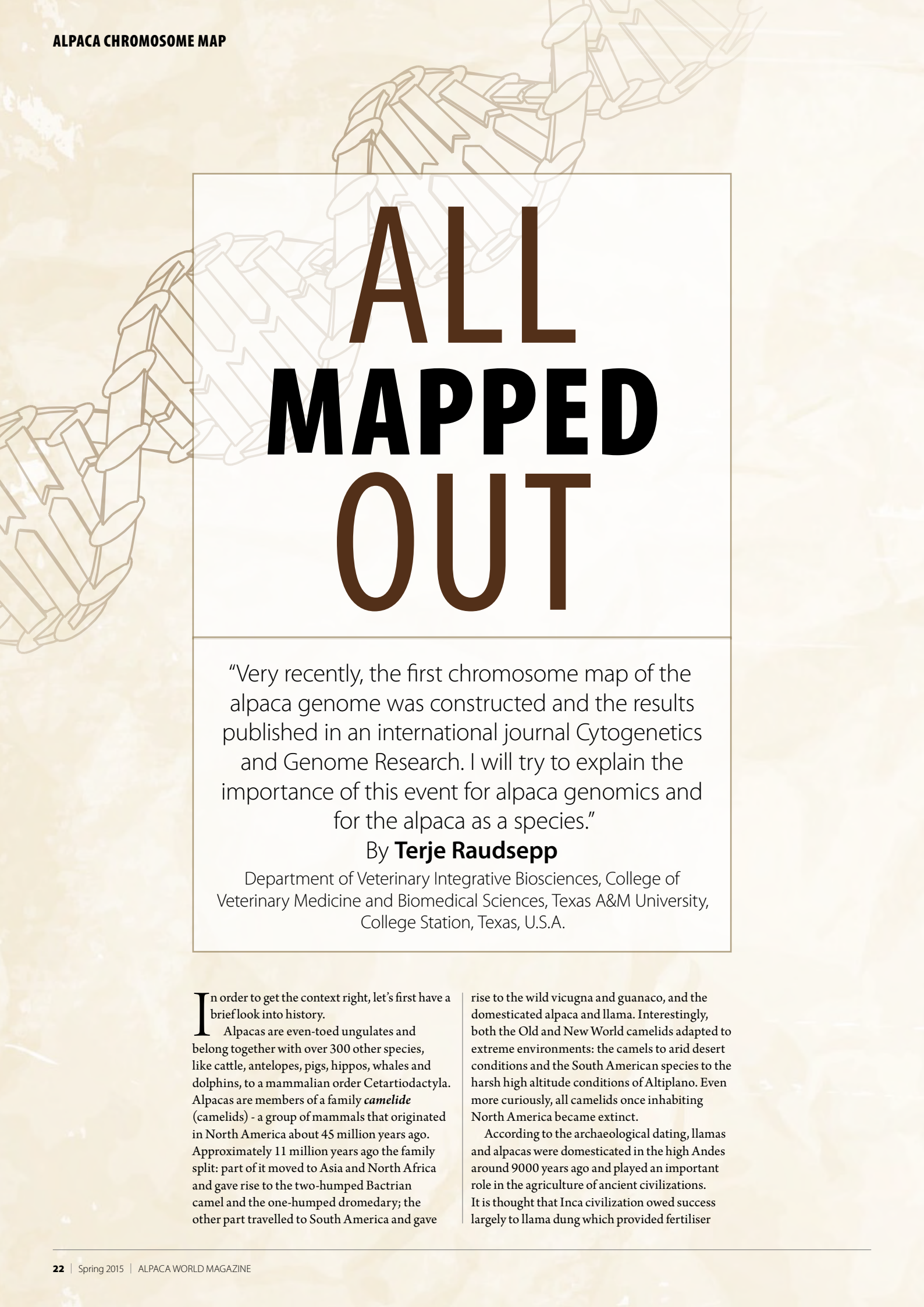




ALL MAPPED OUT

“Very recently, the first chromosome map of the alpaca genome was constructed and the results published in an international journal Cytogenetics and Genome Research. I will try to explain the importance of this event for alpaca genomics and for the alpaca as a species.”

By **Terje Raudsepp**

Department of Veterinary Integrative Biosciences, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, Texas, U.S.A.

In order to get the context right, let's first have a brief look into history.

Alpacas are even-toed ungulates and belong together with over 300 other species, like cattle, antelopes, pigs, hippos, whales and dolphins, to a mammalian order Cetartiodactyla. Alpacas are members of a family *camelide* (camelids) - a group of mammals that originated in North America about 45 million years ago. Approximately 11 million years ago the family split: part of it moved to Asia and North Africa and gave rise to the two-humped Bactrian camel and the one-humped dromedary; the other part travelled to South America and gave

rise to the wild vicugna and guanaco, and the domesticated alpaca and llama. Interestingly, both the Old and New World camelids adapted to extreme environments: the camels to arid desert conditions and the South American species to the harsh high altitude conditions of Altiplano. Even more curiously, all camelids once inhabiting North America became extinct.

According to the archaeological dating, llamas and alpacas were domesticated in the high Andes around 9000 years ago and played an important role in the agriculture of ancient civilizations. It is thought that Inca civilization owed success largely to llama dung which provided fertiliser

and enabled corn to be cultivated at very high altitudes. The Bactrian camel was domesticated in China and Mongolia about 6000 years ago and the dromedary in coastal settlements of the Arabian Peninsula about 4500 years ago. Thus, camelids have been associated with humans for as long as cattle, horses and dogs. Despite this, genome analysis of the alpaca had a late start and lags behind those conducted in other domestic species. In part, this is due to historical, geographic and economic reasons, in part, due to difficulties in studying the alpaca chromosomes.

GENE MAPS

It is well established that the genetic material, the DNA, is carried on from cell to cell, and from generation to generation in the form of chromosomes - defined structures located in the cell nucleus. The set of chromosomes in a cell is known as the **karyotype** and is species specific. Chromosomes also reflect the genetic health of an individual because changes in chromosome number or shape cause congenital disorders, reduced fertility and disease. As chromosomes carry the genetic material, they are important components of gene maps. Such maps show the exact location of genes in chromosomes; determine which genes are on the same, which in different chromosomes; which are neighbours, which are far away from each other, and what is the order of genes in each chromosome.

Even today, when revolutionary technologies, collectively known as **next generation sequencing** (NGS), have made it feasible to sequence the whole genome of any species or individual, the importance of chromosome maps has not reduced. This is because the three billion DNA letters of mammalian genomes cannot be sequenced in one long succession. Instead, genomes are sequenced in millions of short

fragments that have to be assembled back into whole genome, a process similar to building Legos or puzzles. Here, the chromosome-based gene maps serve as instruction sheets by guiding and validating the assembly and anchoring sequences to specific chromosomes.

In the 1960s-1980s karyotypes of all main domestic species, such as cattle, pigs, sheep, goats, horses, dogs, cats and chicken, were characterised in detail. This led to the construction of detailed gene maps which, in turn, set the foundation for the assembly of genome sequence for these species. Chromosome studies in alpacas and other camelids also date back over 50 years with the first publications in 1965. It is established that the alpaca genome is packed into 74 chromosomes which is on the higher side among mammals. For a comparison, pigs and cats have 38, humans 46, sheep 54, cattle and goats 60, horses 64, and dogs and chicken 78 chromosomes.

A karyotype of a female alpaca is shown in **figure 1A**: it comprises 36 pairs of chromosomes known as **autosomes** which are present in both males and females, and one pair of **sex chromosomes** - females have two X chromosomes, males one X and one Y chromosome. However, because of the very similar size and shape of individual chromosomes, there has never been a common agreement among researchers about how to recognise and arrange alpaca chromosomes in the karyotype. This has hindered the development of gene maps and the study of chromosome abnormalities in relation to diseases and congenital disorders.

Curiously, despite millions of years of separate evolution and the very distinct adaptations, all camelids have similar karyotypes: all with 74 chromosomes and all difficult to study. **Figure 1** illustrates this by showing the karyotypes of a female alpaca and a female Bactrian camel.

The alpaca genome is packed into 74 chromosomes which is on the higher side among mammals; pigs and cats have 38, humans 46, sheep 54, cattle and goats 60, horses 64, and dogs and chicken 78 chromosomes.

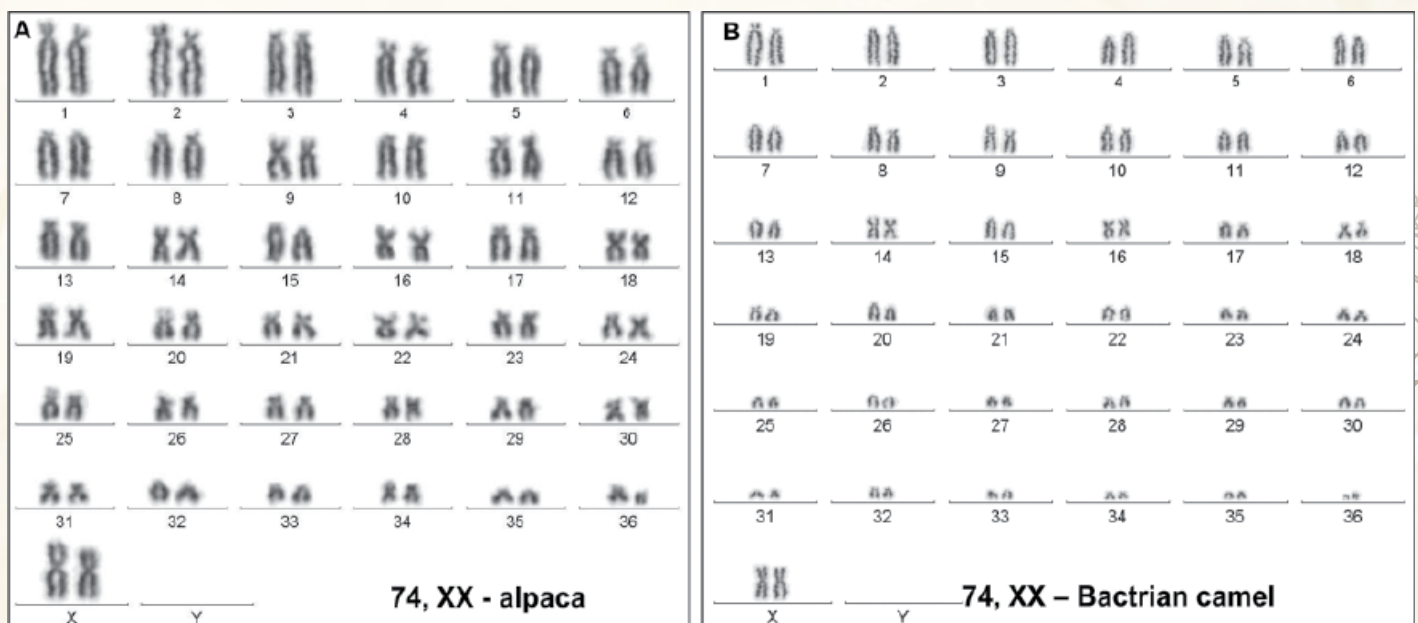


Figure 1: The karyotypes of a female alpaca (A) and a female Dromedary camel. Note the amazing similarity between the chromosomes despite of millions of years of separate evolution.

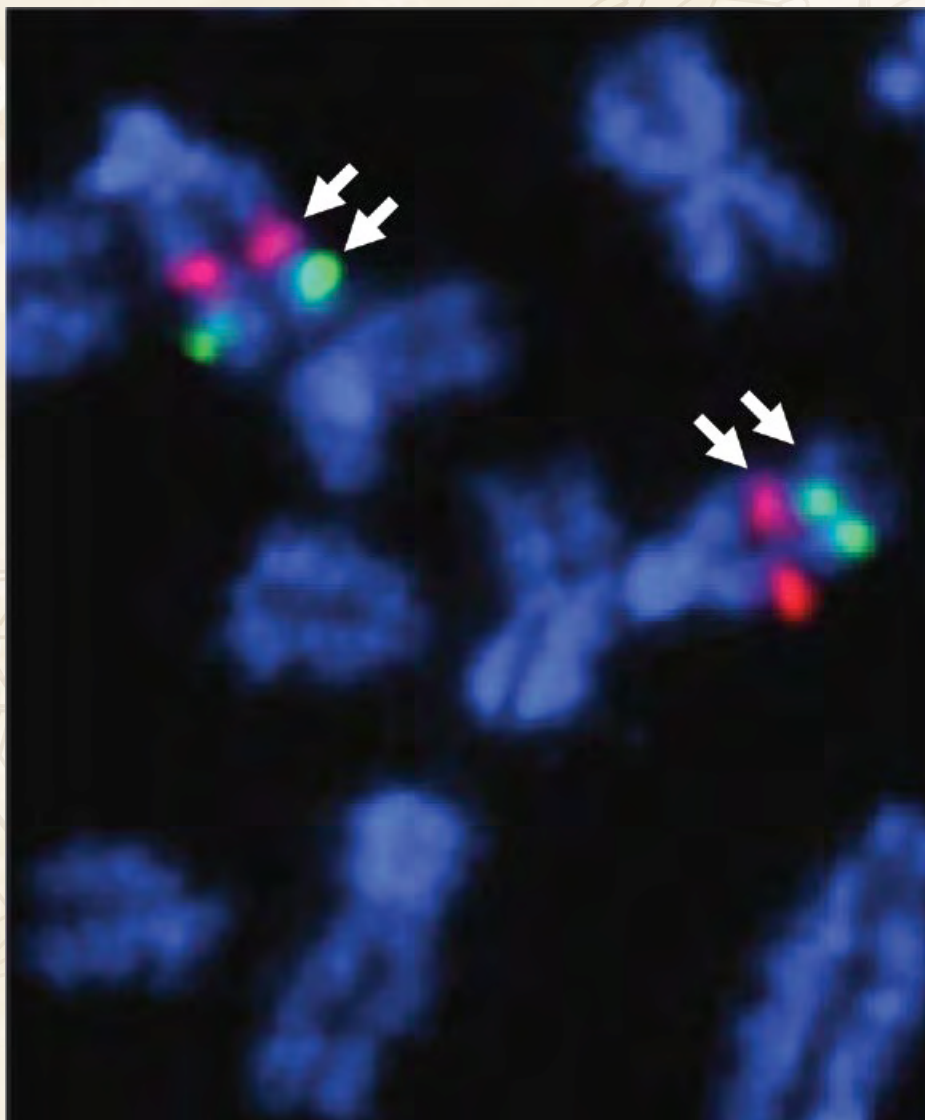


Figure 2 (above): A microscope image showing the results of a two-colour FISH experiment in alpaca chromosome 17. Alpaca chromosomes are stained blue; one gene is labelled red, another green.

The map is an important 'instruction sheet' for the assembly of alpaca and camelid genome sequences and for the discovery of genes related to genetic diseases and traits of biological importance.

Despite the difficulties, these karyotypes have one important advantage: if we construct a chromosome map for one species, this map can be effectively used for all camelids.

WHOLE GENOME SEQUENCING

One of the distinctive signatures of the 21st century genomics is *whole genome sequencing*. Reference genome sequences are available for all main domestic species. Alpaca is not an exception and its genome sequencing project was initiated in 2002. The sequenced animal was a female huacaya alpaca *Nyala Empress Carlotta* (she got her name due to being conceived on a parking lot). *Carlotta's* genome was sequenced two times, first by the traditional (Sanger) sequencing technology and second time, by NGS. This was an outstanding development in alpaca genomics, though with one big problem: in contrast to cattle, horses and dogs, the alpaca had no chromosome-based gene maps to help the sequence assembly - the instruction sheet for a 3 billion piece Lego was missing.

We had the alpaca whole genome sequence but did not know which sequence belongs to which chromosome. That is why in 2009 we proposed to Morris Animal Foundation to construct a whole genome chromosome map for the alpaca

Figure 3 (opposite): Chromosome map of the alpaca genome. The 230 FISH-mapped markers are shown to the right of the schematic drawings of chromosomes.

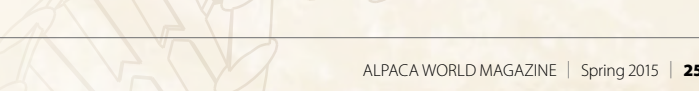
by assigning genes and DNA sequences to all 36 pairs of autosomes and the sex chromosomes.

Chromosome maps or *cytogenetic* maps are based on a method known as Fluorescence In Situ Hybridization or FISH which determines the location of a gene or any other DNA sequence in its original place or *in situ* (Latin) in the chromosome. FISH has two components - the *probe* and the *target*. The targets are chromosomes, while the probe is a DNA sequence to be mapped. The probes are labelled with fluorescent molecules, so that after hybridization one can see their location in chromosomes using a fluorescence microscope.

If probes are labelled with 2-3 spectrally different fluorescent dyes and used in the same FISH experiment, we can see at once the location of two or more genes, decide whether or not they are on the same chromosome, and if they are, determine their relative order. Such a FISH experiment is illustrated in **figure 2** showing the location of two genes in alpaca chromosome 17, one labelled with green, another with red fluorescent dye. The two sequences are in the same chromosome and the green probe is closer to the chromosome end (the telomere) and the red probe closer to the centromere, seen as a constriction. Most importantly, this tells us that all DNA sequences that are in the close vicinity to these two markers in the alpaca genome sequence assembly are also located in chromosome 17.

This way, we designed and fluorescently labelled probes for 230 alpaca genes and DNA markers and constructed maps for all alpaca chromosomes as shown in **figure 3**. Each chromosome has at least two FISH-mapped markers, though most have 5 to 15 markers. The alpaca whole genome chromosome map is the first of its kind in any camelid species and an important landmark for alpaca genomics. The map anchors almost all alpaca genome sequences to defined chromosomes showing which genes are close neighbours and which are located in different chromosomes. The knowledge is important for understanding gene interactions in the normal genome, as well as in case of mutations or chromosome rearrangements.

In addition to the alpaca, we also mapped 100 genes in the dromedary and the llama. As expected, all genes mapped to the same chromosomes and to the same locations in all three species, confirming the extraordinary conservation of these genomes in the course of evolution. Therefore, by mapping 230 genes in the alpaca chromosomes, we indirectly constructed chromosome maps for all camelid species. The map is an important 'instruction sheet' for the assembly of alpaca and camelid genome sequences and for the discovery of genes related to genetic diseases and traits of biological importance.



TRANSLOCATION

Last but not least, the 230 markers mapped to alpaca chromosomes are also tools for telling chromosomes apart, particularly when studying chromosome rearrangements. For example, in case of translocations when chromosomes from different pairs fuse, we can tell which chromosomes are involved and which genes relocated. To illustrate this, we recently studied a chromosome abnormality in a sterile male llama. The llama had 73 chromosomes, thus one less than normal. Because whole chromosome deletions are usually lethal, we suspected a translocation. Indeed, the karyotype contained one large chromosome which was never seen in normal llamas. However, by traditional chromosome analysis we were not able to tell the origin of this abnormal chromosome.

Now, with the alpaca cytogenetic gene map in hand, we selected a set of markers from the most likely candidate chromosomes and carried out a series of FISH experiments in the sterile llama. As a result, we could show that the abnormal chromosome resulted from a fusion of chromosomes 12 and 20 (illustrated in figure 4).

Incidentally, this is the first cytogenetically and molecularly characterised chromosomal translocation in alpacas and camelids.

Taken together, with the whole genome chromosome map available, alpaca finally joins the 'club' with other domestic species and can have the genome sequence properly assigned to chromosomes. Also, for the first time alpaca has a molecular tool kit for clinical cytogenetics. And what is most wonderful, both the map and the molecular tool kit can be shared with its South American cousins and the relatives in Asia and Africa.

ACKNOWLEDGMENTS

I am grateful to a group of excellent and dedicated researchers for their important contributions to alpaca gene mapping and chromosome studies (in alphabetical order): Felipe Avila, Malorie P. Baily, Renuka Chowdhary, Pranab J. Das, Warren E. Johnson, Michelle A. Kutzler, David A. Merriwether, Elaine Owens, Polina Perelman, and Vladimir A. Trifonov. I also thank Morris Animal Foundation and the Alpaca Research Foundation for funding these studies. 🐾

Taken together, with the whole genome chromosome map available, alpaca finally joins the 'club' with other domestic species and can have the genome sequence properly assigned to chromosomes.

Figure 4: Chromosome abnormality in a sterile male llama: the large abnormal chromosome is the result of fusion of chromosomes 12 (green marker) and 20 (red marker). The photo of the llama is courtesy of Dr. Michelle Kutzler.



Sterile llama with an abnormal chromosome

