



Copy number variations of five Y chromosome genes in donkeys

Haoyuan Han, Xiaocheng Zhao, Xiaoting Xia, Hong Chen, Chuzhao Lei, and Ruihua Dang

College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi 712100, China

Correspondence to: Chuzhao Lei (leichuzhao1118@126.com)

Received: 21 July 2017 – Revised: 11 September 2017 – Accepted: 14 September 2017 – Published: 18 October 2017

Abstract. In mammals, the Y chromosome plays a pivotal role in male sex determination and is essential for normal sperm production. A number of studies were conducted on Y chromosome genes of various species and identified single-copy and multi-copy genes. However, limited studies about donkey Y chromosome genes have been done. In this study, 263 male samples from 13 Chinese donkey breeds were collected to analyze the copy number variations (CNVs) of five Y chromosome genes using the quantitative PCR (qPCR) method. These five genes (cullin 4 B Y (*CUL4BY*), equus testis-specific transcript y1 (*ETSTY1*), equus testis-specific transcript y4 (*ETSTY4*), equus testis-specific transcript Y 5 (*ETSTY5*), and sex-determining region Y (*SRY*) were identified as multi-copy, whose median copy numbers (MCNs) were 5, 45, 2, and 2, and 13 with CNV ranges of 1–57, 1–227, 1–37, 1–86 and 1–152, respectively. The CNVs of these five genes were shared in different breeds. Compared to previous studies, the copy numbers of five genes showed some distinct consequences in this study. In particular, the well-known single-copy *SRY* gene showed CNVs in donkeys. Our results provided genetic variations of donkey Y chromosome genes.

1 Introduction

As a consequence of 230–300 million years of independent and non-recombining evolution, the majority (95 %) of the present-day Y chromosome is a male-specific region (MSY) that does not recombine with the X chromosome during meiosis to undergo homologous recombination (Graves, 1998; Quintana-Murci and Fellous, 2011). The MSY is enriched with multi-copy genes and copy number variations (CNVs) (Jobling, 2008; Skaletsky et al., 2003). CNVs are DNA segments of at least 50 bp in size that vary in copy number among individuals of a species (Mills et al., 2011). In many cases, CNV has been linked to gene dosage, gene regulation, gene expression, and phenotype (Cahan et al., 2009; Henrichsen et al., 2009; Redon et al., 2006; Toft and Ross, 2010; Zhang et al., 2009).

Y chromosome is notoriously difficult to be sequenced because of its distinctive and structurally repetitive genomic landscapes. Thus, studies on CNV do not include an analysis of variation on the Y chromosome generally and only focus on a few Y-linked genes (Skaletsky et al., 2003). The testis-specific protein Y-encoded (*TSPY*) gene shows 20–76 copies

and the CNVs have been found to be associated with semen quality and reproduction by regulating cell division in the process of spermatogenesis in men (Nickkholgh et al., 2010; Vodicka et al., 2007). The *TSPY* gene shows 37–200 copies on the bovine Y chromosome (BTAY) (Hamilton et al., 2009; Jakubiczka et al., 1993). The heat-shock transcription factor, Y-linked (*HSFY*) gene contains around 70 copies in bulls (Hamilton et al., 2011). The data from Yue et al. (2013) confirmed that the copy number of the preferentially expressed antigen in melanoma, Y-linked (*PRAMEY*) gene family was highly variable, ranging from 2 to 31 across 15 cattle breeds, and plays an important role in spermatogenesis.

The sex-determining region Y (*SRY*) gene is known as the sex determination gene in mammalian development (Wilhelm et al., 2007). Thus, it is crucial to make clear of its CNV for a better understanding of *SRY* gene in donkeys. The testis-specific transcript Y (*ETSTY*) family is identified to consist of equine-specific genes (Paria et al., 2011); three of them (*ETSTY1*, *ETSTY4*, and *ETSTY5*) were chosen to analyze the CNVs in this study. The copy number of the cullin 4 B Y (*CUL4BY*) gene has been detected in other species, such as

cats and dogs (Murphy et al., 2006; Pearks Wilkerson et al., 2008). The *CUL4BY* gene was involved in this study in order to determine its copy numbers in donkeys. Copy numbers reveal the basic genetic and structural information of Y chromosome genes. Because of the lack of assembling Y chromosome sequences in donkeys, it is difficult to analyze the CNVs of those genes at donkey Y chromosome by re-sequencing. As such, the previous study determining copy numbers on donkey Y chromosome genes was limited.

Population and genetic sources of Chinese native donkeys are abundant. There are 24 Chinese native donkey breeds, and all of which have important functions in transportation and tourism as well as in the production of meat, skin, and milk. In particular, Dezhou donkeys are famous for their skins, which are great materials for donkey-hide gelatin, and Guanzhong donkeys are famous for improving the body size of small donkeys and producing large mules. However, until now, few analyses of CNVs of donkey Y-linked genes have been reported. In this study, the quantitative PCR (qPCR) method was performed to determine the copy number variations of five Y chromosome genes (*CUL4BY*, *ETSTY1*, *ETSTY4*, *ETSTY5*, and *SRY*) in 263 males from 13 Chinese donkey breeds, which would provide novel genetic variations of donkey genome.

2 Materials and methods

2.1 Sampling

Totaling 263 blood and ear tissue samples were obtained from 13 Chinese donkey breeds (Table 1). One additional female individual was used as female control in order to affirm the male specificity of the horse Y chromosome primers in donkeys. Genomic DNA was extracted from blood using a standard phenol–chloroform method (Sambrook and Russell, 2002). The DNA concentration was diluted to $20 \text{ ng } \mu\text{L}^{-1}$ by ultrapure water and stored at -20° .

2.2 Primer design

Primers of the *beta-actin ACTB* gene were designed using primer 5.0. The male-specific primers of five Y chromosome genes – *CUL4BY*, *ETSTY1*, *ETSTY4*, *ETSTY5*, and *SRY* – were acquired from Paria et al. (2011). Details of primers are shown in Table 2. Two males, one female control, and one negative control were amplified using a routine PCR to identify the male specificity of five pairs of primers in donkeys. The PCR protocol was as follows: each $12.5 \mu\text{L}$ reaction contained 20 ng of pooled genomic DNA, 5 pM each primer ($10 \text{ pmol } \mu\text{L}^{-1}$), $6.25 \mu\text{L}$ of $2 \times \text{PCR mix buffer}$ (including 0.375 U Taq DNA polymerase, $2 \times \text{PCR buffer}$, $18.75 \mu\text{M}$ MgCl_2 and $2.5 \mu\text{M}$ dNTPs), and $4.75 \mu\text{L}$ of distilled water. Thermocycling consisted of an initial denaturation at 95°C for 5 min , followed by 35 cycles of at 94°C for 30 s , at annealing temperature (Table 2) for 40 s , at 72°C for 30 s , and

a final extension at 72°C for 10 min , before finally storing at 4°C .

2.3 Copy number variations (CNVs) estimation

Standard curves and CNVs for each gene were measured by qPCR using Roche Lightcycler 480 system. qPCR was performed in plates with 96 wells, on which a calibrator, a negative control (distilled water), and diluted DNA (60 , 40 , 20 , 10 , 5 , 2.5 , and $1.25 \text{ ng } \mu\text{L}^{-1}$ for standard curves) or testing samples ($5 \text{ ng } \mu\text{L}^{-1}$ for CNV estimation) were set up. qPCR reactions with DNA samples, calibrator, and negative control were run in triplicate. Each reaction contained $10 \mu\text{L}$ of SYBR PCR master mix (TaKaRa, Dalian, China), $0.8 \mu\text{L}$ of primers ($10 \text{ pmol } \mu\text{L}^{-1}$), $6.8 \mu\text{L}$ of distilled water, and $1.6 \mu\text{L}$ of DNA template ($5 \text{ ng } \mu\text{L}^{-1}$). The qPCR was run with a program of the following steps: predenaturation at 95°C for 10 min , followed by 40 cycles of denaturation at 95°C for 5 s and annealing at appropriate temperature (Table 2) for 30 s . A melting curve was then generated by taking fluorescent measurements every 0.11°C from 60 until 95°C .

The cycle threshold (C_T) values generated from DNA of graded concentrations for reference (*ACTB*) and targeted genes (*CUL4BY*, *ETSTY1*, *ETSTY4*, *ETSTY5*, and *SRY*) were used to calculate the standard curves and their linear regression equations, the slopes of which were applied to measure primer efficiencies according to the equation $E = 10^{(-1/\text{slope})}$. Then, relying on additional three equations (listed as follows) described in Hamilton et al. (2009), the CNVs of *CUL4BY*, *ETSTY1*, *ETSTY4*, *ETSTY5*, and *SRY* were estimated. In the equations, the C_T values of the calibrator were determined by the average of all C_T values obtained in different plates for each gene, and $\Delta C_T = C_T$ of the calibrator – C_T of the test sample.

$$\text{Copy number}_{\text{calibrator}} = \frac{(E_{\text{reference}})^{C_{T\text{reference}}}}{(E_{\text{target}})^{C_{T\text{target}}}}$$

$$\text{Ratio} = \frac{(E_{\text{target}})^{\Delta C_{T\text{target}}(\text{calibrator-sample})}}{(E_{\text{reference}})^{\Delta C_{T\text{reference}}(\text{calibrator-sample})}}$$

$$\text{Copy number}_{\text{test sample}} = (\text{copy number}_{\text{calibrator}}) \times (\text{ratio}) \times 2$$

2.4 Statistical analysis

In order to minimize technical error and to obtain an accurate copy number (CN) estimation, raw qPCR data showing a coefficient of variation (CV) $> 1 \%$ between the duplicates were excluded from further analysis. The Kolmogorov–Smirnov and Shapiro–Wilk normality tests were used to analyze the normality of the CN data (Justel Peña and Zamar, 1997; Shapiro and Wilk, 1965). Box plot analysis of the CN data was conducted to detect the outliers in all the breeds as a whole. Multiple pair-wise comparisons of median copy

Table 1. Sample information of 13 donkey breeds in China.

Breed	Abbreviation	Sample size	Collection location
Qinghai donkey	QH	30	Minhe county, Qinghai Province
Gunsha donkey	GS	25	Yulin city, Shaanxi Province
Huaibei donkey	HL	6	Huaibei city, Anhui Province
Jiami donkey	JM	19	Mizhi county, Shaanxi Province
Liangzhou donkey	LZ	11	Wuwei city, Gansu Province
Kulun donkey	KL	6	Kulun county, Inner Mongolia Region
Qingyang donkey	QY	16	Qingyang city, Gansu Province
Guanzhong donkey	GU	9	Fufeng county, Shaanxi Province
Dezhou donkey	DZ	60	Dezhou city, Shandong Province
Biyang donkey	BY	39	Biyang county, Henan Province
Xiji donkey	XIJ	16	Xiji county, Ningxia Region
Xinjiang donkey	XJ	21	Yining city, Xinjiang Region
Changyuan donkey	CY	5	Changyuan county, Henan Province
Total		263	

Table 2. QPCR primer information for the *ACTB* gene and five Y chromosome genes.

Gene		Primer F	Primer R	T_m (°)	Length (bp)	Accession number	Reference
<i>ACTB</i>		CACGCCTTTTGCTCAC TTGTC	ACGGGCTTTGTCAC ACGAG	60–65	203	NC_009156	this study
<i>CUL4BY</i>	cullin 4 B Y	TGTGGGGTTCGTGTG AAATA	CAAGGATCGCTGGGT CTTAC	60	172	EU687546	Paria et al. (2011)
<i>ETSTY1</i>	equus testis-specific transcript Y1	GACGGACGACCTTGT GTTTT	CTAGTGCGAGTCCT TTTGG	63	234	EU687549	Paria et al. (2011)
<i>ETSTY4</i>	equus testis-specific transcript Y4	GCGTCTGTGCAGATG TGTCT	GCTCATGCAGTCAAA CAGGA	65	175	EU687552	Paria et al. (2011)
<i>ETSTY5</i>	equus testis-specific transcript Y5	CAAAACCAAGAGGA GGACCA	CTCCAGAGGCAGGT ACTTCG	65	210	EU687553	Paria et al. (2011)
<i>SRY</i>	sex-determining region Y	CCAACGCTTTATCTT CGCA	CCATTTCCTCTTGTT TCACTCC	63	203	NM_001081810	this study

number (MCN) between breeds were analyzed using a non-parametric Mann–Whitney *U* test (Mann and Whitney, 1947) with a Bonferroni correction (Dunn, 1961).

3 Results and discussion

3.1 Male specificity of primers

The donkey Y chromosome sequence is not available. In an investigation of the equid Y chromosome, the horse and donkey MSY gene contents were compared. Comparative analysis of horse and donkey MSY revealed that the *CUL4BY*, *ETSTY1*, *ETSTY4*, *ETSTY5*, and *SRY* genes are shared in the genome of these two species; the sequence identity was 95–100 % between the homologous horse and donkey transcripts

according to Paria et al. (2011), who also verified that the primers for these five genes were male-specific in both horses and donkeys. In this study, the results of DNA fragments amplified with two males, one female control, and one negative control are shown in Fig. 1; *ACTB* showed a single band of expected size in male and female samples. The *ETSTY4* and *SRY* genes displayed a male-specific band. The *CUL4BY*, *ETSTY1*, and *ETSTY5* genes showed different sized fragments in females. However, the amplified band for each gene was specific and correctly sized in males; thus, it is reasonable to use the primers for our CNVs analysis.

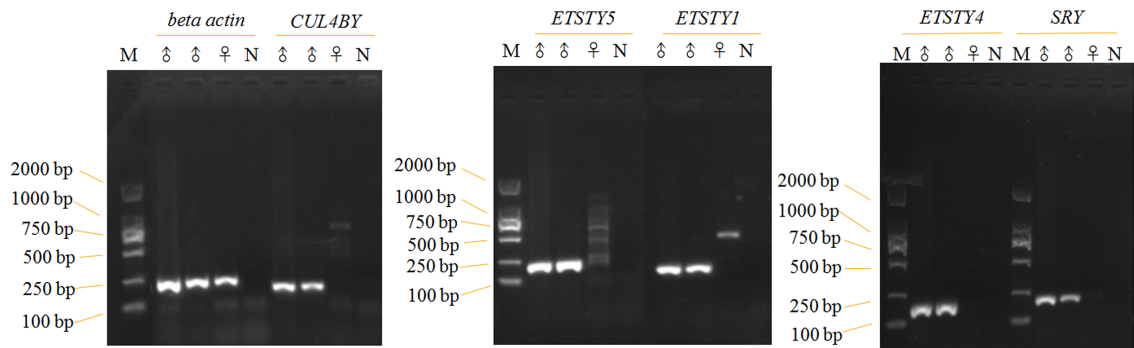


Figure 1. Gel electrophoresis of PCR products of six donkey genes. The primers of these five Y-linked genes amplified male-specific bands with expected fragment size labeled above the band. M: 2 kb DNA ladder; ♂: male donkey genomic DNA; ♀: female donkey genomic DNA; N: negative control (distilled water).

Table 3. Correlation coefficient of standard curve and primer efficiency for *ACTB* and five Y chromosome genes.

Gene	Correlation coefficient (R^2)	Primer efficiency (E)
<i>ACTB</i>	0.9995	2.14
<i>ETSTY1</i>	0.9913	1.90
<i>SRY</i>	0.9933	1.93
<i>ACTB</i>	0.9973	1.90
<i>ETSTY4</i>	0.9976	2.01
<i>ETSTY5</i>	0.9941	1.97
<i>ACTB</i>	0.9915	1.92
<i>CUL4BY</i>	0.9961	1.98

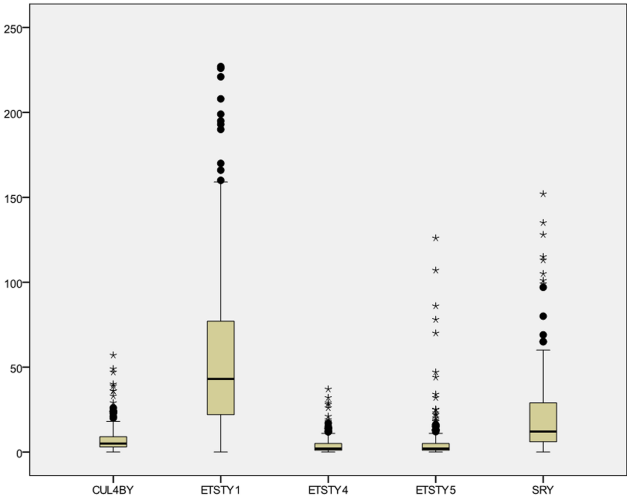


Figure 2. Box plot analysis of five genes' CNs in donkeys. The outliers are indicated by a solid circle or an asterisk (extremely high CN), which include 10–15 donkeys in all breeds for five Y chromosome genes.

3.2 The CNVs of five Y chromosome genes in donkeys

According to the linear regression equations of standard curves and the equation $E = 10^{(-1/\text{slope})}$, correlation coefficients (R^2) were all higher than 0.99 and primer efficiencies (E) were higher than 1.90 for the *ACTB*, *CUL4BY*, *ETSTY1*, *ETSTY4*, *ETSTY5*, and *SRY* genes (Table 3). In total, the CNVs of 263 males from 13 Chinese donkey breeds were estimated. The *CUL4BY*, *ETSTY1*, *ETSTY4*, *ETSTY5*, and *SRY* were all multi-copy with median copy numbers (MCNs) of 5, 45, 2, 2, and 13, respectively. The CNVs ranges were 1–57, 1–227, 1–37, 1–86, and 1–152, respectively (Table 4). Based on Kolmogorov–Smirnov and Shapiro–Wilk normality tests, the copy numbers (CNs) of the five donkey Y chromosome genes did not fit the normal distribution in Chinese horses ($P < 0.0001$). Box plot analyses of the CNs data revealed that approximately 10–15 individuals (4–6 %) were outliers that had significantly higher CN for each gene (Fig. 2).

For the *CUL4BY* gene, the CNVs ranged from 1 to 57 with MCNs of 2–9 in 13 donkey breeds, while Guanzhong (GZ) donkeys had the highest MCN of 9 (Table 4). Dezhou donkeys showed significantly different MCN from nine other breeds for the *CUL4BY* gene (Table S1 in the Supplement).

Similarly, the highly amplified *CUL4BY* gene has been identified in the cats and dogs. (Murphy et al., 2006; Pearks Wilkerson et al., 2008). The horse *CUL4BY* showed intermediate expression being transcribed in testis, heart, spleen, and kidney (Paria et al., 2011); thus, the horse *CUL4BY* might be related to testis and male germ cell proliferation (Petroski and Deshaies, 2005). Considering the close relationship of donkeys and horses, we assumed that the donkey *CUL4BY* had an important function in male fertility.

The *ETSTY* family is novel and equid-specific, the orthologue of which is not found in non-equid species. The range of CNVs of the *ETSTY1* gene was 1–227, with the MCNs varying from 12 (GZ donkey) to 99 (Changyuan donkey) in different breeds (Table 4), indicating that the *ETSTY1* gene was a largely amplified gene with multiple copies. GZ donkeys exhibited significant difference of MCNs from 10 other

Table 4. Copy number variations of five Y chromosome genes in donkeys.

Breed	No. of samples	Median copy number (copy number variation)				
		<i>CUL4BY</i>	<i>ETSTY1</i>	<i>ETSTY4</i>	<i>ETSTY5</i>	<i>SRY</i>
QH	30	3 (1–14)	64 (11–166)	3 (1–14)	2 (1–20)	12 (1–42)
GS	25	2 (1–5)	19 (4–90)	1 (1–21)	1 (1–18)	4 (1–33)
HL	6	2 (1–3)	17 (14–59)	1 (1–2)	2 (1–10)	10 (6–38)
JM	19	7 (4–49)	61 (24–226)	5 (2–8)	25 (3–86)	63 (11–152)
LZ	11	4 (2–6)	65 (25–195)	1 (1–7)	1 (1–10)	8 (7–21)
KL	6	2 (1–4)	51 (16–94)	1	1 (1–5)	2 (1–11)
QY	16	4 (1–11)	70 (23–193)	3 (1–6)	5 (1–18)	12 (4–59)
GZ	9	9 (4–13)	12 (1–44)	2 (2–19)	3 (1–5)	30 (4–97)
DZ	60	8 (2–47)	29 (5–142)	2 (1–37)	2 (1–31)	7 (1–101)
BY	39	8 (2–57)	57 (2–227)	4 (1–29)	2 (1–12)	33 (1–65)
GY	16	3 (2–11)	55 (32–162)	2 (1–3)	2 (1–9)	17 (6–34)
XJ	21	2 (1–36)	50 (6–102)	1 (1–5)	3 (1–16)	11 (4–51)
CY	5	3 (3–5)	99 (73–163)	3 (2–4)	1	13 (11–20)
	263	5 (1–57)	45 (1–227)	2 (1–37)	2 (1–86)	13 (1–152)

Abbreviations: QH, Qinghai donkey; GS, Gunsha donkey; HL, Huaibei donkey; JM, Jiami donkey; LZ, Liangzhou donkey; KL, Kulun donkey; QY, Qingyang donkey; GU, Guanzhong donkey; DZ, Dezhou donkey; BY, Biyang donkey; XJ, Xiji donkey; XJ, Xinjiang donkey; CY, Chuangyuan donkey.

breeds (Table S2). For the *ETSTY4* gene, Jiami (JM) donkeys possessed the highest MCN of 5 (Table 4). The MCN of JM donkeys was significantly different from that of 10 other breeds (Table S3). The highest MCN of the *ETSTY5* gene was found in the JM breed (MCN of 25), while the MCNs only varied from 1 to 5 in 12 other breeds (Table 4). The comparisons of the MCN of the *ETSTY5* gene showed that JM donkeys exhibited different MCNs from 12 other breeds (Table S4). *ETSTY1*, *ETSTY4*, and *ETSTY5* showed multiple copies according to cDNA FISH (Paria et al., 2011), the CNVs of which were confirmed in this study. These genes were expressed exclusively in testis of horses (Paria et al., 2011), but the expression profile remained uncertain in donkeys. More transcriptional analysis of donkeys should be conducted. It is hypothesized that these novel equine transcripts encode novel proteins or may be non-coding transcription units (Hughes et al., 2010; Skaletsky et al., 2003); this needs further verification.

The *SRY* gene has the function in sex determination at early stages of mammalian embryonic development (Wilhelm et al., 2007). From a functional standpoint it is expected that the *SRY* gene would commonly be found in multiple copies to buffer against mutations in any single copy, thus maintaining enough intact copies to perform the gene's normal function (Lahn et al., 2001). *SRY* has been assumed to be present as multiple functional copies in rabbits (Gerald and Ferrand, 2006), rats (Turner et al., 2007), and cats (Pearks Wilkerson et al., 2008) Y chromosomes. Our results further supported the *SRY* function assumption that *SRY* was found as a multi-copy gene in donkeys with an MCN of 13 and a variable CNV range (1–152). The MCN of *SRY* was highest in JM donkeys (63) and lowest in KL donkeys (2) (Table 4).

The MCNs of Gunsha, JM and Kulun donkeys showed significantly different MCN from 10 breeds (Table S5). However, *SRY* is considered as a single-copy gene in the human and mouse Y chromosome sequences (Hughes et al., 2010; Skaletsky et al., 2003; Toure et al., 2004). In horses, it remains disputed whether *SRY* was single-copied by cDNA FISH (Paria et al., 2011). Therefore, more studies, including gene copy number analysis and detailed expression profiling, will be needed to determine the homology between the two Y chromosomes in more detail.

4 Conclusions

The copy numbers of five donkey genes on Y chromosome were determined in current study. The *CUL4BY*, *ETSTY1*, *ETSTY4*, *ETSTY5*, and *SRY* genes are multi-copy and have MCNs of 5, 45, 2, 2, and 13 with CNV ranges of 1–57, 1–227, 1–37, 1–86 and 1–152, respectively. We first investigated the copy numbers of donkey Y chromosome genes, providing basic genetic information for future studies in donkeys.

Data availability. The original data of the paper are available upon request from the corresponding author.

The Supplement related to this article is available online at <https://doi.org/10.5194/aab-60-391-2017-supplement>.

Author contributions. HH, CL, and RD designed this project; XZ and XX collected blood from animals; HH performed experiments in the laboratory; and HH wrote the paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was supported by National Natural Science Foundation of China (31072001, 31272399, 81270439).

Edited by: Steffen Maak

Reviewed by: two anonymous referees

References

- Cahan, P., Li Y., Izumi, M., and Graubert, T. A.: The impact of copy number variation on local gene expression in mouse hematopoietic stem and progenitor cells, *Nat. Genet.*, 41, 430–437, 2009.
- Dunn, O. J.: Multiple comparisons among means, *J. Am. Stat. Assoc.*, 56, 52–64, 1961.
- Geraldes, A. and Ferrand, N.: A 7-bp insertion in the 39 untranslated region suggests the duplication and concerted evolution of the rabbit SRY gene, *Genetics, selection, evolution, Genetics, selection, evolution: GSE*, 38, 313–320, 2006.
- Graves, J. A.: Evolution of the mammalian Y chromosome and sex-determining genes, *J. Exp. Zool.*, 281, 472–481, 1998.
- Hamilton, C. K., Favetta, L. A., Di Meo, G. P., Floriot, S., Percatt, A., Peippo, J., Kantanen, J., Eggen, A., Iannuzzi, L., and King, W. A.: Copy number variation of testis-specific protein, Y-encoded (TSPY) in 14 different breeds of cattle (*Bos taurus*), *Sex. Dev.*, 3, 205–213, 2009.
- Hamilton, C. K., Revay, T., Domander, R., Favetta, L. A., and Allan King, W. A.: large expansion of the HSFY gene family in cattle shows dispersion across Yq and testis-specific expression, *PLoS One*, 6, e17790, <https://doi.org/10.1371/journal.pone.0017790>, 2011.
- Henrichsen, C. N., Chaignat, E., and Reymond, A.: Copy number variants, diseases and gene expression, *Hum. Mol. Genet.*, 18, R1–8, 2009.
- Hughes, J. F., Skaletsky, H., Pyntikova, T., Graves, T. A., van Daalen, S. K., Minx, P. J., Fulton, R. S., McGrath, S. D., Locke, D. P., Friedman, C., Trask, B. J., Mardis, E. R., Warren, W. C., Repping, S., Rozen, S., Wilson, R. K., and Page, D. C.: Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content, *Nature*, 463, 536–539, 2010.
- Jakubiczka, S., Schnieders, F., and Schmidtke, J.: A bovine homologue of the human TSPY gene, *Genomics*, 17, 732–735, 1993.
- Jobling, M. A.: Copy number variation on the human Y chromosome, *Cytogenet. Genome Res.*, 123, 253–262, 2008.
- Justel Peña, A. D. and Zamar, R.: A multivariate Kolmogorov-Smirnov test of goodness of fit, *Stat. Probab. Lett.*, 35, 251–259, 1997.
- Lahn, B. T., Pearson, N. M., and Jegalian, K.: The human Y chromosome, in the light of evolution, *Nat. Rev. Genet.*, 2, 207–216, 2001.
- Mann, H. B. and Whitney, D. R.: On a test of whether one of two random variables is stochastically larger than the other, *Ann. Math. Stat.*, 18, 50–60, 1947.
- Mills, R. E., Walter, K., Stewart, C., Handsaker, R. E., Chen, K., Alkan, C., Abyzov, A., Yoon, S. C., Ye, K., Cheetham, R. K., Chinwalla, A., Conrad, D. F., Fu, Y., Grubert, F., Hajirasouliha, I., Hormozdiari, F., Iakoucheva, L. M., Iqbal, Z., Kang, S., Kidd, J. M., Konkel, M. K., Korn, J., Khurana, E., Kural, D., Lam, H. Y., Leng, J., Li, R., Li, Y., Lin, C., Luo, R., Mu, X. J., Nemesh, J., Peckham, H. E., Rausch, T., Scally, A., Shi, X., Stromberg, M. P., Stutz, A. M., Urban, A. E., Walker, J. A., Wu, J., Zhang, Y., Zhang, Z. D., Batzer, M. A., Ding, L., Marth, G. T., McVean, G., Sebat, J., Snyder, M., Wang, J., Ye, K., Eichler, E. E., Gerstein, M. B., Hurler, M. E., Lee, C., McCarroll, S. A., Korb, J. O., and 1000 Genomes Project: Mapping copy number variation by population-scale genome sequencing, *Nature*, 470, 59–65, 2011.
- Murphy, W. J., Pearks Wilkerson, A. J., Raudsepp, T., Agarwala, R., Schaffer, A. A., Stanyon, R., and Chowdhary, B. P.: Novel gene acquisition on carnivore Y chromosomes, *PLoS genet.*, 2, e43, <https://doi.org/10.1371/journal.pgen.0020043>, 2006.
- Nickkholgh, B., Noordam, M. J., Hovingh, S. E., van Pelt, A. M., van der Veen, F., and Repping, S.: Y chromosome TSPY copy numbers and semen quality, *Fertil. Steril.*, 94, 1744–1747, 2010.
- Paria, N., Raudsepp, T., Wilkerson, A. J. P., O'Brien, P. C. M., Ferguson-Smith, M. A., Love, C. C., Arnold, C., Rakestraw, P., Murphy, W. J., and Chowdhary, B. P.: A gene catalogue of the euchromatic male-specific region of the horse Y chromosome: comparison with human and other mammals, *PLoS One*, 6, e21374, <https://doi.org/10.1371/journal.pone.0021374>, 2011.
- Pearks Wilkerson, A. J., Raudsepp, T., Graves, T., Albracht, D., Warren, W., Chowdhary, B. P., Skowa, L. C., and Murphy, W. J.: Gene discovery and comparative analysis of X-degenerate genes from the domestic cat Y chromosome, *Genomics*, 92, 329–338, 2008.
- Petroski, M. D. and Deshaies, R. J.: Function and regulation of cullin-RING ubiquitin ligases, *Nat. Rev. Mol. Cell Biol.*, 6, 9–20, 2005.
- Quintana-Murci, L. and Fellous, M.: The human Y chromosome: The biological role of a “functional wasteland”, *J. Biomed. Biotechnol.*, 1, 18–24, 2001.
- Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Daniel Andrews, T., Fiegler, H., Shaper, M. H., Carson, A. R., Chen, W., Kyung Cho, E., Dallaire, S., Freeman, J. L., González, J. R., Gratacòs, M., Huang, J., Kalaitzopoulos, D., Komura, D., MacDonald, J. R., Marshall, C. R., Mei, R., Montgomery, L., Nishimura, K., Okamura, K., Shen, F., Somerville, M. J., Tchinda, J., Valsesia, A., Woodwark, C., Yang, F., Zhang, J., Zerjal, T., Zhang, J., Armengol, L., Conrad, D. F., Estivill, X., Tyler-Smith, C., Carter, N. P., Aburatani, H., Lee, C., Jones, K. W., Scherer, S. W., and Hurler, M. E.: Global variation in copy number in the human genome, *Nature*, 444, 444–454, 2006.
- Sambrook, J. and Russell, D. W.: Molecular cloning: A laboratory manual, translated by: Huang, P. T., Beijing: Science Press; Beijing, China, 2002.
- Shapiro, S. S. and Wilk, M. B.: An analysis of variance test for normality (complete samples), *Biometrika*, 52, 591–611, 1965.
- Skaletsky, H., Kuroda-Kawaguchi, T., Minx, P. J., Cordum, H. S., Hillier, L., Brown, L. G., Repping, S., Pyntikova, T., Ali, J., Bieri, T., Chinwalla, A., Delehaunty, A., Delehaunty, K., Du, H.,

- Fewell, G., Fulton, L., Fulton, R., Graves, T., Hou, S. F., Latrielle, P., Leonard, S., Mardis, E., Maupin, R., McPherson, J., Miner, T., Nash, W., Nguyen, C., Ozersky, P., Pepin, K., Rock, S., Rohlfing, T., Scott, K., Schultz, B., Strong, C., Tin-Wollam, A., Yang, S. P., Waterston, R. H., Wilson, R. K., Rozen, S., and Page, D. C.: The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes, *Nature*, 423, 825–837, 2003.
- Toft, M. and Ross, O. A.: Copy number variation in Parkinson's disease, *Genome Medicine*, 2, p. 62, 2010.
- Toure, A., Szot, M., Mahadevaiah, S. K., Rattigan, A., Ojarikre, O. A., and Burgoyne, P. S.: A new deletion of the mouse Y chromosome long arm associated with the loss of Ssty expression, abnormal sperm development and sterility, *Genetics*, 166, 901–912, 2004.
- Turner, M. E., Martin, C., Martins, A. S., Dunmire, J., Farkas, J., Ely, D. L., and Milsted, A.: Genomic and expression analysis of multiple Sry loci from a single *Rattus norvegicus* Y chromosome, *BMC Genet.*, 8, p. 11, 2007.
- Vodicka, R., Vrtel, R., Dusek, L., Singh, A.R., Krizova, K., Svacinova, V., Horinova, V., Dostal, J., Oborna, I., Brezinova, J., Sobek, A., and Santavy, J.: TSPY gene copy number as a potential new risk factor for male infertility, *Reprod. Biomed. Online*, 14, 579–587, 2007.
- Wilhelm, D., Palmer, S., and Koopman, P.: Sex determination and gonadal development in mammals, *Physiol. Rev.*, 87, 1–28, 2007.
- Yue, X. P., Chang, T. C., DeJarnette, J. M., Marshall, C. E., Lei, C. Z., and Liu, W. S.: Copy number variation of PRAMEY across breeds and its association with male fertility in Holstein sires, *J. Dairy Sci.*, 96, 8024–8034, 2013.
- Zhang, F., Gu, W., Hurles, M. E., and Lupski, J. R.: Copy number variation in human health, disease, and evolution, *Annu. Rev. Genomics Hum. Genet.*, 10, 451–481, 2009.