EasyNRM: A Visual Basic Application Approach for Computing Numerator Relationship Matrix of Pedigreed Animals

Samet ÇİNKAYA[1], Mustafa TEKERLİ[2], Mustafa DEMİRTAŞ[3], Koray ÇELİKELOĞLU[4]
Department of Animal Science, Faculty of Veterinary Medicine, Kocatepe University, Afyonkarahisar, Turkey
[1]https://orcid.org/0000-0003-0736-4571, [2]https://orcid.org/0000-0002-8634-1193, [3] https://orcid.org/0000-0002-8349-9649
[4]https://orcid.org/0000-0002-1610-2226
✉: scinkaya@aku.edu.tr

ABSTRACT

The relationship and inbreeding are special concepts in animal breeding, for this reason, a free of charge VBA software was developed to calculate the coefficients of relationship and inbreeding in Microsoft Excel. The tabular method algorithm was implemented by using pedigree. The software was called as EasyNRM and verified with the inclusion of bull Comet's pedigree. Finally, the numerator relationship matrix was created and certified with the results of the literature. It would be beneficial for EasyNRM to be used by researchers and breeders in estimating the breeding values of animals.

Pedigrili Hayvanlarda Paydan Akrabalık Matrisini Hesaplamaya Yönelik Bir Visual Basic Uygulaması Yaklaşımı: EasyNRM

ÖZET

Hayvancılıkta akrabalık ve akrabalı yetiştirme ayrıcalıklı öneme sahip kavramlardır. Bu nedenle yaygın kullanılan Microsoft Excel ortamında adı geçen değerleri hesaplamak üzere ücretsiz ve özgün bir yazılım geliştirilmiştir. Programda soykütüğüne dayalı tablo yöntemi uygulanmıştır. Yazılım EasyNRM olarak adlandırılmış ve bu konuda kaynaklar tarafından bilinen Comet isimli boğanın pedigrisi ile denenmiştir. Sonuç olarak Excel sayfalarında paydan akrabalık matrisi ve akrabalı yetiştirme katsayıları hesaplanmış ve bu bulgular literatür ile teyit edilmiştir. Bu yazılımın damızlık değer tahmininde akrabalık ilişkilerini de dikkate almak isteyen araştırmacılar ve yetiştiriciler için yararlı olacağı kanaatine varılmıştır.

To Cite : Cinkaya S, Tekerli M, Demirtaş M, Çelikeloğlu K 2019. EasyNRM: A Visual Basic Application Approach for Computing Numerator Relationship Matrix of Pedigreed Animals. KSU J. Agric Nat 22(Suppl 2): 418-423. DOI: 10.18016/ksutarimdoga.vi.571575

INTRODUCTION

The inbreeding coefficient of an animal is defined as the probability that two alleles in a locus of that animal are identical by descent from a common ancestor. The formula for the inbreeding coefficient is as follows:

$$F_X = \sum_{CA=1}^{k} \left(\frac{1}{2}\right)^{n_1+n_2+1} \cdot (1 + F_A) \qquad (1)$$

Where F_X is the inbreeding coefficient of the individual X; CA is a common ancestor of the sire and dam of X; k is the number of common ancestors in X's pedigree; n_1 is the number of generations separating the common ancestor from the sire of X; n_2 is the number of generations separating the common ancestor from the dam of X; F_A is the inbreeding coefficient of the common ancestor A, from whom the lines of descent arise. The relationship of parents of individual X reflects the genotypic similarity due to common ancestors. The formula for expressing the degree of relationship, that is the genetic covariance between two individuals, is:

$$r_{SD} = \sum \left[\left(\frac{1}{2}\right)^{n_1+n_2} \cdot (1 + F_A)\right] \qquad (2)$$

Due to it is the numerator of Wright's coefficient of relationship (Wright, 1922) between sire (S) and dam (D), it is called as the *numerator relationship*. The complete formula is expressed as follows:

$$R_{SD} = \frac{r_{SD}}{\sqrt{[(1 + F_S)(1 + F_D)]}} \qquad (3)$$

So, the measure of relationship has been corrected for the effects of inbreeding. The numerator relationship for any pair of parents is twice the value of the inbreeding coefficient of the offspring. The numerator relationship matrix (NRM) of animals in a population is needed to estimate more accurate breeding values. The inverse of NRM is used in mixed models to reliably predict genetic parameters and breeding values (Bourdon, 2013; Kerr et al., 2012; Kaygısız and Kösetürkmen, 2011; Nicholas, 2010; Legates and Warwick, 1990; Tier, 1990; Van Vleck et al., 1987; Henderson, 1976; Bıyıkoğlu, 1973; Emik and Terrill, 1949; Lush, 2011; Wright, 1922).

There is a number of softwares (Cole, 2007; Sargolzaei et al., 2006; Gutiérrez and Goyache, 2005; Boichard, 2002) using different computing languages and algorithms for calculating inbreeding and relationship coefficients, NRM and inverse of NRM. The tabular method suggested by different researchers (Nicholas, 2010; Van Vleck et al., 1987; Cruden, 1949; Emik and Terrill, 1949) is practiced for computational ease.

Microsoft Excel [Microsoft Corp., CA, USA] harbours a type of visual programming language named Visual Basic for Application (VBA) to bring under control the different functions. Excel is widely used by animal scientists for data processing. An intuitive VBA program, called EasyNRM, was written by using the tabular method for generating the numerator relationship matrix, inverse of NRM and Wright's relationship coefficient matrix in pedigreed populations.

MATERIAL AND METHODS
An Overview

The program was developed at the Laboratory of Department of Animal Science in the Faculty of Veterinary Medicine, Afyon Kocatepe University. This program was created in .xlam format. It can work in every Excel .xlsx extention files when it is loaded as an add-in. To install the program "File", "Options", "Add-ins" and "Go" are clicked to select an add-in type respectively, and then "Add-ins" are browsed to locate and EasyNRM.xlam is selected by pressing "Ok". After the excel file containing the pedigree was opened, it can be run using the "ctrl+shift+[c]" shortcut. The software comprises of codes in a module and visual components. Each subroutine in the module had been connected with the individual buttons on the userform. The userform including the respective buttons, textboxes showing elapsed time for calculation, and a progress bar was presented in Figure 1.
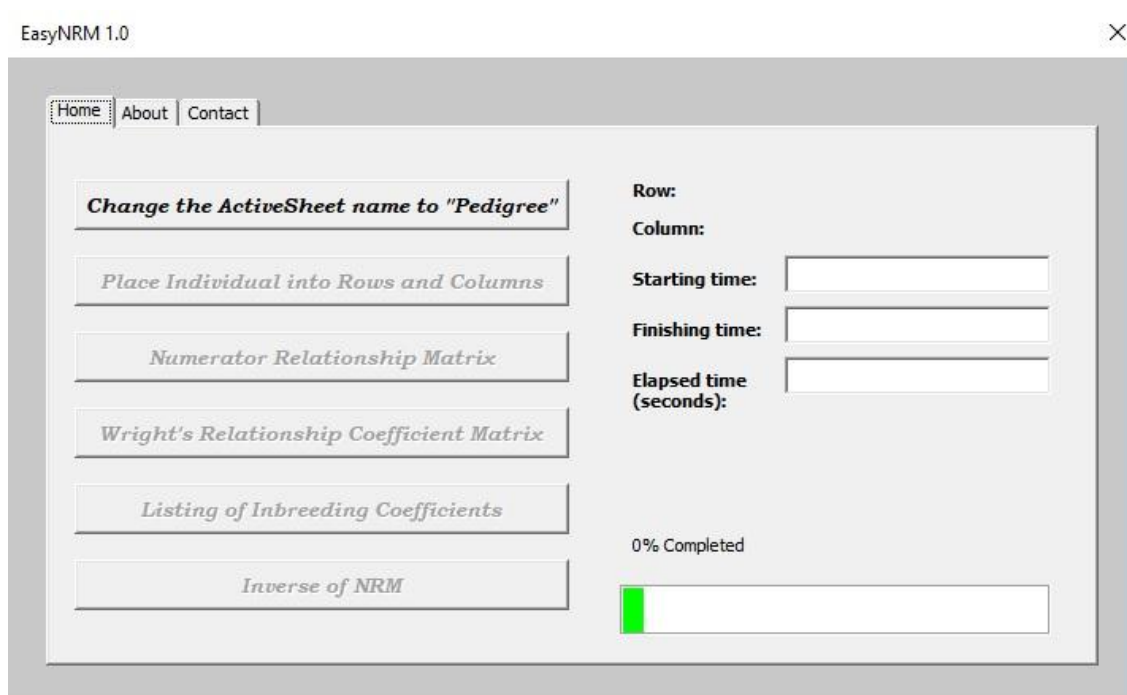

Figure 1. EasyNRM Home page view.

Preparing Pedigree File

Before the inbreeding and relationship coefficients are computed, some prerequisites have to be fulfilled. Alphanumeric individual IDs must be converted to numeric ones. Parents' ID numbers should not be greater than that of their offspring. Animal, sire and dam IDs must be written down the first, second and third columns, respectively. Individual IDs are sorted from the oldest to the youngest according to the date of birth. Animals with problematic records must be pruned from the pedigree file. Editorial softwares such as RE-NUM-OR (Yazgan, 2018) or Pedigri Yıldızı (Tekerli, 2018) would be useful for reordering and renumbering of animals in the pedigree file. According to VanRaden (1992), inbreeding coefficients of the animals with unknown parents may be equalized to

the mean of animals with known parents born in the same period (Aquilar and Misztal, 2008). Because of this fact, the inbreeding coefficients of animals with unknown parents can be manually replaced with the mean after emerging of NRM. Provided that it has .xlsx extensioned any simple name can be given to pedigree file containing animal, sire and dam IDs. It must be considered that the maximum column number for .xlsx extension is 16.384. EasyNRM is restricted with Excel's infrastructure. If the file format is .xls, the number of columns will be limited with 256. The pedigree worksheet is automatically renamed as "*Pedigree*" by pressing the corresponding button.

Functioning of the software

How to use the tabular method to construct the numerator relationship matrix was illustrated by different researchers (Legates and Warwick, 1990; Nicholas, 2010; Tier, 1990; Van Vleck et al., 1987). The EasyNRM constitutes the NRM based on animal pedigrees starting from the individuals of the main population in which the relationship between their parents is unknown or ignored. The algorithm was presented step by step. Every phase is realized by pressing its respective button. The program starts with the button named "*Change the ActiveSheet name to Pedigree*" and traces the following steps.

Step 1: Placing the IDs of individuals and their parents into rows and columns

The individual IDs are written to the second row and column on the newly created worksheets called "*NumeratorRelationshipMatrix*" and "*RelationshipCoefficientMatrix*", from up to down, and left to right. Parent IDs (Sire and Dam) are written to column cells in the first row on the top of the individuals which are placed in the second row. "0" is placed instead of the IDs for unknown parents. The operations until now can be executed by pressing the button "*Place Individuals into Rows and Columns*" on the userform (Figure 1).

Step 2: Computing the Numerator Relationship Matrix

The calculation is launched from the first cell of diagonal. When the presence of a relationship between parents, since the relationship of an individual with itself (X) is 1 by definition, the animal's own relationship coefficient (R_{XX}) is equal to 1 plus its inbreeding coefficient (F_X). Inbreeding coefficient (F_X) of the individual, is half of the relationship between parents and is calculated according to the formula $F_X = \left(\frac{1}{2}\right).r_{SD}$ from the cell where parents' rows and columns intersect. This is the formula of the inbreeding coefficient for an individual. The values of each off-diagonal cells are computed by getting half of the relationship between the individual (Y) in the row and the parents (S and D) of the individual (X) in the

column according to the formula $r_{XY} = \left(\frac{1}{2}\right).(r_{SY} + r_{DY})$. After every individual row has been finished, its transpose is copied to the column having the same ID. The computation continues until the NRM is completed. The interpreter executes the codes associated with numerator relationship matrix after pressing the button "*Numerator Relationship Matrix*" (Figure 1).

Step 3: Computation of Wright's Relationship Coefficients Matrix

The denominator of the relationship coefficient may not be higher than 1.0 in the majority of the herds. Therefore, this part of Equation 3 can be ignored in real life. However, Lush (2011) declared that the inbred lines in a population show the tendencies for drifting apart from each other and so the inbreeding makes the population more variable than expected. Therefore, terms in the denominator of Equation 3 must be used to consider the fact that a couple of animals may be inbred individually. In this connection, the numerator relationship at the intersecting cell of two individuals is determined. Then the value at this cell divided by $\sqrt{(1+F_X).(1+F_Y)}$ to uncover the true relationship coefficient of Wright. This process is performed by pressing the "*Wright's Relationship Coefficient Matrix*" button in EasyNRM's home page window (Figure 1).

Step 4: Listing of Inbreeding Coefficients

After the formation of the numerator relationship matrix, the inbreeding coefficient ($F_X = r_{XX} - 1$) of each individual is calculated from the diagonal of the NRM sheet and listed on the newly created worksheet called "*InbreedingCoefficients*". This running can be accomplished by pressing the "*Listing of Inbreeding Coefficients*" button.

Step 5: Calculating the inverse of Numerator Relationship Matrix

NRM is used in BLUP (best linear unbiased prediction) analysis. So, NRM is useful for adjusting biases in the estimation of breeding values in mixed models (Henderson, 1976; Kim et al., 2016). The inverse of NRM is applied for this purpose. Through, by pressing the last button called "*Inverse of NRM*" this aim can be achieved.

RESULTS AND DISCUSSION
Testing of the Software with an Approved Example.

The EasyNRM program, running with an algorithm according to the tabular method was used to calculate the inbreeding coefficient of a bull named Comet with the pedigree presented in Figure 2. This sample was given in previous literature (García-Cortés et al., 2010; Wright, 1922).

Firstly, the names were converted to IDs manually

before the program was executed (Figure 3).

All the outcomes were obtained by clicking the respective buttons. Figures 4, 5, 6, 7 and 8 represent the results as individuals and their parents in the rows and columns, NRM, and Wright's relationship coefficient matrix, individual inbreeding coefficient and inverse of NRM, respectively.

Figure 2. The real pedigree of the Comet bull (García-Cortés et al., 2010; Wright, 1922).

Figure 3. The numerical pedigree of the Comet bull after conversion.

Comet's inbreeding coefficient was found to be 0.4687. This was consonant with the findings by Garcia-Cortes et al. (2010) and Wright (1922). After computing the NRM according to Equation 2 it was recalculated with Equation 3 for Wright's Relationship Coefficient Matrix. Individual inbreeding coefficients were listed on the "*InbreedingCoefficients*" page (Figure 7).

In the last step, the inverse of NRM is found and shown on the page "NRMinverse" (Figure 8).

Figure 4. Placing of individual animals and their parents on NumeratorRelationshipMatrix and RelationshipCoefficientMatrix sheet.

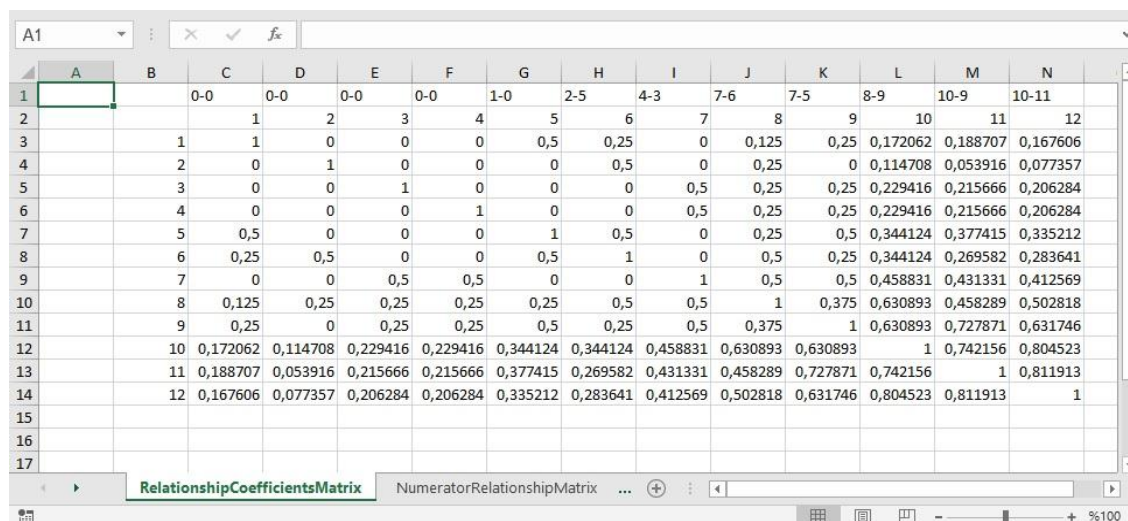Figure 5. Numerator Relationship Matrix.

Figure 6. Wright's Relationship Coefficients Matrix.

Due to its visual components and being hosted in excel the EasyNRM is different from the other similar softwares such as PEDIG (Boichard, 2002), PyPedal (Cole, 2007), ENDOG (Gutiérrez and Goyache, 2005), Numericware N (Kim et al., 2016), CFC (Sargolzaei et al., 2006) and RELAX2 (Stranden and Vuori, 2006). These programs written in Fortran (Boichard, 2002; Stranden and Vuori, 2006), Python (Cole, 2007) and C++ (Kim et al., 2016; Sargolzaei, 2006) are available for pedigree analyses. EasyNRM shows a difference due to written in VBA. All inputs and outputs can be viewed in a single Excel workbook. Different types of data files are easily read in Excel. Whereas the majority of the others works with only a given type of data files. For example, PyPedal (Cole, 2007) and INBUPGF90 (Aquilar and Misztal, 2008) handle with ASCII formatted files for input. However, ENDOG (Gutierrez and Goyache, 2005) can use .xls files.
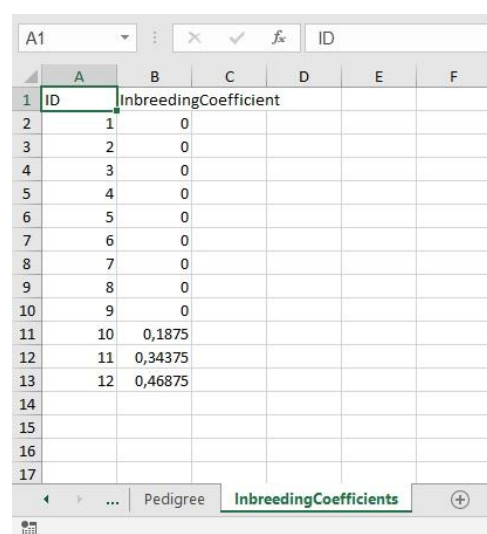

Figure 7. Inbreeding coefficients of individuals.

Figure 8. Inverse of NRM

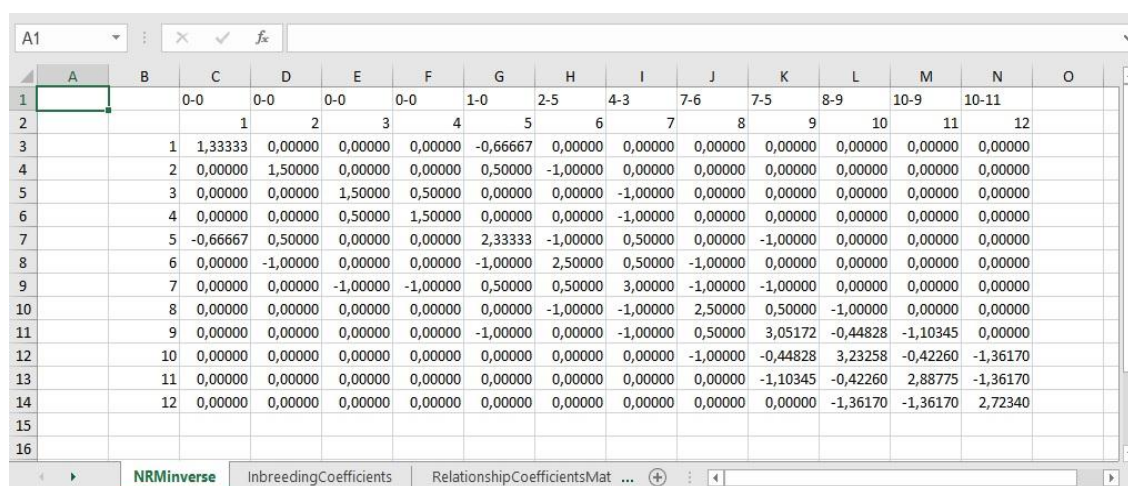EasyNRM was manufactured by using visual components of Excel VBA. Therefore, it is easy to use by the users relative to above mentioned programs. Especially, beginners can learn and operate this program with relative ease. In addition, Excel functionality gives some possibilities to extend the software. For example, Excel's MInverse function can be used to find the inverse of NRM, without extra codes.

CONCLUSION

EasyNRM is user friendly and can be applied to plant and human pedigrees besides animals. It has also potential as a component for using in any software incorporating BLUP subroutines. It would be useful for the researchers interested with statistical mixed models, breeders who want to develop the successful mating strategies for their herds and the students of population genetics. The EasyNRM is free of charge. The software, Comet's pedigree sample, and instruction video are available from the www.sametcinkaya.com.

REFERENCES

Aguilar I, Misztal I 2008. Recursive algorithm for inbreeding coefficients assuming nonzero inbreeding of unknown parents. Journal of Dairy science, 91(4): 1669-1672.

Bıyıkoğlu MK 1973. Hayvancılıkta Uygulanan Yetiştirme Sistemleri (in Turkish). Erzurum, Turkey: Atatürk Üniversitesi Yayınları.

Boichard D 2002. PEDIG: A Fortran package for pedigree analysis suited for large populations. Proceedings of the 7th world congress on genetics applied to livestock production, Montpellier, France.

Bourdon RM 2013. Understanding Animal Breeding 2nd ed. Pearson Education Limited, Essex, England.

Cole JB 2007. PyPedal: A computer program for pedigree analysis. Computers and Electronics in Agriculture, 57 (1): 107-113. Doi: 10.1016/j.compag.2007.02.002

Cruden D 1949. The computation of inbreeding coefficients for closed populations. Journal of Heredity, 40 (9): 248-251.

Emik LO, Terrill CE 1949. Systematic procedures for calculating inbreeding coefficients. Journal of Heredity, 40 (2): 51-55. doi: 10.1093/oxfordjournals.jhered.a105986

García-Cortés LA, Martínez-Ávila JC, Toro MA 2010. Fine decomposition of the inbreeding and the coancestry coefficients by using the tabular method. Conservation Genetics, 11 (5): 1945-1952. doi: 10.1007/s10592-010-0084-x

Gutiérrez JP, Goyache F 2005. A note on ENDOG: a computer program for analysing pedigree information. Journal of Animal Breeding and Genetics, 122 (3): 172-176. Doi: 10.1111/j.1439-0388.2005.00512.x

Henderson CR 1976. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics, 32 (1): 69-83. Doi: 10.2307/2529339

Kaygısız A, Kösetürkmen E 2011. İsviçre Esmeri Sığırlarda Akrabalı Yetiştirme Düzeyi ve Akrabalı Yetiştirmenin Süt Verim Özellikleri Üzerine Etkileri (in Turkish). Kahramanmaraş Sütçü İmam Üniversitesi Doğa Bilimleri Dergisi, 14(2): 1-6.

Kerr RJ, Li L, Tier B, Dutkowski GW, McRae TA 2012. Use of the numerator relationship matrix in genetic analysis of autopolyploid species. Theoretical and Applied Genetics, 124 (7): 1271-1282. Doi: 10.1007/s00122-012-1785-y

Kim B, Beavis WD, Léon J 2016. Numericware N: numerator relationship matrix calculator. Journal of Heredity, 107 (7): 686-690. Doi: 10.1093/jhered/esw057

Legates JE, Warwick EJ 1990. Breeding and Improvement of Farm Animals. McGraw Hill Inc, New York, USA.

Lush JL 2011. Animal breeding plans (Second Edition). Breiningsville, PA USA.

Nicholas F 2010. Introduction to Veterinary Genetics. 1st ed. Wiley-Blackwell, Singapore.

Sargolzaei M, Iwaisaki H, Colleau JJ 2006. CFC: A tool for monitoring genetic diversity. 8th World Congress Genetics Applied Livestock Producion, Belo Horizonte, Brazil.

StackOverflow-community, Test or check if sheet exists, https://stackoverflow.com/questions/6688131/test-or-check-if-sheet-exists, 2019.

Strandén I, Vuori K 2006. RelaX2: pedigree analysis programme. In Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Minas Gerais, Brazil, 13-18 August, 2006 (pp. 27-30). Instituto Prociência.

Tekerli M 2018. Pedigri Yıldızı, Excel based unpublished pedigree software.

Tier B 1990. Computing inbreeding coefficients quickly. Genetics Selection Evolution, 22 (4): 419-430. Doi:10.1186/1297-9686-22-4-419

VanRaden PM 1992. Accounting for inbreeding and crossbreeding in genetic evaluation of large populations. Journal of Dairy Science, 75(11): 3136-3144.

Van Vleck LD, Pollak EJ, Oltenacu EA 1987. Genetics for the Animal Sciences. W.H. Freeman and Company, New York, USA.

Wright S 1922. Coefficients of Inbreeding and Relationship. The American Naturalist 56 (645): 330-338. Retrieved from http://www.jstor.org/stable/2456273

Yazgan K 2018. RE-NUM-OR: Python-based renumbering and reordering software for pedigree files. Czech Journal of Animal Science, 63(2): 70-77.