Encyclopedia

Genomic analysis of sites that present somatic mutations in mice

Created by: Katia Avina-padilla
Version received: 11 December 2018

The recognition that genomes not only contain all the genetic material of a particular organism, but also have their historical information, has increased the range of phylogenetic studies, which provide evolutionary information. The development of complete genome sequencing techniques, in conjunction with the development of bioinformatics has led to a vertiginous growth in the amount of information deposited in the data banks, and in novel tools for analysis. Genomic analyzes allow us to obtain a comprehensive study of the functioning, content, evolution and origin of genomes. It has been determined that there are dramatic rearrangements in the genomes, thus demonstrating genomic plasticity. It has been concluded that genomic rearrangements can occur as a consequence of events in which sites of small or large regions within the genome can be deleted, moved, exchanged, or inserted. Within these mentioned events are the deletions or deletions that consist in the loss of a DNA fragment of a chromosome, the size includes a range so diverse ranging from a nucleotide to large cytogenetically visible regions. There is a large number of studies that indicate that in various animal groups there are species in which genomic rearrangements occur during development whose functional relevance is still unknown. This project aims to focus on the study of recurrent deletion sites using the mouse as a study model with the main objective of identifying, characterizing and evaluating the presence of these genomic rearrangements in different genomes of mammals using the mouse genome as a study model for determine its possible functional and evolutionary relevance. The results obtained from this proposal will allow us to generate a general overview of the role played by these sequences, their possible correlation with the genome of other mammalian species will allow us to generate phylogenetic reconstructions.

a. **Background**

The recognition that genomes not only contain all the genetic material of a particular organism, but also have their historical information, has increased the range of phylogenetic studies, which provide evolutionary information. The development of complete genome sequencing techniques, in conjunction with the development of bioinformatics has led to a vertiginous growth in the amount of information deposited in the data banks, and in novel tools for analysis. Genomic analyzes allow us to obtain a comprehensive study of the functioning, content, evolution and origin of genomes. It has been determined that there are dramatic rearrangements in the genomes in mammals, thus demonstrating genomic plasticity [1]. Studies have concluded that genome rearrangements may occur as a consequence of events in which small or large regions within the genome can be eliminated, moved, exchanged, or inserted during embryonic development or during the differentiation of specific cell types[1]. Within these mentioned events are the deletions (deletions), which consist of the loss of a DNA fragment of a chromosome, the size includes a range as diverse as going from a nucleotide (point deletion) to large cytogenetically visible regions. The origin of genetic deletions can be due to different causes, infection of a virus, interaction with mutagens, or be the result of genetic recombination events. The consequences caused by these events will depend on the size of the deleted segment, the location within the genome (the number and functions of the genes that contain the deletion of genetic material).

The mouse genome was an effort of international scientific collaboration and since then it has been extremely important as a computational tool to understand the content of the human genome in multiple experimental investigations[2]. The evolution of the genome of eutherian mammals is relatively conservative. The genome of the mouse is grouped into 20 pairs of chromosomes, while humans are 23. Both are very similar, the human genome contains 2,900 million base pairs, while the mouse genome is constituted by approx. 2,600 million. There is evidence that some genes found on the same human chromosome are located on different chromosomes in mice. For example, genes GADP, TPI, LDHB, KRAS2, INT1, GL1 and LALB are grouped together on chromosome 12, whereas in the mouse these genes are distributed on chromosomes 6, 15, 10 and 4 The human chromosome 7 has homologies in seven different chromosomes of *Mus musculus*. Chromosome 11 of the mouse has homologies in at least six human chromosomes[3]. The genetic similarity

between the two species allows the genes to be compared almost directly, routes and mechanisms of human diseases, because they also occur in mice. There are examples, among others, such as the case of the murine mutation related to obesity, which is homologous to the *attractrin* gene, which encodes a glycoprotein associated with obesity in humans.

b.Hypothesis

There is a diversity of genomic rearrangements in animal species, including recurrent deletion events of specific fragments that could have a regulatory role in the normal control of cell development and / or differentiation, and whose functional relevance is still unknown. Events could result in significant alterations of said physiological processes.

The aim of this research project is to analyze potential sites of recurrent deletion present in the mouse genome and comparative genomics in other mammals through bioinformatic strategies and molecular biology techniques to determine their functional role. In order to achieve this goal we proceed to the following specific objectives. First the analysis through bioinformatic strategies a set of DNA sequences obtained from to identify potential sites of recurrent deletion. After that we determine and classify the diversity and frequency of deletions not associated with repetitive regions in mouse and other organisms. The experimental development started from a database of reading files of amplified sequences of DNA from mouse kidney and liver tissue cells that were obtained by sequencing by the Illumina Hiseq2000 / 2500 method, and which are contained in the NCBI database. The evaluation of the quality of the readings will be made using the Trimmomatic tool that will allow us to cut or adjust the readings as well as remove adapters that may be present. The criteria required for each reading will be PHRED 33. Once the sequences are selected, they will be aligned using the Segemehl software that allows us to detect the deletions, for which the genome of the mouse reported by the University of California will be used as the reference genome. in Santa Cruz version Dec.2011. Potential sites identified with recurring deletion will be processed using Linux commands and programming with Perl language to determine factors such as frequency, nature, and length.

This project aims to focus on the study of recurrent deletion sites in mammalian genomes, particularly the human genome, using the mouse as a study model with the main objective of identifying, characterizing and evaluating the presence of these genomic rearrangements and their possible functional relevance and evolutionary. The results obtained from this proposal will generate a general overview of the role played by these sequences, their correlation with the genome of other mammals, including the human genome will allow us to obtain phylogenetic reconstructions. Studies and analysis in the mouse genome plasticity during embryonic development and cell differentiation will allow subsequent analysis in comparative genomics with the human and other mammalian genomes, the identification, phylogenetic reconstruction and molecular characterization of rearrangements and genomic rearrangements will be able to determine the functional and evolutionary role.

References

1. Glenn Tesler; Pavel Pevzner; Genome Rearrangements in Mammalian Evolution: Lessons From Human and Mouse Genomes. *Genome Research* **2002**, *13*, 37-45, 10.1101/gr.757503.
2. Deanna M. Church; Leo Goodstadt; LaDeana W. Hillier; Michael C. Zody; Steve Goldstein; Xinwe She; Carol J. Bult; Richa Agarwala; Joshua L. Cherry; Michael DiCuccio; et al.Wratko HlavinaYuri KapustinPeter MericDonna MaglottZoë BirtleAna Claudia MarquesTina GravesShiguo ZhouBrian TeagueKonstantinos PotamousisChristopher ChurasMichael PlaceJill HerschlebRon RunnheimDaniel ForrestJames Amos-LandgrafDavid C. SchwartzZe ChengKerstin Lindblad-TohEvan E. EichlerChris P. PontingThe Mouse Genome Sequencing ConsortiumMouse Genome Sequencing Consortium Lineage-Specific Biology Revealed by a Finished Genome Assembly of the Mouse. *PLOS Biology* **2009**, *7*, e1000112, 10.1371/journal.pbio.1000112.
3. Ruiz García, Manuel; Álvarez, Diana . Evolución comparada de genomas animales y su relación con el genoma humano; Editorial Panamericana: Bogotá, Colombia, 2002; pp. Cap 6.

Keywords

genetic plasticity, phylogenetic reconstruction