

Research Article

Construction of a Llama Bacterial Artificial Chromosome Library with Approximately 9-Fold Genome Equivalent Coverage

K. W. Airmet,¹ J. D. Hinckley,¹ L. T. Tree,² M. Moss,¹ S. Blumell,¹ K. Ulicny,¹ A. K. Gustafson,¹ M. Weed,¹ R. Theodosis,¹ M. Lehnardt,¹ J. Genho,¹ M. R. Stevens,² and D. L. Kooyman¹

¹ Department of Physiology and Developmental Biology, Brigham Young University, Provo, UT 84602, USA

² Department of Plant and Wildlife Sciences, Brigham Young University, Provo, UT 84602, USA

Correspondence should be addressed to K. W. Airmet, kairmet@hotmail.com

Received 10 March 2012; Revised 21 May 2012; Accepted 4 June 2012

Academic Editor: Noelle E. Cockett

Copyright © 2012 K. W. Airmet et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The llama is an important agricultural livestock in much of South America. The llama is increasing in popularity in the United States as a companion animal. Little work has been done to improve llama production using modern technology. A paucity of information is available regarding the llama genome. We report the construction of a llama bacterial artificial chromosome (BAC) library of about 196,224 clones in the vector pECBAC1. Using flow cytometry and bovine, human, mouse, and chicken as controls, we determined the llama genome size to be 2.4×10^9 bp. The average insert size of the library is 137.8 kb corresponding to approximately 9-fold genome coverage. Further studies are needed to further characterize the library and llama genome. We anticipate that this new library will help facilitate future genomic studies in the llama.

1. Introduction

There are six species in the family Camelidae; namely, *Camelus bactrianus*, *Camelus dromedarius*, *Lama glama* (llama), *Lama guanicoe* (guanaco), *Lama pacos* (alpaca), and *Vicugna vicugna* (vicuña). Llamas are especially important for agriculture in Andean South America. They are also indispensable in the Bolivian and Peruvian high plain (Altiplano) where they are well adapted to the harsh conditions. In these locations, the uses of llamas include fiber, meat, and draft. In the impoverished communities of the Altiplano, the utilization of llamas, in these ways, may represent a considerable amount of a family's disposable income. Improvement in the production of llamas would aid the rural families of the Altiplano to have a steady food source and income as well as aid in becoming self-sufficient [1]. Llamas and alpacas are also rising in popularity in the United States.

Our long-term goal is to increase the understanding of the genetic basis of phenotypic differences in camelids. In most common livestock and companion animal species, bacterial artificial chromosome (BAC) libraries have facilitated genetic work. However, until recently, a BAC library has not existed for camelids, which has impeded the progress of

genetic research in this family. Therefore, there was a critical need for a llama BAC library. It is important to recognize that the llama genome size has already been determined [2]. This demonstrates the growing importance of this livestock species.

BAC libraries can provide a way of obtaining complete gene sequences, including intron sequences. Intron sequences are important because they contain more polymorphisms than exons. In addition, introns allow for the design of primers, which amplify entire exons. BAC libraries have been developed for a number of plant and animal species of major and minor economic/agricultural impacts including durum wheat [3], coffee [4], pepper [5], ginseng [6], quinoa [7], bovine [8, 9], domestic dog [10], and sheep [11]. These BAC libraries have been used to facilitate physical mapping, gene cloning, and many other types of studies [12].

A linkage map for the llama species is yet to be created. However, this and other future projects will help in identifying the causative genetic differences in Camelids to improve selection for particular traits. Such identification of mutations responsible for phenotypes, for instance, has already been performed in sheep [13] and pigs [14] as well as a number of other species.

A llama BAC library will aid in research related to llama production throughout the world.

2. Materials and Methods

2.1. Preparation of Megabase Size DNA. Blood was drawn from llamas after which red blood cells were separated from the white blood cells by centrifugation. The white cells were resuspended in sterile normal saline (0.9% NaCl). This procedure was repeated three times and white cells were counted using a haemocytometer. The cells were diluted in sterile normal saline to a concentration of 1×10^8 /mL. Eight milliliters of the white blood cells suspension was combined with 8 mL of low melting point (LMP) agarose to 100 μ L form plugs. The formed plugs were incubated at 4 C for 15 min. and harvested into 50 mL centrifuge tubes. Forty-five ml of lysis buffer (0.5 M EDTA, pH 9.0; 1% lauryl sarcosine and 1 mg/mL proteinase K) was added to each tube and they were incubated at 50 C for 48 hours with gentle rotation.

2.2. Size Selection. While still in the plug, the purified DNA was washed and partially digested with either EcoRI or BamHI. The plugs were then loaded in a CHEF Gel (BioRad Laboratories, Hercules, CA USA) and DNA was separated/fractionated by pulse field electrophoresis. The resulting fractionated DNA was divided into three size categories, cut from the gel, resuspended, and purified. The size categories we used, based upon Lambda DNA standard, were 100–200 kb, 200–300 kb, and 300–400 kb. The sizes were subsequently identified as low, medium, and high molecular weight, respectively.

2.3. Creation of Libraries. The libraries were constructed with approximately 150 ng llama digested DNA, which was used at a rate of 0.6–1.0 ng/ μ L in ligation reactions of linearized dephosphorylated pECBAC1 according to Zhang [15], with minor modifications. At the recommended 1:4 molecular weight ratio (vector to DNA), ElectroMax DH10B cells (Invitrogen, Carlsbad, CA) were transformed with ligation mix by electroporation with a BioRad (Hercules, CA) Gene Pulser at a voltage of 2.5 kV, a capacitance of 25 μ F, and an impedance of 100 Ω , all according to the manufacturer's specifications. Cells were placed on LB agar containing 12.5 μ g/mL chloramphenicol, 15 μ g/mL IPTG, and 60 μ g/mL X-gal and incubated at 37°C for 24–36 h. White colonies were manually picked and transferred directly to 384-well microtiter plates containing 50 μ L LB freezing medium, with 12.5 μ g/mL of chloramphenicol in each well according to Zhang et al. [16]. After incubation at 37°C for 24 h, microtiter plates were placed in a –80°C freezer for long-term storage.

BAC DNA was isolated from 200 and 280 clones from the BamHI and EcoRI libraries, respectively, by an alkaline lysis method, digested with NotI, and separated on 1% agarose CHEF gels. The insert size was determined by comparison with a molecular-weight ladder on the gel, and an average insert size was calculated for the set of clones tested from each library.

Colonies containing individual BACs were double-spotted in high-density arrays on nylon membranes using a

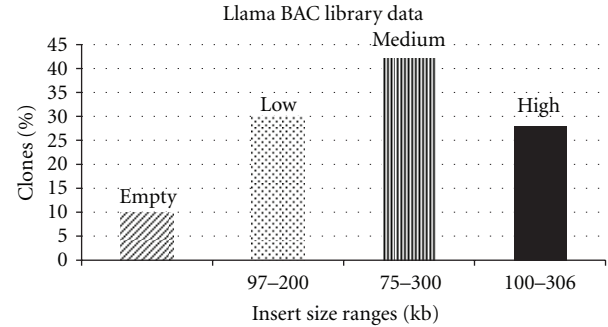


FIGURE 1: Characterization of library average insert size. The average insert sizes within the ranges of the three insert categories (low, medium, and high) for the ligations. The relative percentage of clones represented in the library that came from the size fractions is indicated. The percentage of clones without inserts is estimated.

Genetix Q-Bot at the Arizona Genomics Institute (Tucson, AZ, USA) as described in [17].

2.4. Determining Complexity of the Llama Genome. The llama genome size was determined by flow cytometry as previously described [18].

3. Results and Discussion

3.1. Creation of a Llama BAC Library. Our laboratory succeeded in developing the first and currently the only llama BAC library reported to date. Our llama BAC library contains an average insert size of 137.8 kb consisting of 196,224 clones, together equaling nearly a 9-fold genome coverage. While we calculated the size of the llama genome, by flow cytometry, to be 2.4×10^9 bp, we note that Romanini reported, in 1985, the llama genome size to be about 3.2×10^9 bp [2]. We deferred to this number and used it in all subsequent calculations.

3.2. Characterization of the Library. Clonal BAC DNA was isolated from a number of individual clones and restricted with the enzyme NotI to release the insert. The digestion products were size fractionated in an agarose gel to determine length. Insert ends were sequenced. The insert sizes range from both libraries was from 75 kb to 306 kb. The average insert size of both libraries was 137.8 kb. The inserts selected as low inserts averaged 112.5 kb and ranged from 97 to 200. The inserts selected as medium inserts averaged 134 kb and ranged from 75 to 300 kb. The inserts selected as high inserts averaged 170.6 kb and ranged from 100 to 306 kb. After examination of 300 clones from all three size ranges, we estimate the number of clones without inserts to be 10% (see Figure 1).

The average insert size of the llama BAC library analyzed in this project was 137.8 kb. The llama genome was reported to be 3.2×10^9 base pairs by Romanini [2]. This means that the library would have to contain 196,224 different inserts to cover the entire genome. Of course, this operates on the assumption that each segment of the genome was only copied into one insert. In reality, there are probably certain

segments from the genome that are copied multiple times in several inserts.

Currently, there are 511 plates containing 384 wells each representing 36 μ L of ligation reaction. Given an average insert size of 137.8 kb, this means that we currently have about 196,224 inserts represented in the library. We have an additional 950 μ L of ligation reaction to electroporate and evaluate.

We hybridized the membranes with two low-copy probes, coat color genes, bovine tyrosinase (GenBank NM_174480.3), and Lama pacos MC1R (GenBank EU135880). We observed 38 hits for the tyrosinase probe and 25 for MC1R. The average number of positive hits for these hybridizations was 31.5, which is higher than the estimated genome coverage of the library. This discrepancy may be due to an underestimation of the genome coverage of the library or it may reflect the presence of duplicate genetic loci for these genes in the llama genome.

4. Conclusions

We successfully created about a 9-fold llama BAC library. Using this library, we characterized 480 of our 196,224 clones. The llama BAC library is an important genomic resource that can be utilized for the improvement of camelid production for subsistence farmers on the Altiplano and commercial producers throughout the Andean Region. The library can also be used in the assembly of contigs for generation of a physical map. The physical map will be an important step in unraveling the ancestry of domesticated South American camelids, llama, and alpaca, to better understand its relationship with wild relatives, vicuna and guanaco, that are found throughout the region. Moreover, physical mapping would allow one to do synteny studies utilizing the DNA sequence information from camelids. We envision that the genomic resources developed for South American camelids may be useful in comparative analyses with closely related species of economic importance in Africa and Asia.

In addition, our llama BAC library can be used to physically map genes to specific loci using fluorescence in situ hybridization (FISH) technology. Previously, mapping of genomes in sheep and pigs has led to significant increases in addressing questions in genetic research. For instance, marker-assisted selection has been used to remove deleterious traits within a short time using genome mapping [13, 14]. Additionally, a wide array of research has been furthered using the RPCI-11 human BAC library created in 2001, that is, insert-end sequencing, clone fingerprinting, high-throughput sequence analysis, and diagnostic studies [19]. Our project and the future projects using our BAC library will aid in llama production throughout the world.

Acknowledgments

This project was funded, in part, by a grant from the Utah Department of Agriculture and the Brigham Young University Office of Research and Creative Activities. A special thanks to Marcie Young, David White, Dan Allen, Danielle Call, Chad Marsing, Victor Parraguez, Etel Latorre, Jaime

Ruiz, Luis Espinoza, and Cleria Espinoza for their help and insight in this work.

References

- [1] H.-B. Zhang, *Construction and Manipulation of Large-Insert Bacterial Clone Libraries*, Texas A&M University, Department of Soil and Crop Sciences and Crop Biotechnology Center, 2000.
- [2] M. G. Romanini, "The nuclear content of deoxyribonucleic acid and some problems of Mammalian phylogenesis," *Mammalia*, vol. 49, pp. 369–385, 1985.
- [3] A. Cenci, N. Chantret, X. Kong et al., "Construction and characterization of a half million clone BAC library of durum wheat (*Triticum turgidum* ssp. *durum*)," *Theoretical and Applied Genetics*, vol. 107, no. 5, pp. 931–939, 2003.
- [4] S. Noir, S. Patheyron, M.-C. Combes, P. Lashermes, and B. Chalhoub, "Construction and characterisation of a BAC library for genome analysis of the allotetraploid coffee species (*Coffea arabica* L.)," *Theoretical and Applied Genetics*, vol. 109, no. 1, pp. 225–230, 2004.
- [5] E. Y. Yoo, S. Kim, Y. H. Kim, C. J. Lee, and B.-D. Kim, "Construction of a deep coverage BAC library from Capsicum annum, 'CM334,'" *Theoretical and Applied Genetics*, vol. 107, no. 3, pp. 540–543, 2003.
- [6] C. P. Hong, S. J. Lee, J. Y. Park et al., "Construction of a BAC library of Korean ginseng and initial analysis of BAC-end sequences," *Molecular Genetics and Genomics*, vol. 271, no. 6, pp. 709–716, 2004.
- [7] M. R. Stevens, C. E. Coleman, S. E. Parkinson et al., "Construction of a quinoa (*Chenopodium quinoa* Willd.) BAC library and its use in identifying genes encoding seed storage proteins," *Theoretical and Applied Genetics*, vol. 112, no. 8, pp. 1593–1600, 2006.
- [8] W. Warren, T. P. L. Smith, C. E. Rexroad III et al., "Construction and characterization of a new bovine bacterial artificial chromosome library with 10 genome-equivalent coverage," *Mammalian Genome*, vol. 11, no. 8, pp. 662–663, 2000.
- [9] L. Cai, J. F. Taylor, R. A. Wing, D. S. Gallagher, S. S. Woo, and S. K. Davis, "Construction and characterization of a bovine bacterial artificial chromosome library," *Genomics*, vol. 29, no. 2, pp. 413–425, 1995.
- [10] C. Schelling, J. Schläpfer, A. Billault et al., "Construction of a canine bacterial artificial chromosome library for screening with PCR," *Journal of Animal Breeding and Genetics*, vol. 119, no. 6, pp. 400–401, 2002.
- [11] A. Ratnakumar, E. F. Kirkness, and B. P. Dalrymple, "Quality control of the sheep bacterial artificial chromosome library, CHORI-243," *BMC Research Notes*, vol. 3, p. 334, 2010.
- [12] A. P. Monaco and Z. Larin, "YACs, BACs, PACs and MACs: artificial chromosomes as research tools," *Trends in Biotechnology*, vol. 12, no. 7, pp. 280–286, 1994.
- [13] P. Mulsant, F. Lecerf, S. Fabre et al., "Mutation in bone morphogenetic protein receptor-IB is associated with increased ovulation rate in Booroola Mérimo ewes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 9, pp. 5104–5109, 2001.
- [14] D. Milan, J. T. Jeon, C. Looft et al., "A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle," *Science*, vol. 288, no. 5469, pp. 1248–1251, 2000.
- [15] H.-B. Zhang, *Construction and Manipulation of Large-Insert Bacterial Clone Libraries Manual*, Texas A&M University, College Station, Tex, USA, 2000.

- [16] H.-B. Zhang, S. Choi, S.-S. Woo, Z. Li, and R. A. Wing, "Construction and characterization of two rice Bacterial Artificial Chromosome libraries from the parents of a permanent recombinant inbred mapping population," *Molecular Breeding*, vol. 2, no. 1, pp. 11–24, 1996.
- [17] J. P. Tomkins, R. Mahalingam, H. Smith, J. L. Goicoechea, H. T. Knap, and R. A. Wing, "A bacterial artificial chromosome library for soybean PI 437654 and identification of clones associated with cyst nematode resistance," *Plant Molecular Biology*, vol. 41, no. 1, pp. 25–32, 1999.
- [18] K. Arumuganathan and E. D. Earle, "Estimation of nuclear DNA content of plants by flow cytometry," *Plant Molecular Biology Reporter*, vol. 9, no. 3, pp. 229–241, 1991.
- [19] K. Osoegawa, A. G. Mammoser, C. Wu et al., "A bacterial artificial chromosome library for sequencing the complete human genome," *Genome Research*, vol. 11, no. 3, pp. 483–496, 2001.