PUB: 60/2011

Review Article

Chromosomes to Genes: Implications of Array Comparative Genomic Hybridization (array-CGH) in Medical Practice

Krishna Murthy D S

Genomics and Central Research Laboratory, Sri Devaraj Urs Academy of Higher Education and Research, Kolar, Karnataka

HUMAN GENOME

Every cell has an estimated 30,000 genes that function in an orderly manner and maintain the complex cellular functions such as DNA replication, repair and gene expression. Each and every cell in the human body carries an genetic information in DNA, of which genes are specific parts that encode for proteins to allow biological activity to occur. The activity/inactivity of a gene(s) can be measured using microarrays, which can probe tens of thousands of genes simultaneously.

The DNA maintains its structural integrity through every cell division. Any alteration in these vital functions leads to a disease condition. Identification of the exact nature and cause of fundamental basis of genetic diagnosis, disease prognosis, decision making of a treatment protocol and development of a new therapeutics. With the completion of the Human Genome Project and simultaneous advancement in technology, the understanding of the genetic basis of these has improved rapidly, and novel genetic diseases are being identified as well. Recent advances biotechnology have contributed immensely in precise and rapid diagnosis of genetic disorders which has huge impact on human health care and development. This in turn facilitates in better patient care, genetic counseling and disease prevention by selective screening of high risk population, leading to a healthy community, as well as reduce the economic burden of the society with respect to patient care and management.

CHROMOSOMES AND DISEASE

Genomic imbalance due to the loss or gain of a chromosome (aneuploidy Ex. Trisomy 21, 13, 18; 47,XXY; 45,X) and loss or gain of small chromosomal segments (segmental aneusomy) is the major cause of complex genetic conditions such as mental retardation, developmental delay, learning disability, dysmorphism, autism, multiple congenital anomalies, history of repeated pregnancy loss and cancer. An English Physician John Langdon Down first described the clinical features in Down syndrome in the year 1866 as a distinct clinical condition with mental disability. But the genetic cause of this syndrome was known only in the year 1959, when a French Human Geneticist Jerome Lejeune demonstrated for the first time that it was due to an extra chromosome G (later confirmed to be Trisomy 21). Down syndrome is a well recognized condition and is known to be the most common cause of mental retardation. After this finding, chromosomal basis of disease was identified such as Turner syndrome (45,X), Klinefelter syndrome (XXY), Patau syndrome (trisomy 13), Edward syndrome (trisomy18), a

number of microdeletion syndromes such as DiGeorge, Prader-Willi, William's syndrome and so on. It occurs in 0.6% of newborn population. Whole genome analysis by karyotyping is the method of choice for routine genetic of complex diseases.

CHROMOSOME TO GENE

Mental retardation affects 3% of the population and a considerable number of them can be explained by the presence of gross chromosomal abnormalities. 10-15% of abortions and still-born with malformations are due to chromosomal abnormalities, and about 30% of children with malformation or multiple congenital anomalies have chromosome abnormality that lead to infant mortality or long term morbidity. However, the fact remains that 50% of them remain undiagnosed and are classified as idiopathic. This is because of the hidden complex genetic etiology tht remains undetected due to the limitations of the microscopic resolution, which is limited to 5-10 Mb. This leads to a great dilemma, particularly in genetic counseling. Late 80's saw a new revolution in genetic diagnosis when Pinkel et al in 1989 first demonstrated the use of a molecular cytogenetic technique called Fluorescence In Situ Hybridization (FISH). It is a targeted molecular technology where fluorescently labeled DNA probes are hybridized onto metaphase spreads (or interphase nuclei) of the patient that enables localization of the gene locus. It improved the resolution of cytogenetic detection to less than 1 Mb and proved to be very useful in prenatal, postnatal and cancer diagnosis to detect cryptic interstitial or sub-telomeric rearrangements. However, the method is still limited to the chromosomal and light microscopic resolution. Micro deletions or duplications are still not detected. Therefore there is a need for better detection method with high through put and highest resolution.

MICROARRAYS AND THEIR APPLICATIONS

Microarray satisfies the requirements needed for a high resolution and high trough put genetic analysis. It evolved in late 90s and developed rapidly since then. Numerous leading research centers and institutes use microarray for biological and medical research. Scientists have shown that in a single experiment once can know the expression levels of hundreds and thousands of genes spotted on the array. The fact that it enables a genome wide analysis at the highest molecular level is one of the significant technological findings of the decade. Different types of arrays are available that can measure not only expression but also copy number changes (variations CNVs), mutation and polymorphism. Custom arrays are designed to address specific questions related to a group of patients, genes or pathways. These studies have uncovered a wealth of information about the genomic imbalances and their corresponding phenotypes.

ARRAY BASED COMPARATIVE GENOMIC HYBRIDIZATION (ARRAY-CGH)

Array based comparative hybridization (array-CGH) derived from a conventional metaphase CGH works on the principle of identifying DNA copy number changes as compared to a reference

DNA. Patient and reference DNA labeled with fluorescent dye are co-hybridized onto a matrix based genomic array, either BAC (bacterial clones) or oligos. The genomic resolution depends on the physical distance of the two clones and the size of the clones used on the arrays, with BAC arrays giving a resolution of 1 Mb and oligo arrays upto 5-50 kb. ("*An oligonucleotide, or oligo as it is commonly called, is a short fragment of a single-stranded DNA that is typically 5 to 50 nucleotides long).* Array-CGH is a promising and a high throughput delivering high resolution and precise mapping of the DNA segment and sequence causing aberration. Large multi-institutional studies by array-CGH have shown that 25% of the cases that are classified as idiopathic and with normal karyotypes, are indeed due to submicroscopic genomic imbalance (CNVs). Array-CGH has been the technique of choice in the diagnosis and classification of tumors, risk assessment and disease prognosis.

COPY NUMBER VARIATIONS (CNVS)

Some large data bases are established from multicentral studies on copy number variants in the normal population (polymorphisms). Studies by Roden and Feuk et al (2006) for the first time constructed a first generation copy number variant (CNV) map for the human genome through a study of four populations with ancestry of Europe, Africa or Asia (the HapMap collection). Using the DNA reference sequence as a guide, genome scanning technologies, such as microarray-based comparative genomic hybridization (array-CGH) and genome-wide single nucleotide polymorphism (SNP) platforms, have now enabled the detection of a previously unrecognized degree of larger-sized (non-SNP) variability in all genomes. This heterogeneity can include copy number variations (CNVs), inversions, insertions, deletions and other complex rearrangements, most of which are not detected by standard cytogenetics or DNA sequencing. Although these genomic alterations, collectively termed as structural variants or polymorphisms, have been described are now known to be more global in occurrence. CNVs can contain entire genes and their number can correlate with the level of gene expression. It is also plausible that structural variants may commonly influence nearby genes through chromosomal positional or domain effects. They describe the prevalence of structural variants in the human genome and how they might influence phenotype, including the continuum of etiologic events underlying monogenic to complex diseases. All these progresses have set the stage for a golden era of combined microscopic and sub-microscopic (cyto-genomic)-based research of chromosomes leading to a more complete understanding of the human genome.

SINGLE NUCLEOTIDE POLYMORPHISM (SNP) AND MUTATION ARRAYS

Monogenic conditions are caused by the mutations of a single gene and follow a definite Mendelian pattern of inheritance. However, most of the human traits or disease conditions are polygenic, that is they involve a complex interaction of several genes and environmental factors. Identification of the genetic basis of

these complex traits/conditions, such as color of the eyes or hair and disease conditions like developmental delay, learning disabilities, autistic behavior, dysmorphism, and mental retardation is difficult. Inheritance of these complex traits is based on genome wide variant patterns. Single nucleotide polymorphism (SNP) analysis allows the detection of polymorphic markers related to a disease gene. Population based association studies are more useful in identifying the genetic markers linked to a disease, that require genotypes of a large number of polymorphisms, SNPs across the genome, each of which is tested for an association with a phenotype.

Monozygotic twin studies in complex condition like autism show 40-60% concordance for autism and 70-90% for Autistic Spectrum Disorder (ASD) as compared to 0-25% concordance in di-zygotic twins. This suggests a genetic basis of autism and ASD. Large linkage and association studies have suggested several susceptibility genes (AUTS-13 and AUTSX1-3) on chromosomes 1,2,3,7,12,13,17,21 and X respectively, may be associated to contribute to the phenotype.

EXPRESSION ARRAYS

*(**"Gene expression"** is the term used to describe the transcription of the information contained within the **DNA**, the repository of genetic information, into messenger RNA (mRNA) molecules that are then translated into the proteins that perform most of the critical functions of cells")*

Researchers have shown that expression arrays provide us with enormous information on the gene expression profile, which is the functional genome of a tissue. This information is now used efficiently in disease sub-typing, diagnosis and prognosis, drug discovery and designing individualized treatment protocols (personalized medicine). A break through example is the development of an expression array for breast cancer diagnosis and prognosis detection by Agendia using Agilent oligo-array technology.

An expression array technique has its application in Drug development, drug response and therapy development.

Microarray can help in disease classification, prognosis and treatment options. The expression profile gives information about the disease subtype, stage of the disease and the biological pathways involved in the disease. Pharmaceuticals are using microarrays to study the possible drug targets.

IMPLICATIONS OF MICROARRAYS IN HEALTH CARE

Array-CGH also known as chromosomal microarray is equivalent to performing hundreds of sub-telomeric and locus specific FISH tests simultaneously. It reliably detects deletions and duplications, not seen by metaphase FISH. Thus Chromosomal Micro Array offers a high throughput alternative in detecting and characterizing genomic imbalances (CNVs). Targeted arrays will enable in the detection of clinically relevant genetic conditions. Microarray technology is an integral part of drug discovery. It is used for target identification in investigating the gene expression profile of disease. Interpretation and genetic counseling is

challenging while dealing with CNVs.

India is a vast country with a large and diverse population, which has great potential in health care system development- diagnostics and human genetic research. However, Genetic diagnostic services are lagging far behind compared to developed countries. Genetic diagnostic being a complex specialty, extra caution is administered in the western countries in establishing and providing these services.

REFERENCES

1. Tonnies H. Modern molecular cytogenetic techniques in genetic diagnostics. Trends in Mol Med. 2002;8:246-50.

2. McNeil n and Ried T. Novel molecular cytogenetic techniques for identifying complex chromosomal rearrangements: technology and applications in molecular medicine. Expert reviews in Molecular medicine. http://www.ermm.cbcu.cam.ac.uk (accessed from).

3. Collins FS, Green ED, Guttmacher AE and Guyer MS. A vision for the future of genomics research. *Nature*. 2003;422:835-47.

4. Dominique FC Smeets M. Historical prospective of human cytogenetics: from microscope to microarray. Clinical Biochem. 2004;37:439-46.

5. Down JL. Observations on an ethnic classification of idiots. Lond Hosp Clin Lect Rep. 1866; 3: 259-62.

6. Lejeune J, Gautier M and Turpin R. Etude des chromosomes somatiques de neuf enfants mongoliens. C.R. Hebd Sciences Acad Sci. 1959;248:1721-22.

7. Shaw-Smith C, Redon R, Rickman L, Willatt RML, Fiegler H, Firth H, et al.
Microarray based comparative genomic hybridisation (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. *J Med Genet*. 2004;41:241-48.

8. Sagoo, GS, Butterworth AS, Sanderson S, Shaw-Smith C, Higgins JP, Burton H.
Array CGH in patients with learning disability (mental retardation) and congenital anomalies: updated systematic review and meta-analysis of 19 studies and 13,926 subjects. *Genet Med*. 2009;11:139-46.

9. Shaw-Smith C, Redon R, Rickman L, Rio M, Willatt L, Fiegler H, et al. Microarray-based comparative genomic hybridisation (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. J. Med. Genet. 2004);41:24148.

10. Rosenberg C, Knijnenburg J, Chauffaille Mde L, Brunoni D, Catelani AL, Sloos W, et al. Array CGH detection of a cryptic deletion in a complex chromosome rearrangement. Hum Genet. 2005;116:390-94.

11. Pinkel D, Segraves R, and Sudar D. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet. 1998;20:107-211.

12. Pinkel D and Albertson DG. Array

comparative genomic hybridization and its applications in cancer. Nat. Genet. 2005;37: S11S17.

13. Lisenka ELM, Vissers JA, Kessel VAG and Brunner HG. Identification of disease genes by whole genome CGH arrays. Human Molecular Genetics, 2005;14(2).

14. Redo R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, et al. Global variation in copy number in the human genome. *Nature*. 2006; 444:444-54.

15. Lee C, Iafrate AJ and Brothman AR. Copy number variations and clinical cytogenetic diagnosis of constitutional disorders. *Nat Genet.* 2007; 39:S48-54.

16. Barrett MT, Scheffer A, Ben DA, Sampas N, Lipson D, Kincaid R, et al. Comparative genomic hybridization using oligonucleotide microarrays and total genomic DNA. Proc. Natl Acad. Sci. USA. 2004;101:1776570.

17. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, et al. Large-scale copy number polymorphism in the human genome. Science. 2004;305:52528.

18. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, et al. Detection of large-scale variation in the human genome. Nat. Genet. 2004;36:94951.

19. Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, et al. Segmental duplications and copy-number variation in the human genome. Am. J. Hum. Genet. 2005;77:7888.

20. The International HapMap 3 Consortium. Integrating common and rare genetic variation in diverse human populations. Nature. 2010;467:5258.

21. Rauch A, Ruschendorf F, Huang J, Trautmann U, Becker C, Thiel C, et al. Molecular karyotyping using an SNP array for genome wide genotyping. J. Med. Genet. 2004;41:91622.

22. Sachidanandam R, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature. 2001:*409: 928-33.

23. Albertson D and Pinkel D. Genomic microarrays in human genetic disease and cancer. Hum Mol Genet. 2003;12:145-52.

24. Van de Viyer M, et al. A gene expression signature as a predictor of survival in breast cancer. N Eng J Med. 2002; 347:1999-2009.

25. Green ED, Guyer MS & National Human Genome Research Institute. Charting a course for genomic medicine from base pairs to bedside. Nature. 2011;470:204.

Email: dskmurthy50@gmail.com

Source of Support: Nil Conflict of Interest: Nil