—Full Paper—

Whole-genome resequencing using genomic DNA extracted from horsehair roots for gene-doping control in horse sports

Teruaki TOZAKI[1][#]*, Aoi OHNUMA[1][#], Mio KIKUCHI[1], Taichiro ISHIGE[1], Hironaga KAKOI[1], Kei-ichi HIROTA[1], Natasha A. HAMILTON[2], Kanichi KUSANO[3] and Shun-ichi NAGATA[1]

[1]Genetic Analysis Department, Laboratory of Racing Chemistry, Tochigi 320-0851, Japan
[2]Equine Genetics Research Centre, Racing Australia, NSW 2337, Australia
[3]Equine Department, Japan Racing Association, Tokyo 106-8401, Japan

Gene doping is prohibited in horseracing and equestrian sports. In previous studies, we developed non-targeted transgene and genome editing detection methods based on whole genome resequencing (WGR) using genomic DNA extracted from whole blood. In this study, we aimed to develop a WGR method using DNA extracts from hair roots. Hair roots are a preferred substrate because their collection is less invasive than blood collection. Hair is also easier to store for long periods of time. Although almost all genomic DNA extracted from hair root samples stored for years at room temperature was degraded, the quality of genomic DNA from samples stored for years at refrigerated temperatures (4–8°C) was maintained. High-molecular-weight genomic DNA was isolated from hair roots using a magnetic silica beads method of extraction, enabling WGR from horsehair root extracts. Nucleotide sequencing results and numbers of single-nucleotide polymorphisms and insertions/deletions concurred with those previously reported for WGR of DNA extracted from whole blood. Therefore, we consider that storing hair samples at refrigerated temperatures prevents degradation of DNA, allowing the detection of gene doping in these samples based on WGR. It is likely this finding will also have a deterrent effect, as it is now possible to test horses with archived samples even if they or their parents are deceased. To our knowledge, this is the first report employing WGR on horsehair roots stored for a long term.

Key words: *gene doping, hair root, horse racing, next generation sequencer, Thoroughbred*

The Thoroughbred is the horse breed best known for racing worldwide [3], while several different horse breeds are used in equestrian sports such as show jumping, dressage, and eventing [1, 6]. These breeds were developed by selective breeding over many generations to achieve ideals set out in the breed standards. Consequently, many equine Stud Books require parental verification of registered horses to guarantee their eligibility. Therefore, the genetic integrity of these horses is essential for registration, fair competition, and the integrity of the breeds themselves. The International Federation of Horseracing Authorities (IFHA) [11] and Fédération Equestre Internationale (FEI) [2] are the internationally recognized industry bodies that oversee the governance of horseracing (IFHA) and the Olympics and World Championships (FEI). Both organizations prohibit gene doping.

Recent developments in molecular genetics have enabled the use of genetic therapies in humans and animals [5, 7, 8]. In horses, clinical studies of gene therapies using viral and plasmid vectors have been reported [13, 17]. Whilst these therapies are being developed as treatments for injury or chronic disease in horses, their use in healthy horses may be defined as gene doping in horseracing and equestrian sports. As these therapies are still under development, gene doping is considered a theoretical problem for authorities.

However, many laboratories are already developing anti-gene doping tests to address this issue.

Gene doping can be divided into two types. Gene doping based on traditional gene transfer therapy is most likely to be used on horses in training and competition. This type of doping comprises injection of a gene (called a transgene) packaged in a carrier, usually a plasmid or viral vector, into a horse to enhance its performance [4, 18]. More recently, the emergence of genome editing—also known as CRISPR—as a more refined gene therapy technique has highlighted the potential for editing an egg or embryo to produce genetically modified animals [15, 24]. This type of doping is a concern for racing Thoroughbreds, as the changes introduced through genome editing of an embryo or egg would be heritable. The Thoroughbred Stud Book, for example, prohibits the use of any type of reproductive technology, and thus manipulation of the genome could lead to a Thoroughbred being excluded from the Stud Book.

Several detection methods using quantitative polymerase chain reaction (qPCR) have been reported for gene transfer-based gene doping control in horses [10, 21, 22]. For detection of inserted transgenes, primers are designed to anneal to different exons, and a hydrolysis probe is designed on their exon/exon junction. Recently, a less specific genome-wide detection method using whole genome resequencing (WGR) was developed to detect genome editing and inserted transgenes through the identification of DNA variation, including single nucleotide polymorphisms (SNPs), insertions/deletions (INDELs), and structural variants (SVs) [23].

Parentage verification in horses is performed by examining the inheritance of microsatellite DNA markers [12, 20]. The DNA may be extracted from blood or the follicle (root) of mane or tail hair. According to an investigation by the International Stud Book Committee (ISBC) in 2016, at least 42 Thoroughbred Studbooks have used hair samples, while at least 19 have preferred blood samples. Many laboratories use hair samples because they are easier and less invasive than blood to collect and are cheaper to store. Using these same samples for gene doping control would be highly efficient, as they are already archived following a horse's registration, minimizing the number of samples that need to be collected. On the other hand, genomic DNA for WGR is conventionally extracted from fresh or cryo-preserved blood samples because high molecular weight genomic DNA is required for WGR library preparation. In this study, we compared methods for the extraction of high molecular weight genomic DNA from hair root samples and examined the effect of storage conditions on DNA stability.

Materials and Methods

Animal samples

Hair root and blood sample collection was approved by the Animal Care Committee of the Laboratory of Racing Chemistry (LRC, approval number 20-2). Hair samples from three Thoroughbred horses were collected at the Hidaka Training and Research Center, Japan Racing Association (JRA). Blood samples from three Thoroughbred horses were collected into EDTA tubes at the Miho Training Center, JRA. Hair and blood samples were stored at room temperature and −30°C, respectively, for approximately 1 month before use.

The LRC obtained hair samples from other breeds submitted by owners and/or veterinarians to the Japan Equine Affairs Association (JEAA) for pedigree registration. Consent for their use was obtained during the sample submission process. Hair root samples were collected in 2010 (three Japanese heavy draft horses: F_2010_1, foal; F_2010_2, sire; and F_2010_3, dam), 2012 (one horse of unknown breed: F_2012_1), 2015 (one horse of unknown breed: F_2015_1, mane, and F_2015_2, tail), and 2017 (one horse of unknown breed: F_2017_1) and stored at 4–8°C until 2020. Additionally, hair root samples were collected in 2009 (three Japanese native horses: R_2009_1, R_2009_2, and R_2009_3), 2010 (three Japanese native horses: R_2010_1, R_2010_2, and R_2010_3), 2011 (three Japanese native horses: R_2011_1, R_2011_2, and R_2011_3), and 2016 (three Japanese native horses: R_2016_1, R_2016_2, and R_2016_3) and stored at room temperature until 2020.

Extraction of genomic DNA using spin columns

Genomic DNA was extracted according to the manufacturer's recommendations from horsehair roots and whole blood using a DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany). For the hair samples, 15 hair roots were cut and collected into a 1.5 m*l* tube, digested for 60 min at 56°C in 180 *µl* of ATL Buffer plus 20 *µl* of Proteinase K solution, and then incubated for 10 min at 70°C after adding 200 *µl* of AL Buffer (Qiagen). For the blood samples, 200 *µl* of whole blood was digested for 10 min at 56°C in 200 *µl* of AL Buffer plus 20 *µl* of Proteinase K solution (Qiagen). For both types of samples, the solutions were then purified in DNeasy Mini spin columns (Qiagen) on a QIAcube automated system (Qiagen). Finally, genomic DNA was eluted in 200 *µl* of Milli-Q water.

Extraction of genomic DNA using silica beads

Genomic DNA was extracted from horsehair roots using a MagExtractor Genome Kit (Toyobo, Osaka, Japan), which utilizes a silica beads method. This was performed according

to the manufacturer's instructions with minor modifications. Briefly, 15 hair roots were cut, placed in a 1.5 m*l* tube, and digested for 60 min at 56°C in 90 μl of ATL Buffer plus 10 μl of Proteinase K solution from a DNeasy Blood & Tissue Kit (Qiagen). Next, 750 μl of Lysis & Binding Solution and 40 μl of magnetic silica beads from a MagExtractor Genome Kit (Toyobo) were added to the digested solution and mixed well for 10 min using a tube mixer. The magnetic silica beads that combined with genomic DNA were separated from the solution using a magnet stand, and the supernatant was removed. The beads were washed twice with 900 μl of Washing Solution (Toyobo) and twice with 900 μl of 70% ethanol by the magnetic separation. Finally, genomic DNA was eluted in 100 μl of Milli-Q water.

Quality and quantity checking of extracted DNA

Extracted genomic DNA was quantified by mass concentration (ng/μl) using a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, U.S.A.), according to the manufacturer's instructions. Degradation of genomic DNA was evaluated based on the genomic DNA quality score (GQS) assigned by a LabChip GX Touch 24 Nucleic Acid Analyzer (PerkinElmer, Waltham, MA, U.S.A.) using a Genomic DNA Reagent Kit (PerkinElmer). The GQS ranges from 0 to 5, with 5 representing the highest quality DNA. Extracted genomic DNA was also visualized using electrophoresis on a 1.0% agarose gel stained with ethidium bromide.

PCR amplification of extracted genomic DNA

Eighteen microsatellite markers employed in Japan for parentage verification in Thoroughbred racehorse registration were used in this study: AHT4 (Y07733), AHT5 (Y07732), ASB2 (X93516), ASB17 (X93531), ASB23 (X93537), CA425 (U67406), HMS2 (X74631), HMS3 (X74632), HMS6 (X74635), HMS7 (X74636), HTG4 (AF169165), HTG10 (AF169294), LEX3 (AF075607), LEX33 (AF075635), TKY19 (AB048330), TKY28 (AB048335), TKY321 (AB034629), and VHL20 (X75970). For amplification of these markers, we performed multiplex PCR based on the procedure described by Kakoi *et al.* [12]. The resulting PCR products were electrophoresed using a 3130xl Genetic Analyzer (Thermo Fisher Scientific), and analyses were performed with the GeneMapper Software® (Thermo Fisher Scientific).

For genotyping the grey coat colour gene (*STX17* gene), we employed the method reported by Rosengren Pielberg *et al.* [19], with minor modifications. Genomic DNA (2.0 μl, approximately 50 ng/μl) was mixed with 10 μl of 5× Buffer, 2.5 μl of DMSO, 2.5 μl of 10 mM dNTPs, 0.375 μl of 100 μM DupForward primer (GGAACATAAAGTAGATTTG-GTGGGAAAG), 0.250 μl of 100 μM DupReverse-N primer

(TTCTGATAAATGCATAAACCCACGTAAC), 0.50 μl of 100 μM DupReverse-D primer (TTCCAATTCTGAGA-TTTTGCATTTCTAA), and 1.0 μl of Taq polymerase (5.0 units/μl, Expand Long Range dNTPack, Roche Diagnostics, Rotkreuz, Switzerland) in a total volume of 50 μl. PCR was performed using a SimpliAmp Thermal Cycler (Thermo Fisher Scientific) under the following conditions: initial denaturation for 2 min at 92°C, followed by 35 cycles of denaturation for 10 sec at 92°C, annealing for 15 sec at 55°C, and extension for 6 min at 68°C. After final extension for 7 min at 68°C, PCR products were visualized using electrophoresis on a 1.0% agarose gel.

Whole genome resequencing

For library preparation, 200 ng of genomic DNA extracted from F_2010_1 by silica beads method was used for whole-genome resequencing (WGR). A genomic library (550-bp) for WGR was prepared using a TruSeq Nano DNA Library Prep Kit (Illumina, Inc., San Diego, CA, U.S.A.), according to the manufacturer's recommendations. Paired-end sequencing (150–150 bp) was carried out on a NovaSeq 6000 sequencing platform (Illumina, Inc.) by Macrogen Japan (Sakyo, Kyoto, Japan).

SNPs and INDELs were detected by mapping to the horse reference genome sequence EquCab3.0 (GenBank assembly accession: GCA_002863925.1, 2,506,966,135 bp) using the Reseq analysis pipeline (Amelieff Corp., Tokyo, Japan), which employed the QCleaner software (Amelieff Corp.), Burrows-Wheeler Aligner (BWA, version 0.7.17) [14], Picard (version 2.13.2; https://sourceforge.net/apps/medi-awiki/picard/), GATK (version 4.0.8.1; https://software.broadinstitute.org/gatk/best-practices/), and SnpEff (version v4_0) [16]. Parameters for all analyses are described in our previous study [23].

Results

Genomic DNA extracted from samples stored less than one month

We compared the quantity and quality (indicated by GQS) of genomic DNA from hair root and whole blood samples extracted with the spin column and silica beads methods (Table 1). The DNA extracted from hair roots stored at room temperature with spin columns was consistently more degraded than the high molecular weight genomic DNA extracted from whole blood stored at −30°C.

The average yield and GQS of DNA extracted from whole blood using spin columns were 6952 ng and 4.49, respectively. On the other hand, the average yield and GQS of genomic DNA extracted from hair roots stored at room temperature with the same method were 1,701 ng and 3.46, respectively. The quality of DNA extracted from hair roots

was improved with the use of the silica beads extraction method, enabling extraction of only high molecular weight genomic DNA (average yield and GQS of 900 ng and 4.48, respectively).

Thus, the genomic DNA extraction method using silica beads was preferable for extracting only high molecular weight genomic DNA from hair roots.

Genomic DNA extracted from hair roots stored for a long term at room temperature

Similarly, the quantity and quality of genomic DNA extracted from hair root samples and stored for a long term were compared between the two extraction methods (Table 2). The average yield and GQS of genomic DNA extracted from hair roots stored at room temperature for

Table 1. Quality and quantity of genomic DNA extracted from hair roots and whole blood stored for one month

Name	Sample	Breed	Storage condition		GQS		Extract (ng)	
			Temperature	Period	Columns	Beads	Columns	Beads
R_H_TB_1	Hair	TB	RT	1 month	3.35	4.54	2,120	1,050
R_H_TB_2	Hair	TB	RT	1 month	3.30	4.39	1,548	536
R_H_TB_3	Hair	TB	RT	1 month	3.51	4.44	1,852	1,080
R_H_TB_4	Hair	TB	RT	1 month	3.54	4.49	1,416	774
R_H_TB_5	Hair	TB	RT	1 month	3.59	4.52	1,568	1,060
Mean					3.46	4.48	1,701	900
B_TB_1	Blood	TB	−30°C	1 month	4.45	-	5,960	-
B_TB_2	Blood	TB	−30°C	1 month	4.34	-	5,920	-
B_TB_3	Blood	TB	−30°C	1 month	4.67	-	8,080	-
B_TB_4	Blood	TB	−30°C	1 month	4.50	-	6,480	-
B_TB_5	Blood	TB	−30°C	1 month	4.50	-	8,320	-
Mean					4.49	-	6,952	-

TB, Thoroughbred; RT, room temperature, GQS, genomic DNA quality score (0, low quality, to 5, high quality).

Table 2. Quality and quantity of genomic DNA extracted from hair roots stored for a long term (>4 years)

Name	Sample	Breed	Storage condition		GQS		Extract (ng)	
			Temperature	Period	Columns	Beads	Columns	Beads
F_2010_1	Hair	Draft	4–8°C	10 years	3.03	4.66	1,948	448
F_2010_2	Hair	Draft	4–8°C	10 years	3.09	4.34	3,620	1,030
F_2010_3	Hair	Draft	4–8°C	10 years	3.14	4.57	2,020	834
F_2012_1	Hair	Unknown	4–8°C	8 years	2.80	4.50	2,740	740
F_2015_1	Hair	Unknown	4–8°C	5 years	2.95	4.20	1,120	328
F_2015_2	Hair	Unknown	4–8°C	5 years	3.13	4.45	1,552	418
F_2017_1	Hair	Unknown	4–8°C	3 years	2.71	4.62	1,104	618
Mean					2.98	4.48	2,015	631
R_2009_1	Hair	Native	RT	11 years	1.53	2.29	1,260	578
R_2009_2	Hair	Native	RT	11 years	1.40	2.55	476	282
R_2009_3	Hair	Native	RT	11 years	1.37	2.09	816	398
R_2010_1	Hair	Native	RT	10 years	1.34	2.17	1,020	314
R_2010_2	Hair	Native	RT	10 years	1.52	2.38	880	482
R_2010_3	Hair	Native	RT	10 years	1.60	2.31	796	370
R_2011_1	Hair	Native	RT	9 years	1.30	2.14	1,044	384
R_2011_2	Hair	Native	RT	9 years	1.75	2.57	828	258
R_2011_3	Hair	Native	RT	9 years	1.21	2.11	524	268
R_2016_1	Hair	Native	RT	4 years	2.64	3.21	1,372	434
R_2016_2	Hair	Native	RT	4 years	2.75	3.59	1,448	644
R_2016_3	Hair	Native	RT	4 years	2.75	3.23	1,364	478
Mean					1.76	2.55	986	408

Draft, Japanese heavy draft; Native, Japanese native breed; RT, room temperature, GQS, genomic DNA quality score (0, low quality, to 5, high quality).

4–11 years by spin columns were 986 ng and 1.76, respectively. Gel electrophoresis confirmed that extracts stored for longer than 10 years had degraded DNA fragments shorter than 3 kbp, while extracts stored for 4 years contained genomic DNA fragments ranging between 5 and 10 kbp (Fig. 1A).

When the same samples were extracted with silica beads, the average yield was 408 ng and GQS was 2.55. Gel electrophoresis confirmed that the conditions of these extracts were similar to those extracted with spin columns, although less genomic DNA was fragmented to <500 kbp (Fig. 1B). Overall, extraction of only high molecular weight genomic DNA from hair roots stored for a long term at room temperature was difficult with both methods, indicating that the DNA had degraded.

Genomic DNA extracted from hair roots stored for a long term at 4–8°C

Since DNA degradation is known to progress during long-term storage, the study also examined the effect of storage temperature on DNA quality. Table 2 shows the yield and GQS of genomic DNA extracted from hair roots stored at 4–8°C with spin columns or the silica beads method.

The average yield and GQS of genomic DNA extracted from these samples with spin columns were 2,015 ng and 2.98, respectively. The qualities of the genomic DNA extracted from hair roots stored for 10, 8, 5, and 3 years at 4–8°C did not appear to differ significantly when examined on agarose gel (Fig. 1A). All extracts exhibited high molecular weight genomic DNA as well as some fragmented DNA.

The average yield and GQS score of genomic DNA extracted from hair roots stored at 4–8°C for 3–10 years by silica beads were 631 ng and 4.48, respectively. The quality
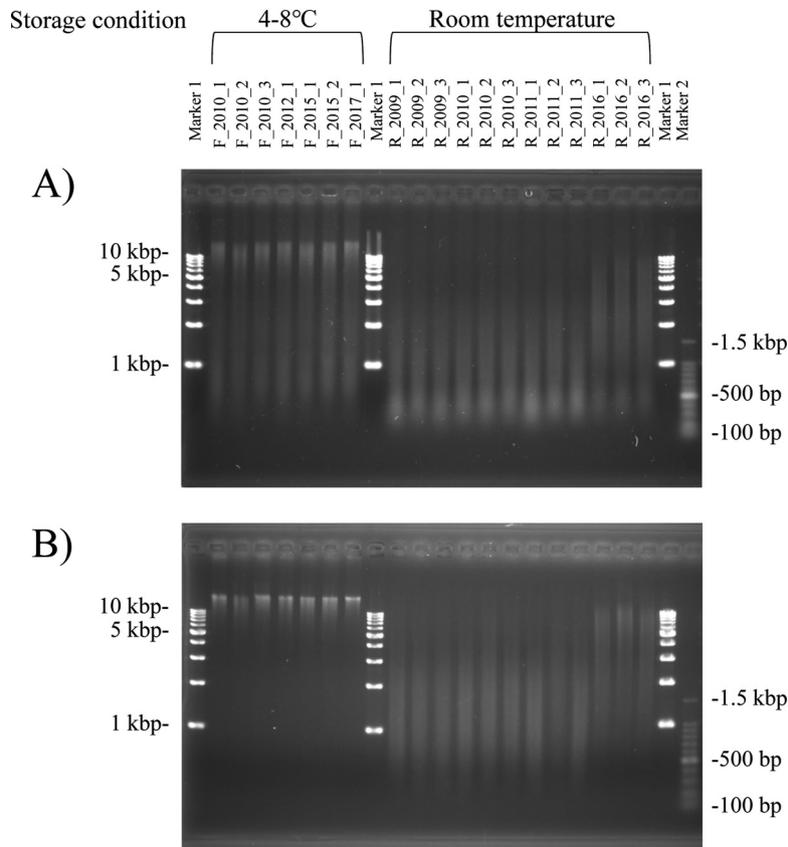
Fig. 1. One percent agarose gel showing genomic DNA extracted from hair roots stored for a long term. Genomic DNA was extracted with spin columns (A) or silica beads (B) methods. F_2010_1, F_2010_2, and F_2010_3 were stored for 10 years at 4–8°C; F_2012_1 was stored for 8 years at 4–8°C; F_2015_1 and F_2015_2 were stored for 5 years at 4–8°C; and F_2017_1 was stored for 3 years at 4–8°C. R_2009_1, R_2009_2. R_2009_3 was stored for 11 years at room temperature; R_2010_1, R_2010_2, and R_2010_3 were stored for 10 years at room temperature; R_2011_1, R_2011_2, and R_2011_3 were stored for 9 years at room temperature; and R_2016_1, R_2016_2, and R_2016_3 were stored for 4 years at room temperature.

value was comparable to that of genomic DNA extracted from whole blood. Gel electrophoresis revealed that only high molecular weight genomic DNA was extracted (Fig. 1B).

The spin columns method was successful in extracting high molecular weight genomic DNA from hair roots stored for a long term at 4–8°C; however, degraded DNA was also observed. Conversely, the silica beads method allowed extraction of only high molecular weight genomic DNA from the same samples. This trend was also observed on the extraction of DNA from similar samples stored for less than one month, indicating that refrigerated storage likely prevents the degradation of genomic DNA in hair roots.

PCR amplification

The quality of genomic DNA extracted from hair roots was also evaluated by PCR amplification of the standard panel of microsatellite markers for parentage verification in Thoroughbred racehorse registration. PCR amplification

and genotyping were successfully conducted for hair roots stored both at room temperature and 4–8°C and extracted by either method. No discrepancies were observed in the relationship between the foal sample (F_2010_1), its sire (F_2010_2), and its dam (F_2010_3). The same genotypes were identified for F_2015_1 and F_2015_2 no matter which method of extraction was used. Genotyping was possible even with degraded DNA because the PCR products of all microsatellite markers were less than 500 bp.

Amplification of the *STX17* 'grey' gene mutation (approximately 5-kbp DNA fragment) was confirmed only in DNA extracted from hair roots stored at 4–8°C by both the spin columns (Fig. 2A) and silica beads (Fig. 2B) methods. DNA amplification of this larger fragment was possible in refrigerated samples which still contained a large proportion of intact DNA. The samples stored at room temperature appeared to be too degraded for the long-range PCR to work. Therefore, these results suggest that extracts from hair roots stored at room temperature are only suitable
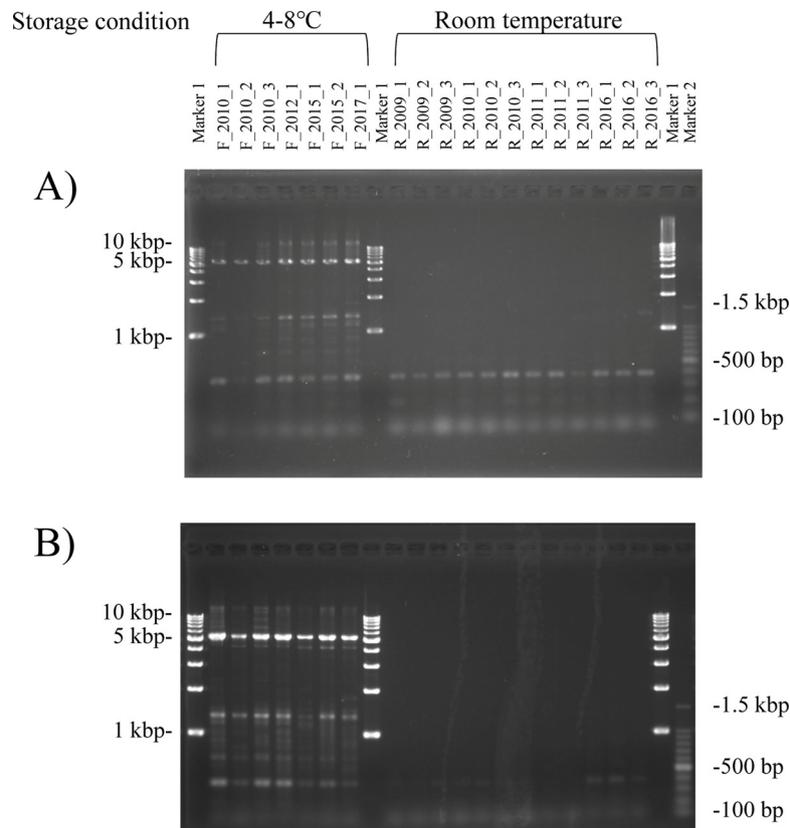
Fig. 2. Amplification of the *STX17* gene using genomic DNA extracted from hair roots stored for a long term. Genomic DNA was extracted with spin columns (A) or silica beads (B) methods. F_2010_1, F_2010_2, and F_2010_3 were stored for 10 years at 4–8°C; F_2012_1 was stored for 8 years at 4–8°C; F_2015_1 and F_2015_2 were stored for 5 years at 4–8°C; and F_2017_1 was stored for 3 years at 4–8°C. R_2009_1, R_2009_2. R_2009_3 was stored for 11 years at room temperature; R_2010_1, R_2010_2, and R_2010_3 were stored for 10 years at room temperature; R_2011_1, R_2011_2, and R_2011_3 were stored for 9 years at room temperature; and R_2016_1, R_2016_2, and R_2016_3 were stored for 4 years at room temperature.

for PCR amplification of smaller-sized fragments, such as microsatellite markers for parentage verification.

Whole genome resequencing

A next generation sequencing library was prepared and sequenced using genomic DNA extracted with the silica beads methods from hair roots refrigerated for 10 years (Fig. 1B). The number of acquired reads was 686,558,582, with 652,488,758 passing quality control and being mapped to the reference genome (Table 3). The available reads were similar to those in our previous study [23]. Within the single horse that was sequenced, 6,064,956 SNPs and 727,565 INDELs were detected (Table 3).

The sample used for WGR analysis originated from a horse with a white coat due to documented variation in the *KIT* gene. This allele was named W17, and it consists of two SNPs which change two amino acids in exon 14 of the gene [9]. These mutations were identified in the previous study by Sanger sequencing of PCR products using DNA extracted from a blood sample. The nucleotide sequence obtained by WGR in the current study was identical to that reported in the previous study. The two mutations were observed on the same reads in this study, indicating that they occurred on the same chromosome.

Discussion

In this study, we demonstrated that high molecular weight genomic DNA could be extracted from horsehair roots in refrigerated storage for up to 10 years using a magnetic silica beads method. The extracted genomic DNA was suitable for WGR library preparation and next generation sequencing. This provides a new option for the long-term storage of samples for gene doping control in horse sports. For the many jurisdictions that currently use hair samples

for parentage verification, collecting and storing blood samples for future gene doping control would be difficult to sell to stakeholders. Blood samples need to be collected by a veterinarian and should be chilled and transported to the laboratory before storage in a freezer, which is expensive. The alternative of using hair samples already collected for parentage verification minimizes the number of samples taken from horses and is preferable for their welfare. This study shows that authorities and laboratories now have the option of at least two different types of samples for gene doping detection based on WGR.

To our knowledge, this is the first report detailing WGR using hair roots from horses stored for a long term as a model study for gene doping control. A limitation of the study is that extraction only succeeded with a combination of Proteinase K (Qiagen) digestion and magnetic silica beads (Toyobo) extraction. Other extraction kits were not suitable for extracting high-quality undamaged genomic DNA from hair roots.

For samples that have been kept for a long term, it may be necessary to check the conditions of the hairs and to keep them dry to prevent mold growth. It is also recommended that samples be washed before DNA extraction to avoid contamination.

It is unlikely that artificially inserted genes would be shed into hair follicles. However, manipulation of the genome performed at the embryonic level, such as that which would occur with gene editing, would change the DNA extracted from the hair root. Currently, all methods of gene editing detection in the horse focus on comparing the DNA of a foal to its parents. The discovery that hair samples could be used for WGR means that stored parent hair samples can be used for this purpose, even if the parent has since passed away. This study is therefore a valuable contribution to gene doping control in horseracing and other equestrian sports.

Table 3. Sequencing and mapping summary

	F_2010_1	C2*	C4*
Total reads	686,558,582	1,081,810,248	989,568,234
Passed reads	652,488,758	632,750,570	845,549,706
Mapped/passed reads	95.04%	99.56%	99.38%
Reads after duplicate reads removed	595,705,705	625,962,801	723,928,404
Coverage >=1	99.20%	97.30%	97.40%
Coverage >=10	97.50%	95.20%	97.10%
Coverage >=30	64.50%	64.20%	82.80%
Average depth >=1	35.90%	38.70%	44.60%
Average depth >=10	36.40%	39.40%	44.80%
Average depth >=30	42.50%	47.90%	48.20%
Identified SNPs	6,064,956	-	-
Identified INDELs	727,565	-	-

SNPs, single nucleotide polymorphisms; INDELs, insertions/deletions. *Tozaki *et al*. 2020 [23].

Acknowledgments

References

1. Ablondi, M., Viklund, Å., Lindgren, G., Eriksson, S., and Mikko, S. 2019. Signatures of selection in the genome of Swedish warmblood horses selected for sport performance. *BMC Genomics* **20**: 717. [Medline] [CrossRef]

2. Anti-Doping Rules, Fédération Equestre Internationale (FEI). https://inside.fei.org/content/anti-doping-rules [accessed on October 1, 2020].

3. Bower, M.A., McGivney, B.A., Campana, M.G., Gu, J., Andersson, L.S., Barrett, E., Davis, C.R., Mikko, S., Stock, F., Voronkova, V., Bradley, D.G., Fahey, A.G., Lindgren, G., MacHugh, D.E., Sulimova, G., and Hill, E.W. 2012. The genetic origin and history of speed in the Thoroughbred racehorse. *Nat. Commun.* **3**: 643. [Medline] [CrossRef]

4. Brzeziańska, E., Domańska, D., and Jegier, A. 2014. Gene doping in sport-perspectives and risks. *Biol. Sport* **31**: 251–259. [Medline] [CrossRef]

5. Chu, Y., Bartus, R.T., Manfredsson, F.P., Olanow, C.W., and Kordower, J.H. 2020. Long-term post-mortem studies following neurturin gene therapy in patients with advanced Parkinson's disease. *Brain* **143**: 960–975. [Medline] [CrossRef]

6. Ducro, B.J., Gorissen, B., van Eldik, P., and Back, W. 2009. Influence of foot conformation on duration of competitive life in a Dutch Warmblood horse population. *Equine Vet. J.* **41**: 144–148. [Medline] [CrossRef]

7. Dupont, J.B., Guo, J., Renaud-Gabardos, E., Poulard, K., Latournerie, V., Lawlor, M.W., Grange, R.W., Gray, J.T., Buj-Bello, A., Childers, M.K., and Mack, D.L. 2020. AAV-mediated gene transfer restores a normal muscle transcriptome in a canine model of X-Linked myotubular myopathy. *Mol. Ther.* **28**: 382–393. [Medline] [CrossRef]

8. George, L.A., Ragni, M.V., Rasko, J.E.J., Raffini, L.J., Samelson-Jones, B.J., Ozelo, M., Hazbon, M., Runowski, A.R., Wellman, J.A., Wachtel, K., Chen, Y., Anguela, X.M., Kuranda, K., Mingozzi, F., and High, K.A. 2020. Long-term follow-up of the first in human intravascular delivery of AAV for gene transfer: AAV2-hFIX16 for severe hemophilia B. *Mol. Ther.* **28**: 2073–2082. [Medline] [CrossRef]

9. Haase, B., Rieder, S., Tozaki, T., Hasegawa, T., Penedo, M.C., Jude, R., and Leeb, T. 2011. Five novel KIT mutations in horses with white coat colour phenotypes. *Anim. Genet.* **42**: 337–339. [Medline] [CrossRef]

10. Haughan, J., Jiang, Z., Stefanovski, D., Moss, K.L., Ortved, K.F., and Robinson, M.A. 2020. Detection of intra-articular gene therapy in horses using quantitative real time PCR in synovial fluid and plasma. *Drug Test. Anal.* **12**: 743–751. [Medline] [CrossRef]

11. International Agreement on Breeding Racing and Wagering (IABRW). International Federation of Horseracing Authorities (IFHA). https://www.ifhaonline.org/Default. asp?section=IABRW&area=15 [accessed on October 1, 2020].

12. Kakoi, H., Nagata, S., and Kurosawa, M. 2001. DNA typing with 17 microsatellites for parentage verification of racehorses in Japan. *Anim. Sci. J.* **72**: 453–460.

13. Kovac, M., Litvin, Y.A., Aliev, R.O., Zakirova, E.Y., Rutland, C.S., Kiyasov, A.P., and Rizvanov, A.A. 2018. Gene therapy using plasmid DNA encoding VEGF164 and FGF2 genes: a novel treatment of naturally occurring tendinitis and desmitis in horses. *Front. Pharmacol.* **9**: 978. [Medline] [CrossRef]

14. Li, H., and Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754–1760. [Medline] [CrossRef]

15. Martin, N.P., Myers, P., Goulding, E., Chen, S.H., Walker, M., Porter, T.M., Van Gorder, L., Mathew, A., Gruzdev, A., Scappini, E., and Romeo, C. 2018. Laser-assisted lentiviral gene delivery to mouse fertilized eggs. *J. Vis. Exp.* **141**: 58327. [Medline]

16. McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R., Thormann, A., Flicek, P., and Cunningham, F. 2016. The Ensembl Variant Effect Predictor. *Genome Biol.* **17**: 122. [Medline] [CrossRef]

17. Moss, K.L., Jiang, Z., Dodson, M.E., Linardi, R.L., Haughan, J., Gale, A.L., Grzybowski, C., Engiles, J.E., Stefanovski, D., Robinson, M.A., and Ortved, K.F. 2020. Sustained interleukin-10 transgene expression following intra-articular AAV5-IL-10 administration to horses. *Hum. Gene Ther.* **31**: 110–118. [Medline] [CrossRef]

18. Neuberger, E.W.I., and Simon, P. 2017. Gene and cell doping: the new frontier−beyond myth or reality. *Med. Sport Sci.* **62**: 91–106. [Medline] [CrossRef]

19. Rosengren Pielberg, G., Golovko, A., Sundström, E., Curik, I., Lennartsson, J., Seltenhammer, M.H., Druml, T., Binns, M., Fitzsimmons, C., Lindgren, G., Sandberg, K., Baumung, R., Vetterlein, M., Strömberg, S., Grabherr, M., Wade, C., Lindblad-Toh, K., Pontén, F., Heldin, C.H., Sölkner, J., and Andersson, L. 2008. A cis-acting regulatory mutation causes premature hair graying and susceptibility to melanoma in the horse. *Nat. Genet.* **40**: 1004–1009. [Medline] [CrossRef]

20. Tozaki, T., Kakoi, H., Mashima, S., Hirota, K., Hasegawa, T., Ishida, N., Miura, N., Choi-Miura, N.H., and Tomita, M. 2001. Population study and validation of paternity testing for Thoroughbred horses by 15 microsatellite loci. *J.*

Vet. Med. Sci. **63**: 1191–1197. [Medline] [CrossRef]

21. Tozaki, T., Ohnuma, A., Kikuchi, M., Ishige, T., Kakoi, H., Hirota, K.I., Kusano, K., and Nagata, S.I. 2020. Microfluidic quantitative PCR detection of 12 transgenes from horse plasma for gene doping control. *Genes (Basel)* **11**: 457. [Medline] [CrossRef]

22. Tozaki, T., Ohnuma, A., Takasu, M., Kikuchi, M., Kakoi, H., Hirota, K.I., Kusano, K., and Nagata, S.I. 2019. Droplet digital PCR detection of the erythropoietin transgene from horse plasma and urine for gene-doping control. *Genes (Basel)* **10**: 243. [Medline] [CrossRef]

23. Tozaki, T., Ohnuma, A., Takasu, M., Nakamura, K., Kikuchi, M., Ishige, T., Kakoi, H., Hirora, K.I., Tamura, N., Kusano, K., and Nagata, S.I. 2020. Detection of non-targeted transgenes by whole-genome resequencing for gene-doping control. *Gene Ther.* (in press). [Medline] [CrossRef]

24. Zhang, J., Liu, J., Yang, W., Cui, M., Dai, B., Dong, Y., Yang, J., Zhang, X., Liu, D., Liang, H., and Cang, M. 2019. Comparison of gene editing efficiencies of CRISPR/Cas9 and TALEN for generation of MSTN knock-out cashmere goats. *Theriogenology* **132**: 1–11. [Medline] [CrossRef]