CONCEPT AND APPLICATION OF GENE MAPPING IN ANIMAL BREEDING

Sushil Kumar, Govind Mohan, Kotresh Prasad C, RevanasidduDeginal and Anand Kumar N
Ph.D. Scholar, ICAR-National Dairy Research Institute, Karnal, Haryana-132001, India

Abstract: Gene mapping is the process of establishing the locations of genes on the chromosomes. Early gene maps used linkage analysis. Gene mapping of the genome involves partitioning this number of genes into maps of individual linkage groups and determining their linear order on each of the chromosomes. The closer two genes are to each other on the chromosome, the more likely it is that they will be inherited together. By following inheritance patterns, the relative positions of genes can be determined. More recently, scientists have used recombinant DNA (r DNA) techniques to establish the actual physical locations of genes on the chromosomes. Application of comparative gene mapping involves predicting the location of genes in one species given their locations in another species. Knowledge of the genetic maps of animals leads to the development of animal breeds that are more nutritious, productive and can better resist diseases, insects and drought. The technique can also be used in organ transplants to achieve better matches between recipients and donors, thus minimizing the risks of complications and maximizing the use of donated healthy organs, a scarce resource.
Keywords: Gene Mapping, Linkage, recombinant DNA, Application of gene mapping.

Introduction

The genetic information contained in mammalian cells is spread over a set of mitochondrial DNA and the nuclear chromosomes. Converging evidence has made it possible to estimate the mammalian genome, which may contain about fifty thousand structural genes, which may be grouped in some three thousand to fifteen thousand clusters. Gene mapping of the genome involves partitioning this number of genes into maps of individual linkage groups and determining their linear order on each of the chromosomes. On dissection of the nuclear genome for mapping can be achieved by meiotic segregation in families, segregation in somatic cell hybrids, natural tagging of particular chromosomes in individuals with chromosome aberrations, by sorting of individual metaphase chromosomes, by analysis of subchromosomal fragments using somatic cell genetics and/or genomic or gene sequencing by molecular biology techniques.

What is Gene Mapping?

Any method used for determining the location of and relative distances between genes on a chromosome. It is assignment of a locus to a specific chromosome or determining the sequence of genes and their relative distance from one another on a chromosome. It determines the relative positions of genes on a DNA molecule (chromosome or plasmid) and of the distance, in linkage units or physical units, between them.

Gene mapping is the process of establishing the locations of genes on the chromosomes. Early gene maps used linkage analysis. The closer two genes are to each other on the chromosome, the more likely it is that they will be inherited together. By following inheritance patterns, the relative positions of genes can be determined. More recently, scientists have used recombinant DNA (rDNA) techniques to establish the actual physical locations of genes on the chromosomes. Genes can be viewed as one special type of genetic markers in the construction of genome maps, and mapped the same way as any other markers.

Types of Maps:-

There are three distinct types of genetic maps i.e linkage, chromosomal and physical map
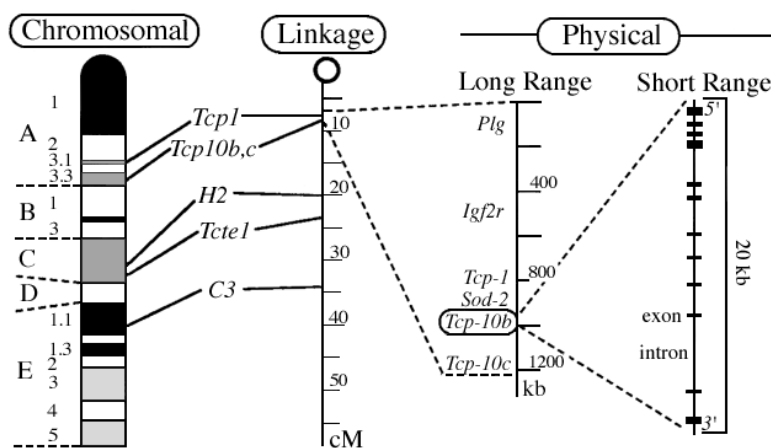

Fig 1. Types of gene maps

Linkage Map:-

The linkage map, also referred to as a recombination map, was the first to be developed soon after the re-discovery of Mendel's work at the beginning of the 20th century. Linkage maps can only be constructed for loci that occur in two or more heritable forms, or alleles. Linkage maps are generated by counting the number of offspring that receive either parental or recombinant allele combinations from a parent that carries two different alleles at two or

more loci. Analyses of this type of data allow one to determine whether loci are "linked" to each other or not. Distances are measured in Centimorgans, with one Centimorgan equivalent to a crossover rate of 1%. The linkage map is the only type based on classical breeding analysis.

Chromosome Map:-

The chromosome map (or cytogenetic map) is based on the karyotype of the mouse genome. All the mouse chromosomes are defined at the cytogenetic level according to their size and banding pattern, and ultimately, all chromosomal assignments are made by direct cytogenetic analysis or by linkage to a locus that has previously been mapped in this way. Chromosomal map positions are indicated with the use of band names. Inherent in this naming scheme is a means for ordering loci along the chromosome.

Today, several different approaches, with different levels of resolution, can be used to generate chromosome maps. First, in some cases, indirect mapping can be accomplished with the use of one or more somatic cell hybrid lines that contain only portions of the mouse karyotype within the milieu of another species' genome. By correlating the presence or expression of a particular mouse gene with the presence of a mouse chromosome or subchromosomal region in these cells, one can obtain a chromosomal, or subchromosomal, assignment. The second approach can be used in those special cases where karyotypic abnormalities appear in conjunction with particular mutant phenotypes. When the chromosomal lesion and the phenotype assort together, from one generation to the next, it is likely that the former causes the latter.

When the lesion is a deletion, translocation, inversion, or duplication, one can assign the mutant locus to the chromosomal band that has been disrupted. Finally, with the availability of a locus-specific DNA probe, it becomes possible to use the method of in situ hybridization to directly visualize the location of the corresponding sequence within a particular chromosomal band. This approach is not dependent on correlations or assumptions of any kind and, as such, it is the most direct mapping approach that exists.

Physical Map:-

Physical maps are based on the direct analysis of DNA. Physical distances between and within loci are measured in basepairs (bp), kilobasepairs (kb) or megabasepairs (mb). Physical maps are arbitrarily divided into short range and long range. Short range mapping is commonly pursued over distances ranging up to 30 kb. In very approximate terms, this is the average size of a gene and it is also the average size of cloned inserts obtained from cosmid-

based genomic libraries. Cloned regions of this size can be easily mapped to high resolution with restriction enzymes and, with advances in sequencing technology, it is becoming more common to sequence interesting regions of this length in their entirety. Direct long-range physical mapping can be accomplished over megabase-sized regions with the use of rare-cutting restriction enzymes together with various methods of gel electrophoresis referred to generically as pulsed field gel electrophoresis or PFGE, which allow the separation and sizing of DNA fragments of 6 mb or more in length. PFGE mapping studies can be performed directly on genomic DNA followed by Southern blot analysis with probes for particular loci.
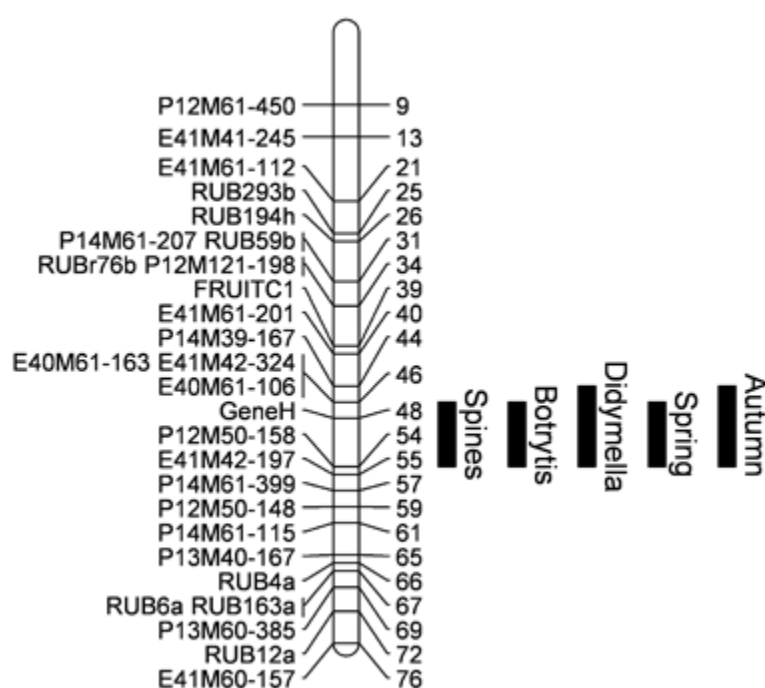
Fig 2. Chromosome map

Long-range mapping can also be performed with clones obtained from large insert genomic libraries such as those based on the yeast artificial chromosome (YAC) cloning vectors, since regions within these clones can be readily isolated for further analysis . In the future, long-range physical maps consisting of overlapping clones will cover each chromosome in the mouse genome. Short-range restriction maps of high resolution will be merged together along each chromosomal length, and ultimately, perhaps, the highest level of mapping resolution will be achieved with whole chromosome DNA sequences.

Genetic distance

Crossing over is a random event that may happen anywhere along the chromosome. If we were to consider a long segment of chromosome having loci labelled sequentially from A to

Z, it is obvious that genes A and Z are wide apart whereas genes J and K are very close together. It is much more likely that crossing over happens somewhere between A and Z than exactly between J and K. Recombination is much more likely between genes A and Z that are far apart than between genes J and K , which are very close to one another. The frequency of recombination of two genes is proportional to the distance between them. The recombination frequency can be measured in families in which the genotypes of all individuals are known. The frequency with which recombination occurs in the offspring is expressed as a percentage. Genes which are very close together (closely linked) will have a very small recombination frequency (e.g. 1%). A recombination frequency of 1% means that only one out of 100 offspring was the combination of two genes different from that in their parents. In contrast, genes that are very far apart on the same chromosome or those that are on different chromosomes are equally likely to be transmitted together or separately and so would have a recombinant frequency of 50%. This leads to the conclusion that the frequency of recombination is directly proportional to the distance between two genes - the smaller the distance, the smaller is the frequency of recombination. Thomas Hunt Morgan, a leading geneticist of the early twentieth century, first presented this theory after careful and painstaking experiments observing the frequencies of inheritance of combinations of characteristics in the fruit fly Drosophila. In recognition of his outstanding work, the unit for measuring genetic distance has been designated as the centiMorgan which is defined as the distance between two genes in which recombination occurs with a frequency of 1%. Two genes are linked if they show a recombinant frequency of less than 50%. However, if linked genes are far apart on a chromosome, it is possible that crossing-over occurs one or more times within that distance. This introduces errors in the estimation of gene distance. Hence measures of genetic distance using recombination frequencies are accurate only if the genes are closely linked i.e. if the gene distance is small. The unit of gene distance is also called a map unit. One map unit is equal to one centi-Morgan.

Methods of Gene Mapping:

Genetic Linkage Mapping:-

In organisms that reproduce sexually, meiosis breaks the parental chromosomes at random, recombines the fragments and segregates the shutted chromosomes into gametes, and thence to offspring. If two markers, A and B, lie close together on a chromosome, then it is unlikely that a meiotic break will occur between them. Hence, A and B will seldom recombine, and will usually segregate together (cosegregate) to the same gametes and hence to the same

offspring. In contrast, if markers B and Z lie at opposite ends of a chromosome, it is much more likely that a meiotic break will fall between them, whereupon they may segregate independently. Markers lying on different chromosomes will also segregate independently. Hence, the distance between any two markers is reflected by their recombination frequency: closely linked markers recombine rarely (or cosegregate often) while distant markers recombine often, and hence cosegregate no more than expected by chance – 50% of the time. This provides a way to estimate the distances between pairs of markers and, if all pairwise distances are known, a map can be made.

There is one complicating factor, however. Meiotic recombination occurs by an exchange of homologous segments between chromosomes. Hence, some way must be found to distinguish between the copies of the markers on each parental chromosome, so that the pattern of segregation can be followed. This is only possible in the case of polymorphic markers, and recombination can be detected only in those cases where the parent is heterozygous at the relevant loci. In addition, three generations (grandparents, parents and offspring) are normally required to be able to deduce completely the pattern of segregation.More often, though, genetic linkage maps are made using markers characterized at the sequence level. The earliest such markers were restriction fragment length polymorphisms or RFLPs – regions of sequence in which the alleles di er in the distribution of sites for a particular restriction enzyme.

RFLPs can be detected by digesting genomic DNA with the restriction enzyme, resolving the fragments electrophoretically, Southern blotting and probing with a labelled piece of DNA complementary to the region of interest. Different alleles then appear as bands of different sizes on the autoradiograph. RFLPs have largely been superseded in linkage mapping by microsatellites, regions of simple repeated sequence (such as CACACACA...) in which the number of repeats differs between alleles. Alleles are scored by measuring (on gels) the length of a PCR product produced by primers which anneal to the unique sequence on either side of the microsatellite. Sequence polymorphisms of a single nucleotide can be detected by PCR- or hybridization-based methods. Such single nucleotide polymorphisms or SNPs are extremely common in the human genome, occurring once every few hundred bases on average.

However, SNPs are more often identified once a region of the genome has been mapped and sequenced, rather than serving as markers for the initial genetic linkage mapping.

Radiation hybrid mapping:-

In the mid-1970s, Stephen Goss and Henry Harris discovered that, if cultured human cells were subjected to high dose of radiation (enough to fragment their chromosomes) and then fused to un-irradiated hamster cells, hybrids would be produced. Such 'radiation hybrids' would initially contain many fragments of human chromosomes (in addition to the complete hamster genome), and hence would express many human proteins. Over time, however, the unstable human fragments would be lost (and with them, the human proteins) until only a few were stably retained. If, in many such hybrids, two human proteins were often retained together, it was likely that their genes lay on the same human chromosome fragment.

HAPPY Mapping:-

In HAPPY mapping, DNA is first prepared from the species whose genome is to be mapped. To prevent unwanted mechanical breakage to the naked, fragile DNA molecules, this is normally done by embedding the living cells in agarose gel, and then treating them with a solution of detergents and proteases. These diffuse into the agarose, lysing the cells and stripping away proteins and other cellular debris, which diffuse out of the agarose; the long molecules of chromosomal DNA remain trapped and protected within the agarose. The DNA is now broken randomly, either by melting the agarose and mechanically shearing. This broken DNA is then diluted to a very low concentration, and about one hundred samples are dispensed into separate tubes. The complete set of samples is referred to as a 'mappingpanel'. Most importantly, each sample is exceedingly small, containing less than one genome's worth of DNA fragments. For example, a mammalian genome contains about 3pg (310212 g) of DNA per haploid copy; in this case, each sample would contain only 1–2pg of DNA fragments.

These minuscule samples are screened by PCR, to determine which markers (STSs or ESTs) are present in each one. Because the samples are so small, each one will contain only a randomly sampled subset of the markers, rather than the complete genome. Hence, any particular marker will be present in some, but not all of the members of the mapping panel. If two markers are close together in the genome, then they will seldom be broken apart and hence will tend to be found together in the same members of the mapping panel – they cosegregate. As the distance between markers increases, it is more likely that the random breakage will separate them, so that they lie on separate DNA fragments and are hence less likely to cosegregate. If the markers are very far apart, they will always be broken apart during the random breakage, and hence will show no particular tendency to co-segregate.

Physical mapping:-

Physical mapping involves finding a continuous series (or 'contig') of cloned DNA fragments which contain overlapping portions of the genome. The starting point for physical mapping is a library of cloned genomic fragments, normally prepared by either random mechanical breakage or partial restriction digestion of genomic DNA. (Complete restriction digestion of the DNA would mean that no clones could be found that overlapped across the restriction sites for that enzyme.) The fragments are usually cloned in bacterial hosts (normally Escherichia coli), using bacteriophage, cosmid, plasmid or other vector systems. For physical mapping of large genomes, it is desirable to use clones containing large inserts, such as P1 artificial chromosomes (PACs) or bacterial artificial chromosomes (BACs) which can carry inserts in excess of 100kb. Even larger fragments (over 1Mb) can be cloned in yeast (Saccharomyces cerevisiae) using yeast artificial chromosome (YAC) vectors; however, such clones are often unstable, undergoing deletions or internal rearrangements.

Once the library is established, overlapping clones can be identified using several different approaches. In STS content mapping, the library is screened to identify all clones that contain a specific STS marker. Screening can be done either by PCR using the appropriate primers.

Fluorescence in situ hybridization:-

Fluorescence in situ hybridization (FISH) is an elegantly direct way of observing the physical arrangement of markers along the chromosomes. Fluorescently labelled probes are hybridized to metaphase chromosomes on a glass slide, and their position observed using a fluorescence microscope. The first requirement for conventional FISH is a population of cells in which a high proportion are in metaphase. Probes are normally cloned fragments of genomic DNA or cDNA, which have been labelled. They are denatured, allowed to hybridize to the corresponding sequences of the metaphase chromosomes, and the surplus washed away. Nonspecific hybridization is normally suppressed by using an unlabelled competitor such as total genomic DNA. To detect the probes, the slide is incubated with fluorescently labelled proteins (such as avidin or anti-digoxigenin antibodies) which bind to the hapten in the probe and render it visible in the fluorescence microscope.

Application of Gene Mapping:-

Scientists have become more proficient in genetic sequencing - the detailed genetic maps that help locate the risk genes for a host of genetic diseases. The ability to investigate the root cause of diseases may one day allow researchers to develop strategies to avoid the environmental conditions that serve as triggers to disease, formulate customized drugs, and

techniques for gene therapy. Knowledge of the genetic maps of animals leads to the development of animal breeds that are more nutritious, productive and can better resist diseases, insects and drought. The technique can also be used in organ transplants to achieve better matches between recipients and donors, thus minimizing the risks of complications and maximizing the use of donated healthy organs, a scarce resource. Locations of murine and human genes are the main operational guide used in domestic animal gene mapping.

References

[1]   http://ghr.nlm.nih.gov/glossary=genemapping

[2]   https://www.genome.gov/10000715

[3]   http://www.mun.ca/biology/scarr/Large-Scale_Mapping.html

[4]   M Gillois; Gene mapping today: applications to farm animals

[5]   http://www.nptel.ac.in/courses/102103013/31