Principal of Repeated DNA

Naila Rozi, Nasir uddin khan, Zuhair Hasnain

Abstract-DNA sequences in the nuclear diploid genome usually exist as two allelic copies (on paternal and maternal homologous chromosomes . in addition to this degree of repetition , approximately 40 % of the human nuclear genome in both haploid and diploid cells is composed of sets of closely related nonallelic DNA sequences (DNA sequences families or repetitive DNA).within the considerable variety of different repetitive DNA sequences are DNA sequence families whose individual members include functional genes (multigene families)& also many examples of nongenic repetitive DNA sequence families.

Index Terms- DNA, Multigene families, homologous chromosomes

I. INTRODUCTION

The Re association kinetics of human DNA sugest 3 broad classes of DNA sequence Re association kinetics .first suggested that complex genomes, such as the human genome comprise different sequence classes on the basis of the copy number. Typically this evolves randomly shearing human DNA (eg.by sanitation) to give fragments whose average size is about 500bp & denaturing the sheared DNA by heating to separate the complementary strand of each fragment. Thereafter the DNA is cooled typically to a temperature of about 20-30 degree centigrade below the melting temperature,Tm (which marks the mid-point of the transiction between the double stranded and single stranded states of DNA heated in solution).The cool DNA renatures but the rate of disassociation depends not only on time(t) but also on the intial concentration(Cot) of that sequence(i.e the Cot value).This type of analysis has suggested that the human genome consists of roughly 3 broad sequence components:

II. SINGLE COPY (or VERY LOW COPY NUMBER) DNA (60%)

Reassociates very slowly because a single strand from a single copy sequence will require some considerable time to find a complementary partner strand given that the vast majority of DNA fragments are unrelated to it. Denaturatation of probe DNA is generally achieved by heating a solution of the labeled DNA to a temperature which disrupts the hydrogen bonds that hold the two complementary DNA strand together. The energy required to separate two perfectly complimentary DNA strand is dependent on a number of factors.

Manuscript received on January, 2013.

Naila Rozi, Department of Mathematics, Sir Syed University of Eng.&Technology University , Karachi, Pakistan.

Nasir Uddin Khan, Department of Mathematics, University of Karachi Name, Karachi,Pakistan.

Zuhair Hasnain, Department of Agronomy, University College of Agriculture, Bahauddin Zakariya University, Multan, Pakistan.

Highly repetitive reassociates very rapidly because there are numerous copies of the same sequences and the chance of quickly finding complementary partners within the mass of different Fragments are high.

III. GENE FAMILIES ENCODING PRODUCTS

The member of some gene families may not very obviously related at the DNA sequences level,but nevertheless encoded gene products that are characterized by common General function and the presence of very short conserved sequencs. In some types of gene family, the gene encoded products that are known to be functionally related in a general sense and show only vey week sequences homology over large segments without very significance conserved amino acid.

IV. IDENTIFYING CODING DNA

From the outset of the Human Genome Project there has been much debate over whether to go for an allout assault of in discriminate sequencing of all 3 billion base of the human genome.Tandom trinucleotides repeat are not infrequent in the human genome. Although there are 64 possible trinucleotides sequences.when allowance is made for cyclic permutations

$(CAG)_n=(AGC)_n=(GCA)_n$ (1)

and reading from either strand [5'$(CAG)_n$ on one strand=5'(CTG)n on the other] (2)

there are only 10 different trinucleotides repeats. Most of them are known as usefully polymorphic microsatellite markers but in addition certain repeats of CAG/CTG and CCG/GGC show anomalous behavior.

Several genes contains $(CAG)_n$ repeats within the coding sequences ,translated a polyglutamine tracts in the protein product. Typically, pthalogical all stable and non-pathological allels have 10-30 repeats, while unstable pathological have modest expansions, often in range of 40-100repeats.

V. COMPETITION HYBRIDIZATION AND COT-1 DNA

Competition (or suppression) hybridization involves blocking a potentially strong repetitive DNA signal which can be obtained when using a complex DNA probe. The labeled probe DNA is denatured and allowed to reassociate in the presence of unlabeled total genomic DNA in solution, or preferably a fraction that is enriched for highly repetitive DNA sequences. In either case, the highly repetitive DNA within the unlabeled DNA is present in large excess over the repetitive elements in the labeled probe. As a result, such sequences will readily associate with complementary strands of the repetitive sequences within the labeled probe, thereby effectively blocking their hybridization to target sequences.

Instead of using total genomic DNA as a blocking agent in hybridization, it is more effectively to use a fraction of total

genomic DNA that is enriched for highly repetitive DNA sequences, such as the *Alu*,LINE-1 and THE repeats o human DNA. From human DNA, and other mammalian DNA where the genome size id much the same as that of the human genome, the latter usually involves preparing a fraction of DNA known as COT-1 DNA. Total purified human genomic DNA is sonicated to an average length of about 400bp, denatured by heating, then allowed to renature in 0.3 M NaCl at 65°C at a starting concentration of χ moles of nucleotides per liter for a time of t sec, where $\chi t = 1.0$.

• Several genes contain (CAG) repeats within the coding sequence, translated as polyglutamine tracts in the protein product. Typically, the stable and non-pathologic alleles have 10-30 repeats, while unstable pathological alleles have modest expansions, often in the range of 40-100 repeats. Transcription and translation of the gene are not affected by the expansion.

• Some (CGG) repeats in noncoding sequences can expand massively from a normal copy number of 10-50 up to hundreds or thousands of repeats. By unknown means, the expanded repeats affect DNA methylation and chromatin structure, producing inducible chromosomal fragile sites.

• Uniquely, a (CTG) repeat in the 3' untranslated region of the myotonic dystrophy kinase gene(DMK) at 19q13 has 5-35 repeats units in normal people, but up to 2000 units in people with myotonic dystrophy (DM, MIM 160900). There is perfect correlation between the repeat expansion and the disease, even though the repeat has no evident effect on transcription or the structure of the gene product.

VI. CONCLUSION

At first sight it may seem surprising that a program designed to calculate scores can also calculate genetic risks –but in fact the two are closely related Linkage analysis programs are general-purpose engines for calculating the likelihood of a pedigree, given certain data and assumptions. For calculating the likelihood of linkage we calculate the ratio.These probes contains the common core sequences of a hypervariable dispersed repetitive sequence (GGGCAGGNG), first discovered by *Jeffreys et all*(1985) in the myolobin gene. They give an individual –specific fingerprint of bands when hybridized to Southern blots. Their chief disadvantage is that it is not possible to tell which pairs of bands in a fingerprint represent alleles.Sequence-tagged sites are important mapping tools simply because the presence o that sequence can be assayed very conveniently by PCR.

REFERENCES

[1] Naila Rozi, Nasir Uddin Khan,2nd ABRC 2009Conference of Comsat Institute Lahore on 22-23 october,2009(Geometric Patterns of DNA Help In Signal Processing

[2] Naila Rozi and Nasir Uddin khan, *Journal of Life Sciences*, U.S.A(ISSN-1934-7391),Termination of DNA replication and role of enzymes in recombination Vol#5,Number2,February2011.

[3] Naila Rozi, Nasir Uddin Khan, ISSN# 2073 7122.*JournalofIJCRB*,Jan.2011,vol#2,page#528 Improve Hidden Markov Model In DNA SequencesHMM).

[4] Naila Rozi, Nasir Uddin Khan2ND ICECS Conferenc,28-30DEC2009Proceeding IEEE,DUBAI,U.A.E,Number Patterns Geometric representation of DNA, ISBN978-7695-3937-9,page#343-344G.

First Author Naila Rozi PhD Assistant Professor Department of Mathematics10 Research Publication member of ISOSS.

Second Author Nasir uddin Khan Provice Chancellor University of KarachiProfessor Department of Mathematics 35 publications.

Third Author Zuhair Hasnain is PhD research student in Agriculture Agronomy. He is recently back from The International Rice Research Institute (IRRI), Philippines, and worked with C4 Rice Project there. He is also the active member of the Editorial Advisory Baord of the Journal of Natural Resources and Development, Chile. While working with rice, his expertise includes nutrient and water management, Rice Agronomy, and major field crop production.

Zuhair is also Dignity Advisor at IRRI-Philippines and a brief overlook about him can have from the following URL:

http://irri.org/index.php?option=com_k2&view=item&id=11718:hasnain-zuhair&lang=en

http://scholar.google.com.pk/citations?user=Y8kz0xkAAAAJ