10.1016/j.livsci.2016.04.006.

VANRADEN, P.M. Accounting for inbreeding and crossbreeding in genetic evaluation of large populations. **Journal of Dairy Science**, v.11, p.3136-3144, 1992. DOI: 10.3168/jds.S0022-0302(92)78077-1.

VANRADEN, P.M. Efficient methods to compute genomic predictions. **Journal of Dairy Science**, v.91, p.4414-4423, 2008. DOI: 10.3168/jds.2007-0980.

WOLC, A.; ZHAO, H.H.; ARANGO, J.; SETTAR, P.; FULTON, J.E.; O'SULLIVAN, N.P.; PREISINGER, R.; STRICKER, C.; HABIER, D.; FERNANDO, R.L.; GARRICK, D.J.; LAMONT, S.J.; DEKKERS, J.C.M. Response and inbreeding from a genomic selection experiment in layer chickens. **Genetics Selection Evolution**, v.47, art.59, 2015. DOI: 10.1186/s12711-015-0133-5.

WRIGHT, S. Coefficients of Inbreeding and Relationship. **The American Naturalist**, v.56, p.330-338, 1922. DOI: 10.1086/279872.

YANG, J.; LEE, S.H.; GODDARD, M.E.; VISSCHER, P.M. GCTA: a tool for genome-wide complex trait analysis. **American Journal of Human Genetics**, v.88, p.76-82, 2011. DOI: 10.1016/j.ajhg.2010.11.011.

ZAVAREZ, L.B.; UTSUNOMIYA, Y.T.; CARMO, A.S.; NEVES, H.H.R.; CARVALHEIRO, R.; FERENČAKOVIĆ, M.; PÉREZ O'BRIEN, A.M.; CURIK, I.; COLE, J.B.; VAN TASSELL, C.P.; SILVA, M.V.G.B. da; SONSTEGARD, T.S.; SÖLKNER, J.; GARCIA, J.F. Assessment of autozygosity in Nellore cows (*Bos indicus*) through high-density SNP genotypes. **Frontiers in Genetics**, v.6, p.1-8, 2015. DOI: 10.3389/fgene.2015.00005.

ZHANG, Q.; CALUS, M.P.L.; GULDBRANDTSEN, B.; LUND, M.S.; SAHANA, G. Estimation of inbreeding using pedigree, 50k SNP chip genotypes and full sequence data in three cattle breeds. **BMC Genetics**, v.16, art.88, 2015a. DOI: 10.1186/s12863-015-0227-7.

ZHANG, Q.; GULDBRANDTSEN, B.; BOSSE, M.; LUND, M.S.; SAHANA, G. Runs of homozygosity and distribution of functional variants in the cattle genome. **BMC Genomics**, v.16, art.542, 2015b. DOI: 10.1186/s12864-015-1715-x.