

# Genetic Analysis of the CHAAR horse breed of Kyrgyzstan

E. Gus Cothran, R. Juras, and J. Kjoellerstroem

October 24, 2022

Animal Genetics Lab., VIBS, School of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, TX, 77843 USA

## 1. Objective

The goal of this study was to gain information on the genetic make-up of the Chaar horse breed of Kyrgyzstan for better insight into the breed's history and with additional analysis of the Toktogul breed from Kyrgyzstan. These data have been added to our data set of over 150 horse breeds to help in our overall understanding of genetic diversity in modern domestic horse breeds.

## 2. Material and methods

### 2.1. Sampling

Hair samples from 30 horses of the Chaar horse breed were collected in November of 2012 and also hair samples from 25 Toktogul horses were collected in June of 2022. DNA was extracted from the hair root follicles and typed for 16 equine specific microsatellite loci from these samples. Domestic horse breeds previously collected and DNA typed for the same 16 microsatellite markers were used for comparison purposes (unpublished and published data).

### 2.2. DNA isolation, PCR analysis and STR genotyping

Genomic DNA was isolated from hair samples using manufacturer's (Qiagen) protocols, with some modifications. All horses were tested by PCR for the 16 STR markers (*AHT4*, *AHT5*, *ASB2*, *ASB17*, *ASB23*, *HMS6*, *HTG6*, *HMS7*, *HMS3*, *HMS2*, *HTG4*, *VHL20*, *HTG7*, *HTG10*, *LEX33* and *LEX3*). However, *LEX3* is X-linked marker and was not used for subsequent data analysis. All markers are included in the panel recommended by the International Society for Animal Genetics for diversity studies and parentage verification. Genotyping was performed using an ABI PRISM 3730 sequencer (Applied Biosystems, Foster City, CA, USA) following previously described methods.

### 2.3 Data analysis

A variety of genetic variability measures were calculated from the gene marker data. The measures were observed heterozygosity ( $H_o$ ) which is the actual number of loci heterozygous per individual; expected heterozygosity ( $H_e$ ), which is the predicted number of heterozygous loci based upon gene frequencies; effective number of alleles ( $A_e$ ) which is a measure of marker system diversity; total number of variants ( $TNV$ ); mean number of alleles per locus ( $MNA$ ); the number of rare alleles observed which are alleles that occur with a frequency of 0.05 or less ( $RA$ ); the percent of rare alleles ( $\%RA$ ); and estimated inbreeding level ( $F_{is}$ ) which is calculated as  $1-H_o/H_e$ . These values are shown in Table 1.

The STR relationship between the Chaar horse and a global cohort from 65 additional breeds was determined by majority-rule consensus of restricted maximum likelihood (RML) trees (Figure 1). The consensus tree was generated using the Reynolds' chord distance and based upon 1000 bootstrapped

allele frequency datasets using the PHYLIP 3.69 package. Trees were visualized by TreeView. The consensus tree established a baseline for genetic relationships among the Char horse and other breeds. Based on the consensus tree and geographic location, the genetic relationship among the Char horse and 21 breeds (including the Toktogul horse) of the global cohort were mapped out using GENETIX to generate a factorial correspondence analysis (FCA).

### **3. Results**

DNA types for Kyrgyzstan horses presented in Tables 2 and 3. All 15 autosomal STR loci tested were found to be polymorphic and did not deviate from HWE in the Char population or Toktogul. Table 1 gives the values for the genetic variability measures of the Char horse. Also shown in Table 1 are values from a representative group of domestic horse breeds (including the Toktogul). The breeds were selected to cover the range of variability measures in domestic horse populations. Mean values for domestic breeds (based upon 80 domestic horse populations) also are shown. No variants were observed which have not been seen in horse breeds. A total of 101 variants were detected in the Char population, well higher than average for domestic horse breeds, of these 31 had frequencies below 0.05%. This percentage of variants at risk of further loss is above average for domestic horse populations. Observed and expected heterozygosity is slightly above the domestic mean, the same is true for the effective number of alleles.

### **4. Discussion**

Genetic studies in the horse have unveiled the diversity of many horse breeds, the ancestry of these breeds, the impact of losing a breed to the species' overall diversity, and the effect of inbreeding on animals' health, especially in small populations. In this study, we examined the genetic variation and the genetic relationships of the Char horses of Kyrgyzstan and compare them to other breeds from around the World. Comprehensive information about genetic diversity is important for establishing appropriate conservation and breeding strategies and can provide the baseline information for conservation efforts.

To better understand the relationship between the Char horses and other horse breeds that may have contributed to its history, a majority-rule consensus tree was generated, Figure 1. The breeds that were most like the Char was the Mongolian horse, which paired directly with the Char, and the Yakut (a Siberian horse and Exmoor Pony (which is a primitive type of British Native Pony breed)) which joined with the Char/Mongolian Horse branch of the tree. The Toktogul clusters within a group of breeds that have strongly influenced by contributions from the English Thoroughbred.

The similarity of the Char horse and other breeds was also examined using the factorial correspondence analysis FCA for 21 other breeds representing breeds from Figure 1 that showed closest resemblance to the Char as well as to the Toktogul – Figure 2. FCA results are very similar to the RML tree as close genetic relationships of Mongolian and Yakut horses is quite clear. Here we see association to the Char to other Asian breeds as well as the South American Peruvian Paso. This latter result is likely due to the limited number of breeds in the analysis which caused to Peruvian to fall in a place that was not natural for it.

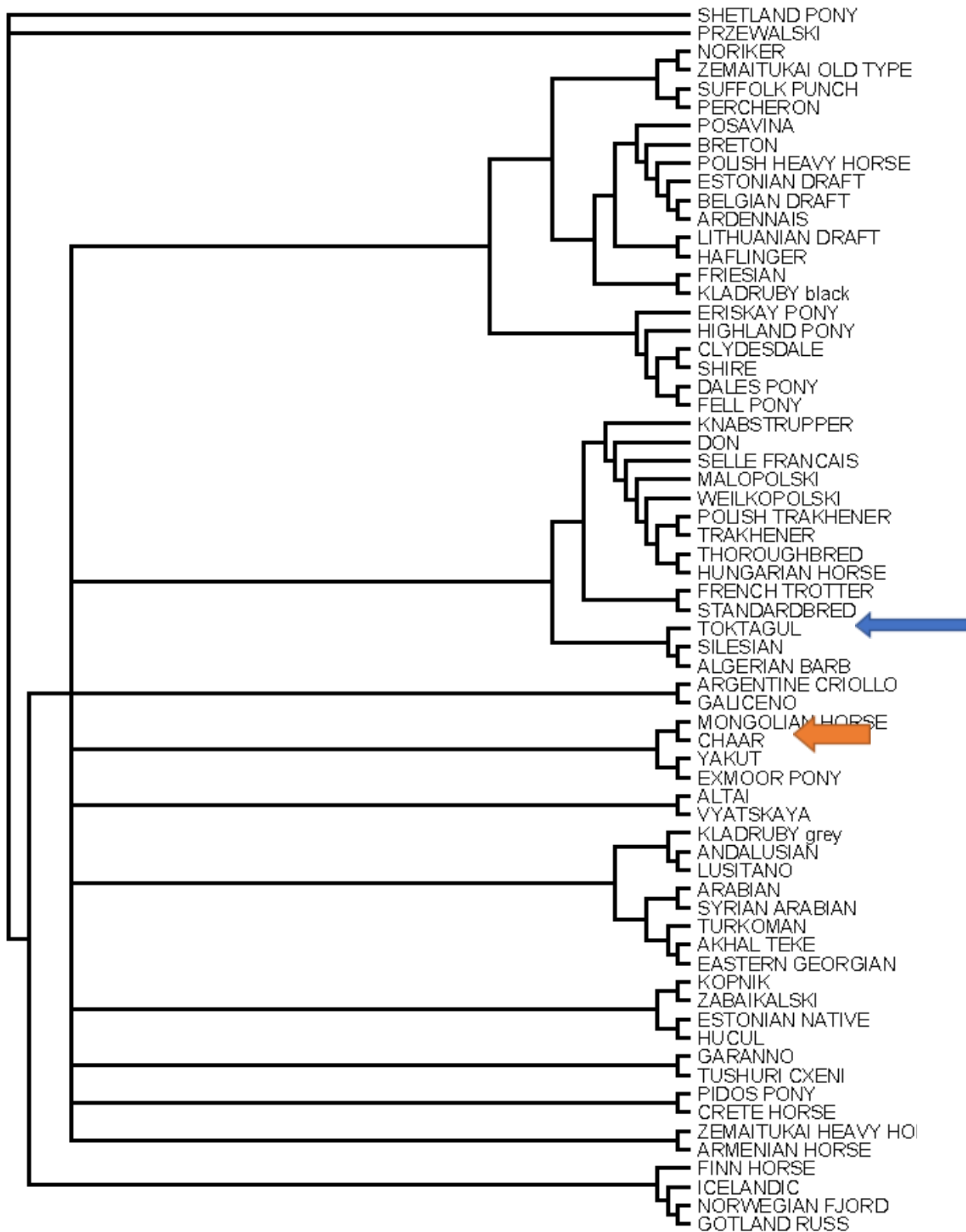
Table 1. Genetic diversity measures based upon 15 microsatellite loci.

BREED	N	Ho	He	Fis	Ae	TN V	MN A	%RAR E
<b>CHAAR</b>	<b>30</b>	<b>0.753</b>	<b>0.732</b>	<b>-0.028</b>	<b>3.969</b>	<b>101</b>	<b>6.733</b>	<b>0.307</b>
<b>TOKTOGUL</b>	<b>25</b>	<b>0.811</b>	<b>0.778</b>	<b>-0.055</b>	<b>4.820</b>	<b>116</b>	<b>7.733</b>	<b>0.301</b>
AKHAL TEKE	29	0.726 4	0.643 1	-0.1295	3.341	81	5.400	0.222
AMERICAN CREAM DRAFT	25	0.725 3	0.711 4	-0.0196	3.735	92	6.133	0.261
AMERICAN SADDLEBRED	179	0.697 3	0.705 7	0.0120	3.948	106	7.067	0.358
ANDALUSIAN	53	0.688 5	0.716 5	0.0391	3.968	98	6.533	0.327
ARABIAN	95	0.643 2	0.695 1	0.0748	3.694	102	6.800	0.343
BELGIAN DRAFT	20	0.653 3	0.656 1	0.0042	3.419	83	5.533	0.253
BOULONNAIS	38	0.657 9	0.614 2	-0.0712	3.023	74	4.933	0.203
BRETON	26	0.730 8	0.657 2	-0.1119	3.415	77	5.133	0.208
CLEVELAND BAY	90	0.628 9	0.618 6	-0.0166	2.870	73	4.867	0.274
CLYDESDALE	32	0.602 1	0.581 8	-0.0348	2.563	66	4.400	0.258
DALES PONY	86	0.712 5	0.688 2	-0.0353	3.653	92	6.133	0.283
DARTMOOR PONY	75	0.712 0	0.671 1	-0.0609	3.319	95	6.333	0.326
DON	34	0.688 2	0.718 9	0.0426	3.740	98	6.533	0.286
EXMOOR PONY	46	0.598 0	0.625 0	0.0431	2.841	79	5.267	0.329
FELL PONY	55	0.722 4	0.716 6	-0.0081	3.763	88	5.867	0.250
FRIESIAN	143	0.475 6	0.519 5	0.0845	2.524	76	5.067	0.342
HAFLINGER	46	0.657 6	0.660 0	0.0036	3.388	85	5.667	0.259
HIGHLAND PONY	28	0.683 3	0.663 1	-0.0306	3.165	80	5.333	0.225
NEW FOREST PONY	44	0.805 3	0.782 8	-0.0288	4.940	119	7.933	0.286
PERCHERON	38	0.684 2	0.683 0	-0.0017	3.546	98	6.533	0.337
POLISH HEAVY HORSE	69	0.758 5	0.719 2	-0.0546	3.953	101	6.733	0.248
POSAVINA	27	0.743 2	0.712 5	-0.0431	4.063	98	6.533	0.255
QUARTER HORSE	77	0.689 7	0.751 3	0.0820	4.321	119	7.933	0.378

SHETLAND PONY	112	0.670 0	0.748 9	0.1054	4.536	121	8.067	0.388
SHIRE	32	0.685 4	0.654 5	-0.0473	3.134	81	5.400	0.321
THOROUGHBRED	139	0.651 0	0.707 9	0.0803	3.694	85 80.8	5.667	0.271
DOMESTIC HORSE MEAN	80	0.710	0.720	0.012	4.012	8 16.7	6.74	0.283
Standard Deviation		0.078	0.071	0.086	0.735	9	1.40	0.082

The above measures are observed heterozygosity ( $H_o$ ) which is the actual number of loci heterozygous per individual; expected heterozygosity ( $H_e$ ), which is the predicted number of heterozygous loci based upon gene frequencies; effective number of alleles ( $A_e$ ) which is a measure of marker system diversity; total number of variants ( $TNV$ ); mean number of alleles per locus ( $MNA$ ); the percent of rare alleles ( $\%RA$ ) which are alleles that occur with a frequency of 0.05 or less; and estimated inbreeding level ( $Fis$ ) which is calculated as  $1-H_o/H_e$ .

Figure 1. RML tree of genetic similarity to domestic horse breeds



**Figure 2a.** Factorial correspondence analysis (FCA) of Chaar Horse in comparison with 21 horse breeds (breed list is Table 1S). Each color represents the population and each dot represents individual.

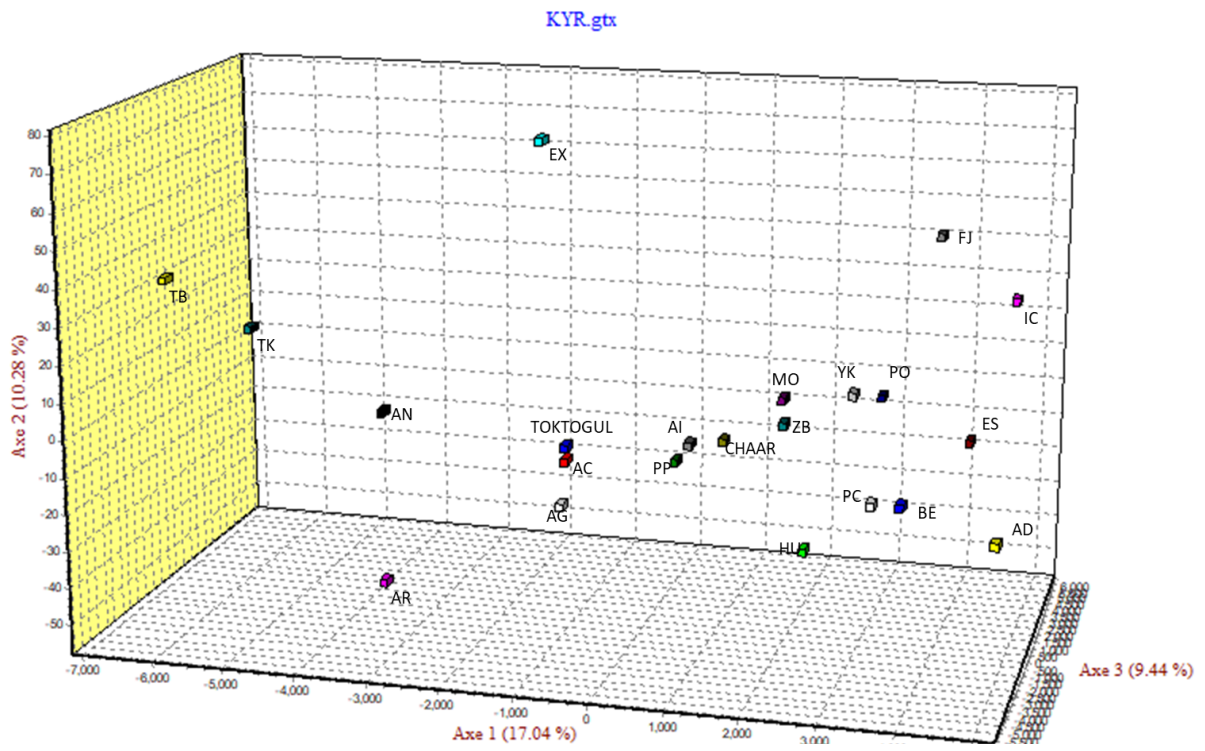


Table S1. Breed identification of animals used in FCA analysis

# Acronym Breed name

---

1	AD	ARDENNAIS
2	BE	BELGIAN DRAFT
3	PC	PERCHERON
4	FJ	NORWEGIAN FJORD
5	IC	ICELANDIC
6	ZB	ZABAILKALSKY
7	PO	POLISH PRIMITIVE
8	ES	ESTONIAN NATIVE HORSE
9	HU	HUCUL
10	CHAAR	CHAAR KYRGYZSTAN
11	MO	MONGOLIAN
12	YK	YAKUT
13	EX	EXMOOR PONY
14	AN	ANDALUSIAN
15	AC	ARGENTINE CRIOLLO
16	PP	PERUVIAN PASO
17	TB	THOROUGHBRED
18	TOKTAGUL	TOKTOGUL
19	AG	ALGERIAN BARB
20	AI	ALTAI NATIVE HORSE
21	AR	ARABIAN
22	TK	TRAKHENER

---





Table 2. Genotypes of Chara horses

6246 3	KZ4 7	LP	KL	JP	LM	O O	JN	KM	MO	O R	NO	M Q	MT	IS	JU	LO	FF
6246 4	KZ4 7	OP	LL	IO	LL	JJ	JO	KM	NO	IO	OO	PQ	M M	IS	JL	LO	LL
6246 5	KZ4 7	OO	LL	HI	LM	O O	N N	MP	MN	RR	NN	MP	M M	NS	LU	Q Q	M M
6246 6	KZ4 7	MR	LM	JO	LL	JO	N N	MP	MN	LO	KM	PP	JM	NO	LU	OT	KK
6246 7	KZ4 7	LO	LM	H H	LN	JO	JJ	M M	NO	O R	NN	PP	M N	NO	JL	Q Q	LL
6246 8	KZ4 7	MO	KL	HI	KL	JO	KN	MP	NO	LR	KN	MR	MT	NS	KU	N Q	LM
6246 9	KZ4 7	OP	LM	IO	LM	JO	JO	M M	MO	IR	NO	PP	M M	NS	LU	LL	FL
6247 0	KZ4 7	JM	KL	LP	LM	JO	JJ	M M	MN	O R	M O	PR	MN	IS	JU	O Q	M M
6247 1	KZ4 7	MR	KL	HJ	KM	JJ	KN	MO	MO	LR	KM	MR	KT	NN	JU	N Q	MP
6247 2	KZ4 7	RR	KL	HP	M O	G O	LL	OP	M M	IO	OO	IM	TT	LN	U U	GL	NN
6247 3	KZ4 7	OP	M M	HI	LM	JO	JJ	M M	NO	O R	KN	MP	KM	NN	JL	LQ	HM
6247 4	KZ4 7	M M	KL	JP	KL	JJ	JK	KO	NO	LR	M O	PR	KM	IP	JJ	O Q	LM
6247 5	KZ4 7	MN	MP	H Q	LO	G G	N N	KP	KK	LQ	KO	IM	JN	IQ	KK	KQ	M M
6247 6	KZ4 7	OO	M M	HI	LN	JO	JJ	MO	MO	IR	KO	PP	M M	NN	LU	LL	HL
6247 7	KZ4 7	M M	LM	HP	LM	JO	JN	KM	MO	RR	M O	PQ	M M	IM	IJ	KO	FL
6247 8	KZ4 7	MO	LM	H H	KN	O O	KN	MP	NO	O R	KN	MP	JM	NN	LU	N Q	M M
6247 9	KZ4 7	IM	LM	HJ	KO	JJ	KO	LO	MN	O R	M O	PP	KM	PR	JQ	LQ	M M
6248 1	KZ4 7	MR	LM	PP	LM	JJ	JK	LM	NO	RR	NO	PQ	M M	IN	IJ	O Q	FL
6248 2	KZ4 7	MN	LL	HI	KO	JO	JN	MO	NN	O R	KO	MP	JT	RS	IU	LN	LL
6248 3	KZ4 7	NR	MO	HP	M O	JO	JK	LP	MN	O R	OO	PP	JM	LN	JL	LO	LL
6248 4	KZ4 7	NP	LL	HI	M O	O O	JJ	MO	M M	O R	OO	NP	JM	RS	IL	LL	LL
6248 5	KZ4 7	LP	LL	JO	LM	JJ	JN	KL	NO	O R	M O	MP	KM	MS	JJ	O Q	FL
6248 6	KZ4 7	LR	KP	O P	LN	O O	IJ	MP	NO	O O	KK	O O	JO	GG	JQ	LL	LN
6248 7	KZ4 7	QR	LM	O P	JJ	O O	JN	M M	MN	O Q	KK	OR	JR	M Q	JJ	LL	FL
6248 8	KZ4 7	JR	LM	HP	LM	JJ	JO	LL	KM	LL	KN	PP	MN	IN	IQ	KQ	LM
6248 9	KZ4 7	JQ	LM	HP	JL	O O	JN	MP	M M	KS	KK	OP	RT	QR	JL	LL	FN
6249 0	KZ4 7	QR	LL	HK	JN	O O	JJ	MP	MN	KO	KO	MP	NR	GR	JL	KL	FL

6249 1	KZ4 7	MR	MP	KO	JO	O O	J M	MO	NP	O R	OO	M O	JN	GK	JJ	GK	FL
6249 2	KZ4 7	M M	LM	KO	LL	IJ	JO	MP	KQ	O R	OO	IQ	KM	NR	JU	L M	HM
6249 3	KZ4 7	MP	M M	HK	JK	IJ	KN	PP	BN	KO	NO	MR	MN	M N	KL	Q Q	HH

Table 3. Genotypes of Toktogul horses

14265 3	KY7 1	IR	M M	JK	KL	M O	MQ	KL	NO	MR	KO	IM	JN	GR	JS	KL	FM
14265 4	KY7 1	IM	KM	JK	JO	JO	NO	LO	MP	OO	KN	II	JJ	QY	K U	KL	JO
14265 5	KY7 1	IN	LM	JK	JJ	GO	LM	PP	KQ	M O	KO	IP	JN	GL	IJ	LO	MO
14265 6	KY7 1	IM	M M	HJ	M N	GG	MN	LM	KQ	LR	OO	OP	LM	HN	SS	KL	FL
14265 7	KY7 1	LR	KM	H O	M N	IJ	JN	LO	NP	RR	OO	PQ	JJ	HN	SS	RR	M M
14265 8	KY7 1	IM	KM	O R	JL	IJ	JO	M M	KM	M Q	NO	MP	NQ	NQ	JL	LL	HH
14265 9	KY7 1	IM	LM	KK	JK	OP	JM	LM	QQ	KM	KO	NR	NN	FG	S U	KO	M M
14266 0	KY7 1	JM	KM	LO	JO	GJ	MO	KO	M M	M O	NO	MP	MN	N W	IK	KL	JN
14266 1	KY7 1	II	M M	H O	JL	JP	JJ	LP	KM	IN	KN	MP	JM	NR	KL	LL	LL
14266 2	KY7 1	PR	MN	IL	QQ	JM	MN	MP	KM	LR	M O	PP	JT	FI	JK	GK	FL
14266 3	KY7 1	IR	LM	H O	JN	JJ	MN	LO	PQ	MR	OO	NP	JN	GH	SS	LR	M M
14266 4	KY7 1	II	LM	JK	JL	GO	MQ	LP	OQ	MR	OO	IP	JN	GR	SS	LQ	FM
14266 5	KY7 1	IM	KL	JK	JK	JO	MN	KP	OO	LM	KO	NR	JN	GL	IS	LO	HM
14266 6	KY7 1	IR	LM	JO	LL	G M	NQ	KL	NO	IR	KO	MP	JL	OR	LS	KQ	FF
14266 7	KY7 1	LQ	M M	O O	LL	OP	M M	PP	OQ	KR	OO	IM	MO	KR	K R	Q R	HH
14266 8	KY7 1	M M	LP	HP	M O	OO	LN	PP	MN	LO	NN	NP	MP	GQ	IJ	LO	FL

14266 9	KY7 1	M M	MP	HJ	JO	GG	MN	PP	NO	OR	NN	MP	JJ	GR	L U	KO	FF
14267 0	KY7 1	NP	MP	JP	KO	OO	LM	M M	KM	RR	KO	IP	JM	JR	JL	LO	HH
14267 1	KY7 1	QR	LO	JK	LN	IO	KO	KM	BM	RR	KN	M Q	KK	QR	JS	LQ	HH
14267 2	KY7 1	OR	M M	O O	LO	GJ	LO	KP	NQ	IR	NO	PR	NR	HN	LL	LQ	LL
14267 3	KY7 1	M M	M M	H O	LQ	OO	KO	PR	KN	IR	M O	PP	NN	RR	K K	L M	PP
14267 4	KY7 1	IP	KM	KO	LO	M O	JO	MO	NP	LO	OO	NP	MT	NT	KS	LQ	LM
14267 5	KY7 1	LQ	MN	J M	LN	OO	MN	KP	MN	M O	OO	QR	KM	GN	S U	L M	IL
14267 6	KY7 1	IR	LM	JO	NN	GJ	JO	OP	IM	IQ	OO	MP	M M	LN	JU	G Q	IP
14267 7	KY7 1	MP	KP	JO	JO	GO	JK	LP	MN	M O	NO	M O	MT	MR	L U	O Q	KM