Genome Analysis and Human Health: A Critical Appraisal

Leena Rawal[1], Neeta Sehgal[2] and Sher Ali[1,*]

[1]*Molecular Genetics Laboratory, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi-110067, India*

[2]*Department of Zoology, University of Delhi, New Delhi-110007, India*

Abstract: Completion of the human genome project and availability of the resultant sequences pampered the optimism of the scientists and policy makers both, as it was construed that the same would fuel the major developments in the areas of disease diagnosis and gene therapy. The access to genome information has indeed made significant contributions in the health care systems world over. However, complexity of the genetic diseases coupled with innate ethnic variations posed technical impediments. This then synergized the development of a series of advancements useful for genome analysis. In the present article, we provide a critical appraisal of some of the important and complex genetic diseases and the currently available technologies used in the context of normal and diseased genomes. We have also attempted to assess the therapeutic impact of RNAi technology in the context of genome analysis and human health. The quest to fight with genetic anomalies supported by ever unfolding newer technologies is envisaged to provide remedial measures to a large number of diseases in the forthcoming future.

Keywords: DNA based diagnosis, Genetic diseases, Genome analysis, Whole genome sequencing.

1. INTRODUCTION

Since the time DNA structure was proposed [1] as the master regulator to all biological lives on earth, attempts were made to unravel its related mysteries.The most important endeavors undertaken in recent years has been the Human Genome Project (HGP), an internationally coordinated undertaking that has culminated into the availability of the entire human genome sequence in the public domain. The resultant outcome of HGP is the introduction of “new genetics” i.e. the ability to identify gene mutations, regulation, modulation and in certain instances, even copy number variation. With this new knowledge, we now know that DNA of all humans is more than 99.9% identical and the human genome contains only about 20,000-25,000 genes. The exact number of genes perhaps would never be known as many genes undergo alternate splicing resulting in generation of several mRNA transcripts. After the DNA code was deciphered for human, the genome-related researches across the species gained momentum.

Genome research pampered our optimism as it was envisaged that this new knowledge would enhance our understanding about the mechanisms of genetic diseases benefiting not only individuals but also the overall public health care system. Since disease prevalence is not uniform across the different regions of the globe, it is therefore important to set the research priority keeping in view the regional epidemiological data. This then would enable us to narrow down the search of causative factor of a given disease in a specific region of the world.

The availability and the integration of genetic information have been the driving forces towards our understanding of normal and abnormal genomes. And now with the human genome sequence nearly completed with over 99% accuracy, determining the precise effect of a gene on disease will become easier. Gene influences every aspect of human health though its density varies across the chromosomes (see Table **1**) [2]. The perceived role of genetics in public health is changing so also the definition of the genetic diseases. Defects in the genetic makeup of a person usually are the cause of “Genetic disease”. Along with this, environment tends to play equally important role. It can either be inherited or can arise from a sporadic mutation acquired during a person’s lifetime. Today, little can be done to treat, let alone cure these diseases. However, information on a gene is mandatory for possible therapeutic intervention.

Technological advances for studying gene expression and functions using cytogenetics tools coupled with recombinant DNA technology allow raw DNA sequence to be converted into wealth of information useful for biological systems. Emergence of newer dimensions on genomics has considerable potential rationalizing gene interaction for a given biological system. Conceptually, advances in genetical knowledge fuelled by the technology could be used to prevent diseases creating much healthier gene pool.

*Address correspondence to this author at the Molecular Genetics Laboratory, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi-110067, India, Tel: 011 26703753; Fax: 011 26742125;
E-mail: alisher@nii.ac.in, sheralib5@hotmail.com

Table 1: A Brief Overview of the Genes on Different Human Chromosome Based on Online Inheritance in Man (OMIM) Database

S.No.	Gene Status	Number of autosomal genes	Number of X-linked genes	Number of Y linked genes	Number of mitochondrial genes	Total number of genes
1	Genes with known sequences	11752	541	48	37	12378
2	Genes with known sequences and phenotype	353	30	N/A	N/A	383
3	Genes with phenotype description and known molecular basis	2119	199	2	26	2346
4	Gene with Mendelian phenotype/locus and molecular basis unknown	1482	136	5	0	1624
5	Other genes with mainly phenotype with suspected Mendelian basis	1943	140	2	0	2085
6	Total	17649	1046	57	63	18816

OMIM database is constantly annotated. Thus, the figures shown here may not correspond to current updates.

This article provides a brief overview on the available key technologies and upcoming molecular approaches useful for genome analysis. Most of the molecular approaches are poised to become an integral part of the routine diagnosis for better management of the genetic anomalies.

2. CURRENT STATUS OF HUMAN GENOME AND GENETIC DISEASES

In the early 1990s, when the second 5-year plan for the HGP with sizable quantum of budget was finalized, common disorders were presented as the future target of genome research. Today, even after more than 15 years, despite billions of dollars spent on genome-wide association studies (GWAS), fewer genetic risk factors for common diseases have been identified. Thus, on the face of it, the enthusiasm for a large scale GWAS has started dwindling.

Genes are responsible for producing proteins that allow cells to perform a variety of functions. Genetic disorders are discrete events affecting set of cells related to each other. A person normally inherits two copies of a gene, one from each parent or an abnormal version of both the copies leading to disorder(s). Genetic disorders involve a permanent change (or mutation) in the genome encompassing chromosomes, mitochondria, a single gene or several (multiple) genes. A brief overview of such genetic anomalies is mentioned hereunder.

2.1. Chromosomal Abnormalities

Chromosomes, distinct structures made up of DNA and protein, are located in the nucleus of each eukaryotic cell. Chromosomal disorders refer to an excess or deficiency of a whole chromosome or its part thereof. As the chromosomes are carriers of genetic material, abnormalities in chromosome structure as missing or extra copies or gross breaks and rejoining (translocations), can result in diseases. Literature is full of reports showing genes, their locations and linked genetic disorders providing a bird's eye view on the abnormal genome (Table **2**). Some genetic disorders are indeed very uncommon. In fact, one in every 200 babies born falls in the category of rare chromosome disorder. Likewise, one in every 1000 babies having symptoms from early childhood is affected when they grow. Chromosomal problems are common, affecting about 0.7 percent of live born infants and account for early miscarriages [3]. A number of complex genetic diseases have been described together with genotype and resultant phenotype. In the process, several markers have been developed (see Table **3**).

Chromosomal anomalies have been known for the past half a century and now despite much acclaimed advancements, no clear cut understanding is still available on a large number of diseases. In other words, several chromosomal abnormalities have still remained unfathomed and no gene therapy whatsoever seems to be in sight.

2.2. Mitochondrial DNA Linked Disorders

Mitochondria found in the cytoplasm of plant and animal cells are small organelles involved in cellular respiration. This organelle generates energy for cellular processes through oxidative phosphorylation and the generation of adenosine triphosphate (ATP). These cellular structures contain their own DNA distinct from

Table 2: Genetic Disorders Related to Different Human Chromosomes

S.No.	Chromosome Number	Size (Mb)	Sequence Determined	Number of genes present	Linked genetic diseases	Genes Involved
1	1	240	90%	3000	Porphyria cutanea tarda Gaucher Disease Glaucoma Prostate Cancer	*UROD* *GBA* *GLC1A* *HPC1*
2	2	240	95%	2500	Essential Tremor Colon Cancer Waardenberg Syndrome	*ETM2* *MSH2, MSH6* *Pax3*
3	3	200	95%	1900	von Hippel Lindau Colon Cancer Lung Cancer Essential Tremor	*VHL* *MLH1* *SCLC1* *ETM1*
4	4	190	95%	1600	Ellis-van-Creveld Huntington Disease Achondroplasia Narcolepsy Parkinson Disease Fibrodyslapsia ossificans progressiva	*EVC* *HD* *FGFR3* *NRCLP* *SNCA* *FOP*
5	5	180	95%	1700	Steroid 5-alpha reductase 1 Cockayne Syndrome Spinal muscular atrophy Diastrophic Dysplasia	*SRD5A1* *CKN1* *SMN1* *DTD*
6	6	170	95%	1900	Spinocerebellar ataxia Haemochromatosis Diabetes Congenital adrenal hyperplasia Epilepsy	*SCA1* *HFE* *IDDM1* *CYP21A* *EPM2A*
7	7	150	95%	1800	Diabetes Williams Syndrome Cystic Fibrosis Obesity	*GCK* *ELN* *CFTR* *OB*
8	8	140	95%	1400	Werner Syndrome Burkitt Lymphoma	*WRN* *MYC*
9	9	130	85%	1400	Malignant Melanoma Freidrich's ataxia Tangier Disease Tuberous Scelrosis Chronic myeloid leukemia	*CDKN2* *FRDA* *ABC1* *TSC1* *ABL*
10	10	130	95%	1400	Rafsum Disease Gyrate Atrophy	*PAHX* *OAT*
11	11	130	95%	2000	Harvey Ras-oncogene Diabetes Long QT Sydrome Best Disease Multiple Endocrine Neoplasia Ataxia telangiectasia	*HRAS* *IDDM2* *LQT* *VMD2* *MEN1* *ATM*

(Table 2). Continued.

S.No.	Chromosome Number	Size (Mb)	Sequence Determined	Number of genes present	Linked genetic diseases	Genes Involved
12	12	130	95%	1600	Zellwager Syndrome Phenylketonuria	*PEX2*
13	13	110	80%	800	Breast Cancer Autosomal recessive neurosensory deafness Retinoblastoma Wilson disease	*BRCA2* *CX26* *RB1* *ATP7B*
14	14	110	80%	1200	Alzeimer Disease Alpha-1-antitrysin deficiency	*PS1 (AD3)* *SERPINA-1*
15	15	110	80%	1200	Prader-Wili Syndrome Angelman Syndrome Marfan Sydrome Tay-Sachs disesae	*SNRPN* *UBE3A* *FBN1* *HEXA*
16	16	90	85%	1300	Alpha thalassemia Polycystic Kidney Disease Familiar Mediterranean fever	*HBA1, HBA2* *PKD1* *FMF*
17	17	90	85%	1300	Tumor Suppressor Protein Charcot-marie-tooth Syndrome Breast Cancer	*p53* *CMT1A* *BRCA1*
18	18	70	85%	1300	Niemann-Pick Disease Pancreatic Cancer	*NPC1* *DPC4*
19	19	60	85%	1700	immunodeficiency Maple Syrup Urine Disease Myotonic Dystrophy Atherosclerosis	*JaK3* *BCKDHA* *DMPK* *APOE*
20	20	60	90%	900	Severe combined immunodeficiency	*ADA*
21	21	40	70%	400	Amylotrophic lateral sclerosis Autoimmune Polyglandular Syndrome	*SOD1* *APS1*
22	22	40	70%	800	Glucose galactose malabsorption DiGeorge Syndrome Neurofibromatosis Chronic Myeloid Leukemia	*SGLT1* *DGS* *NF2* *BCR*
23	X	150	95%	1400	Paroxymal nocturnal hemoglobinuria Duchenne Muscular Dystrophy Menkes Syndrome Alport Syndrome Lesch-Nyhan Syndrome Fragile X Syndrome Adrenoleuko Dystrophy	*PIG-A* *DMD* *ATP7A* *COL4A5* *HPRT1* *FMR1* *ALD*
24	Y	50	50%	200	*SRY* determining factor, *AZFa, AZFb AZFc, DAZ, DBY, UPS9Y* and several other involved in normal spermatogenesis.	The entire MSY regions is prone to alteration

the DNA contained in the cell nucleus. In recent years, more than 20 hereditary disorders have been shown to result from the mutations in mitochondrial DNA [4]. It is inherited maternally and does not recombine. Thus, mutations gained once remain largely unfixed. Mitochondrial DNA is transmitted maternally to the offsprings. This relatively rare type of genetic disorder is caused by mutations in the non-chromosomal DNA of mitochondria. Such disorders can appear at any age with a wide variety of non-specific symptoms. These disorders include metabolic disturbances, developmental delay, blindness, hearing loss, heart

Table 3: Examples of Some known Genetic Disorders in the Human

S.No.	Genetic Disorder	Molecular and Cellular Defects	Prevalence	Features
		I *Chromosomal Disorders*		
1	Down syndrome	It is a genetic disorder caused by the presence of all or part of a third copy of chromosome 21. Down syndrome is the most common chromosome abnormality in humans. Trisomy 21 (47,XX,+21) is caused by a meiotic nondisjunction event.	1 in 1000 births in United States	Mental retardation, characteristic facial features, fingers that curl inward, and there usually is only a single palmar (*i.e.,* simian) crease
2	Turner's syndrome	Characterized cytogenetically by a monosomy of the X chromosome, the presence of an abnormal X chromosome	1 of every 2500 live births, 99% pure XO fetuses are spontaneously aborted in the first trimester.	The females are usually short stature, failure in menstruation and show no signs of secondary sex characteristics
3	Klinefelter's syndrome	It is a condition of testicular dysgenesis accompanied by the presence of one or more X chromosomes in excess of the normal male XY complement. Most males with Klinefelter's syndrome have one extra X chromosome (47, XXY). In rare cases, there may be more than one extra X chromosome (48, XXXY). The presence of the extra X chromosome in the 47,XXY male results from nondisjunction during meiotic division in one of the parents	1 in 500 -1000 births in United States	Enlarged breasts, sparse facial and body hair, small testes, and the inability to produce sperm. Regardless of the number of X chromosomes present, the male phenotype is retained. The infant usually has normal male genitalia, with a small penis.
		II *Mitochondrial Disorders*		
4	Leigh disease	It is caused by deficiency of the pyruvate dehydrogenase complex (PDHC), most commonly involving a PDHC subunit which is encoded by an X-linked gene	1 person per 2000 in Europe	Proximal muscle weakness, sensory neuropathy, developmental delay, ataxia, seizures, dementia, visual impairment due to retinal pigment degeneration, poor sucking ability, loss of head control and motor skills, loss of appetite, vomiting, irritability, continuous crying (in infants) and seizures.
5	Myoclonic epilepsy with ragged red fibers	It is caused by a maternally-inherited mutation at position 8344 in the mitochondrial genome in over 80% of cases. This point mutation disrupts the mitochondrial gene for tRNA-Lys and disrupts synthesis of proteins essential for oxidative phosphorylation.	1/400,000 in Europe	Myoclonic seizures, cerebellar ataxia, mitochondrial myopathy
6	Leber's hereditary optic neuropathy	It is a mitochondrially inherited (transmitted from mother to offspring) degeneration of retinal ganglion cells (RGCs) and their axons that leads to an acute or subacute loss of central vision; this affects predominantly young adult males. Mutations in the *MT-ND1, MT-ND4, MT-ND4L,* and *MT-ND6* genes cause Leber hereditary optic neuropathy.	1:30,000 to 1:50,000 in Europe.	Painless, subacute, bilateral visual loss, with central blind spots (scotomas) and abnormal color vision
7	Mitochondrial Encephalomyopathy	It is a condition that affects many of the body's systems, particularly the brain and nervous system (encephalo-) and muscles (myopathy)	1/2000 in Europe	Muscle spasms (myoclonus), impaired muscle coordination (ataxia), hearing loss, heart and kidney problems, diabetes, epilepsy, and hormonal imbalances
8	Chronic progressive external ophthalmoplegia	A mutation is located in a conserved region of mitochondrial tRNA at nucleotide 3243 in which there is an A to G nucleotide transition.	1 out of 11,000 preschool children in Swedish population	Progressive weakness of the extra ocular muscles
9	Kearns-Sayre syndrome	It is a severe syndromic variant of chronic progressive external ophthalmoplegia (abbreviated CPEO), characterized by isolated involvement of the muscles controlling eyelid movement (levator palpebrae, orbicularis oculi), and those controlling eye movement (extra-ocular muscles).	1-3/100,000 births in United States	Progressive weakness, retinal pigmentation, heart involvement (cardiomyopathy, cardiac conduction defect), skeletal muscle myopathy, intestinal disorders, hormonal deficit (hypoparathyroidism, diabetes) and renal failure
		III *Single-gene disorders*		
		III a. Autosomal Dominant The anomaly appears in every generation. Each child of an affected parent has a 50% chance of inheriting the disease.		
10	Achondroplasia	It is short-limb dwarfism. Mutation in Fibroblast growth factor receptor 3 (FGFR3).	1 in 25,000	Short stature, decreased muscle tone, bowed legs, spinal stenosis, spine curvature (kyphosis) and prominent forehead (frontal bossing)

(Table 3). Continued.

S.No.	Genetic Disorder	Molecular and Cellular Defects	Prevalence	Features
11	Adult polycystic kidney disease	There are three genetic mutations *PKD-1*, *PKD-2*, and *PKD3* with similar phenotypical presentation	1/2500 of European origin	PKD is characterized by the presence of multiple cysts (hence, "polycystic") typically in both kidneys. The disease can also damage the liver, pancreas and even heart and brain in some cases
12	Huntington's chorea	It is a neurodegenerative disorder, caused by an autosomal dominant mutation in one of the two copies of a gene called Huntingtin. The *HTT* gene is located on the short arm of chromosome 4 at 4p16.3	1/10,000 of European origin	The symptom relates to mood or cognition problems coupled with lack of coordination and an unsteady gait, uncoordinated, jerky body movements, decline in mental, behavioral and psychiatric abilities, affecting heart and reducing overall life span.
13	Familial hypercholesterolemia	Characterized by premature atherosclerosis. Patients have mutations in the *LDLR* gene that encodes the LDL receptor protein. The *LDL* gene is located on the short arm of chromosome 19 (19p13.1-13.3).	1:122 data based on French and Canadians subjects	Severe cholesterol deposition such as yellowish patches around the eyelids (xanthelasma palpebrarum), the outer margin of the iris (arcus senilis corneae) and in the form of lumps in the tendons of the hands, elbows, knees and feet
14	Marfan's syndrome	Connective tissue disorder with abnormalities in skeletal, ocular, cardiovascular systems. It affects *fibrillin I,* a major component of microfibrils found in the extracellular matrix. It is located on chromosome 15q21.	2 to 3 per 10,000 individuals	Severe skeletal deformities that include a long, thin body with exceptionally long extremities and long, tapering fingers, hyper extensible joints; and a variety of spinal deformities, including kyphoscoliosis. Chest deformity
15	Von Recklinghausen disease (Neurofibromatosis (NF) type 1)	It is a condition involving neurogenic tumors that arise from Schwann cells and other elements of the peripheral nervous system	1 in 3500 individual	Skeletal lesions such as scoliosis and erosive bone defects, increased risk for development of other nervous system tumors such as meningiomas, optic gliomas and pheochromocytomas
16	Acoustic Neurofibromatosis (Neurofibromatosis type 2)	Characterized by tumors of the acoustic nerve. The gene is located on chromosome 22	Occurrence is not clear	It is associated with intracranial and spinal meningiomas.
17	Osteogenesis imperfecta	Molecular defects of collagen. It is characterized by mutations in the *COL1A1* and *COL1A2* genes	1 per 20,000 live births	Respiratory failure or intracerebraz, hemorrhage resulting in reduced life expectancy, bones deformity
18	von Willebrand's disease	Bleeding disorder. It arises from a qualitative or quantitative deficiency of von Willebrand factor (vWF), a multimeric protein that is required for platelet adhesion	Occurrence not clear	Bleeding tendency, usually in the form of easy bruising, nosebleeds and bleeding gums. Women may experience heavy menstrual periods and blood loss during childbirth
III b. Autosomal Recessive				
Parents of an affected individual may not express the disease. On an average, the chances of an affected child's brothers or sisters having the disease are 1 in 4. For a child to have symptoms of an autosomal recessive disorder, the child must receive the defective gene from both the parents.				
19	Cystic fibrosis	This is caused by a mutation in the gene for the protein cystic fibrosis transmembrane conductance regulator (CFTR). The location of the CFTR gene on chromosome 7. Mutation delta 508 (F508) is known.	1 in 3700 births in United States	Symptoms often appear in infancy and childhood, such as bowel obstruction due to meconium ileus in newborn babies. Salty tasting skin, poor growth and poor weight gain despite a normal food intake, accumulation of thick, sticky mucus, frequent chest infections, and coughing or shortness of breath. Males can be infertile due to congenital absence of the vas deferens.
20	Glycogen storage diseases	Characterized by an inborn error of metabolism (genetically defective enzymes). It is the result of defects in the processing of glycogen synthesis or breakdown within muscles, liver, and other cell types	1 per 20,000-25,000 births in United States	Excess accumulation of glycogen in the liver and striated muscles.
21	Oculocutaneous albinism	It is a group of inherited disorders of melanin biosynthesis *OCA1* is caused by an alteration of the tyrosinase gene. Variants include *OCA1A* (the most severe form), *OCA1B*, OCA1-minimal pigment (OCA1-MP), OCA1-temperature sensitive (OCA1-TS), *OCA2, OCA3, OCA4* and *OCA5*.	1:50,000 births in United States	Hypopigmentation of skin, hair, eyes as result of inability to synthesize melanin

(Table 3). Continued.

S.No.	Genetic Disorder	Molecular and Cellular Defects	Prevalence	Features
22	Phenylketonuria (PKU)	It is a rare metabolic disorder caused by a deficiency of the liver enzyme phenylalanine hydroxylase.	1 in every 15,000 infants in the United States	Symptoms of untreated PKU develop gradually and would often go undetected until irreversible mental retardation had occurred; newborn infants are screened for abnormal levels of serum phenylalanine.
23	Sickle cell disease	Red blood cell defect caused by point mutation in the β-globin chain of haemoglobin. The β-globin gene is found on chromosome 11	1/625 of sub-Saharan African origin	Anemia including the vaso-occlusive crisis, aplastic crisis, sequestration crisis, haemolytic crisis and others
24	Tay-Sachs disease	It is lysosomal storage diseases, known as gangliosidoses. In this case, gangliosides are found in the neurons of the central nervous system and retina because of a failure of lysosomal degradation. It is caused by genetic mutation in the *HEXA* gene on chromosome 15	1/1000 east European Jews	Speech and swallowing difficulties, unsteadiness of gait, spasticity, cognitive decline, and psychiatric illness, particularly a schizophrenia-like psychosis
25	Alpha-1-antitrypsin (AAT) deficiency	It is a genetic disorder that causes defective production of alpha 1-antitrypsin (*A1AT*), leading to decreased *A1AT* activity in the blood and lungs, and deposition of excessive abnormal A1AT protein in liver cells.	1 in 1500 to 3500 individuals with European ancestry	Shortness of breath, wheezing, rhonchi, and rales
26	Familial Mediterranean fever	It is an auto inflammatory disease caused by 25 mutations in *MEFV* gene. The *MEFV* encodes pyrin (marenostrin), a protein implicated in the regulation of neutrophil activity. *MEFV* gene is located on the short arm of chromosome 16 (16p13).	Occurrence is not clear	Abdominal attacks, featuring abdominal pain, scrotal, joint attacks, chest attacks
27	Bardet Biedl syndrome	It is a ciliopathic human genetic disorder affects several body systems	In North America and Europe they are estimated to occur in 1 in 140,000 newborns	Obesity, retinitis pigmentosa, polydactyly, hypogonadism, and renal failure in some cases
28	Infantile osteopetrosis	It is a rare disease that results in a child having abnormal bones	1 in 20,000 births in United States	Large head (macrocephaly), failure to thrive, low platelets, low hemoglobin, large liver or spleen.
III c. X-Linked Recessive The incidence of the disease is much higher in males than females. Since the abnormal gene is carried on the X chromosome, males do not transmit it to their sons - they do transmit it to their daughters. The presence of one normal X chromosome masks the effects of the X chromosome with the abnormal gene. So, almost all of the daughters of an affected man appear normal, but they are all carriers of the abnormal gene. The sons of these daughters then have a 50% chance of receiving the defective gene.				
29	Hemophilia A	Bleeding disorder. It affects gene for factor VIII	1-2/10,000 males in United States	Bleeding joints, soft tissues and muscles, mouth from a cut, bitten tongue or loss of a tooth (especially in children), blood in the urine (hematuria) and surface bruising.
30	Duchenne dystrophy	Muscular dystrophy. It affects Dystrophin	1/3500 males in United States	Muscle weakness, enlargement of calf and deltoid muscles (pseudohypertrophy), low endurance, fibrosis, intellectual impairment abnormal bone development leading to skeletal deformities, including curvature of the spine eventually leading to paralysis.
31	Fragile X syndrome	Approximately 20% of males with fragile X mutation are clinically and cytogenetically normal. Because male carriers transmit the trait through all their daughters (who are phenotypically normal) to affected grandchildren, they are called *transmitting males.* Approximately 50% of female carriers are affected (mentally retarded), a proportion that is higher than with other X-linked disorders.	1 in 1000 male infants	Affected males are mentally retarded and share a common physical phenotype that includes a long face with large mandible and large, averted ears. Hyper extensible joints, a high-arched palate, and mitral valve prolapse. The distinctive feature, which is present in 90% of prepubertal males, is macro-orchidism or large testes.
III d. X-linked dominant inheritance The presence of the defective gene appears in females even if there is also a normal X chromosome present. Since males pass the Y chromosome to their sons, affected males will not have affected sons, but all of their daughters will be affected. Sons or daughters of affected females will have a 50% chance of getting the disease				
32	Hypophosphatemia rickets	The mutation results in altered (or missing) activity of the PHEX protein, which inactivates hormone-like substances (phosphatonins) that promote phosphate excretion. XLH is associated with a mutation in the *PHEX* gene sequence. It is located on the human X chromosome at location Xp22.2-p22.1.	The prevalence of the disease is 1:20,000	Short stature, dental problem, bone softening, weakness, knee pain, muscle pain

(Table 3). Continued.

S.No.	Genetic Disorder	Molecular and Cellular Defects	Prevalence	Features
		IV *Multifactorial Diseases*		
33	Cardiovascular disease	It refers to any disease that affects the cardiovascular system, principally cardiac disease, vascular diseases of the brain and kidney, and peripheral arterial disease. The causes of cardiovascular disease are diverse but atherosclerosis and/or hypertension are the most common.	Cardiovascular disease is the leading cause of deaths worldwide.	Currently, biomarkers which may reflect a higher risk of cardiovascular disease include: Coronary artery calcification. Carotid intima-media thickness.
34	Cancer	It is a disease of tissue growth regulation failure. In order for a normal cell to transform into a cancer cell, the genes which regulate cell growth and differentiation must be altered.http://en.wikipedia.org/wiki/Cancers - cite_note-pmid18234754-44 Several genes are involved, some of these are *BRCA1/2, MLH1, MSH2, P53.*	Occurrence not clear	Unintentional weight loss, fever, being excessively tired, and changes to the skin. Hodgkin disease, leukemias, and cancers of the liver or kidney can cause a persistent fever of unknown origin.http://en.wikipedia.org/wiki/Cancers - cite_note-Card10-7
35	Obesity	It is a medical condition in which excess body fat has accumulated to the extent that it may have an adverse effect on health, leading to reduced life expectancy and/or increased health problems.	As of 2008 the WHO estimates that at least 500 million adults (greater than 10%) are obese, with higher rates among women than men.	Excessive body weight is associated with various diseases, particularly cardiovascular diseases, diabetes mellitus type 2, obstructive sleep apnea, certain types of cancer, osteoarthritis and asthma. As a result, obesity has been found to reduce life expectancy.
36	Alzheimer's disease	It is the most common form of dementia, caused by mutations in one of three genes: amyloid precursor protein (APP) and presenilins 1 and 2. Most mutations in the *APP* and presenilin genes increase the production of a small protein called $A_{\beta}42$, which is the main component of senile plaques.http://en.wikipedia.org/wiki/Alzheimer's_disease - cite_note-pmid8938131-78	In 2006, there were 26.6 million sufferers worldwide.	Loss of memory, confusion, irritability, aggression, mood swings, trouble with language, and long-term memory loss.
37	Cleft lip (cheiloschisis)	It is a variation of a type of clefting congenital deformity caused by abnormal facial development during gestation. A number of genes are involved including cleft lip and palate transmembrane protein 1 and *GAD1*, one of the glutamate decarboxylases http://en.wikipedia.org/wiki/Cleft_palate -cite_note-Kanno2004-22.	1 in 700 children born have a cleft lip or a cleft palate or both	Clefts can also affect other parts of the face, such as the eyes, ears, nose, cheeks, and forehead.

rhythm troubles, short stature and gastrointestinal problems (Table **3**).

2.3. Single-Gene Disorders

These are caused by a single defective or mutant gene. The defective gene may be present on an autosome or the X chromosome, and one or both the copies affected. Single-gene defects follow the Mendelian pattern of inheritance, often called *Mendelian disorders.* Single gene disorders usually show a characteristic family history of a specific genetic disease and affect about 2 percent of the population. There are about 6000 known single gene disorders and these occur in about 1/200 births of which some examples are listed in Table **3**.

2.4. Multifactorial (also Called Complex or Polygenic) Disorders

These anomalies are the result of both genetic and environmental factors where multiple genes are affected despite the absence of clear cut genotype-phenotype correlation. The genetic testing of the multifactorial disorders can be performed within the framework of either case-control or prospective studies. Methodologies in detecting the most common polymorphisms are fully established and several of these have been correlated with diseases. Likewise, GWAS are frequently used to ascertain risk factors in spite of its indirect approach. Some common chronic disorders are heart disease, high blood pressure, Alzheimer's, arthritis, diabetes, obesity, age-related muscular degeneration (AMD) and several types of cancers (refer Table **3**) [5-7].

2.5. Human Y Chromosome and Genetics Disorders

Every chromosome is prone to mutations often resulting in genetic anomalies. However, all of them cannot be described in the context of their known mutations and associated diseases. Here we discuss the genomics of the human Y chromosome in the light of available information on related genetic anomalies. Human Y chromosome has more than 60 million DNA

sequences and represents around 2–3% of a total haploid genome [8]. The Y chromosome contains the largest non-recombining region in the human genome, spanning almost the entire length called MSY region [9]. A number of Y- linked genes have been described which are involved in control and regulation of in/fertility. These are *CDY1*, *DAZ1*, *DAZ2*, *DDX3Y*, *HSFY1* and *RBMY1A1*. The *DYZ1* and *DYZ2* represent satellite fractions useful for detecting aberrant Y chromosome and gender identification [10].

Y chromosome is heavily involved in male in/fertility, occurring approximately 1 in 2000 men. There are different categories where production of normal sperm is affected to varying degrees. Consequently, an azoospermic male produces no sperm. On the other hand, an oligospermic male produces reduced level of sperm compared to that of the normal one. The genes on the human Y involved in germ cell differentiation and spermatogenesis are clustered within the azoospermia encompassing *AZFa*, *AZFb* and *AZFc* regions of Yq11 [11,12]. In rare cases, changes to a single gene called *usp9y*, located in the *AZFa* region of the Y chromosome, can cause infertility. Deletion of *USP9Y* (X-degenerate) and *DBY* (ampliconic) gives rise to the Sertoli-cell-only type I syndrome. On the other hand, partial *AZFa* deletions with loss of *USP9Y* and presence of *DBY* exhibit mild oligozoospermia [13]. The Y chromosome has often been discussed in the context of sex chromosome related anomalies like XY gonadal dysgenesis, Swyer syndrome (SWY), XYY males, recurrent spontaneous abortions (RSA), Klinefelter's and Turner Syndromes. In a recent study, very high level of Y chromosome mosaicism has been demonstrated amongst the Turner patients [14]. However, in all such sex chromosome related genetic anomalies, all the known Y-linked genes have not been systematically analyzed. A detailed study on this line would resolve most commonly affected genes under a given genetical anomaly.

Involvement of the Y chromosome in human oncogenesis is yet another concern that has attracted a great deal of attention [15]. Gain, loss and rearrangements of Y have been associated with bladder cancer [16], male sex cord stroma tumors [17], lung cancer [18] and esophageal carcinoma [19]. Further, its instability was detected in 55.6% of 27 cases of Finnish non-Hodgkin lymphomas [20]. Significantly, complete loss of the Y in all the cells is frequently observed in cases of chronic myelogenous leukemia (CML), myeloproliferative diseases (MPD) and myelodysplastic syndrome (MDS) commonly found in the patients over 60 years of age [21]. The gonadoblastoma gene (*GBY*) is putatively located in the proximal region on the long arm of the Y chromosome distinct from the testis-determining factor [22]. Sequence-tagged sites were used to perform deletion mapping in sex-reversed female with a Y chromosome and gonadoblastoma. It was discovered that *TSPY* gene (testis-specific protein, Y-encoded) is up regulated in gonadoblastoma tissues in addition to other testicular and prostate cancers [23]. Despite genetical advancement, thus far, the exact location of the gene for gonadoblastoma is not known. Seemingly, only indirect evidence about the existence of this gene on the Y chromosome is available [24]. A total of 13 chromosomes including X and Y have been reported to be involved in prostate cancer [25]. The role of Y chromosome in the maintenance of prostate gland is not yet defined. However, its analysis in the patients is of relevance as the same is male specific cancer particularly in the context of development of DNA based marker. Identification of DNA based marker for prostate cancer (PC) would require massive hunt of the genes and loci across the different grades of samples to uncover a consensus sequence. Chromosome transfer studies indicate that the Y chromosome suppresses tumorigenicity of human prostate cell lines *in-vivo* implying that it may have tumor suppressor gene(s) [26]. Y chromosome has been reported to be unstable in prostate cancer cell line and perhaps in tissues as well [27]. Initial studies showed *SRY* gene to be a negative regulator of the androgen receptor but the same is down-regulated in PCs [28].

A perusal of literature indicates that different ethnic groups have varying level of predisposition towards PC [29]. Paracchini *et al*., (2003) showed the significant statistical predisposition of PC only in one out of the 4 lineages of Japanese groups [30]. On the contrary, in the Korean population, no association was found between Y haplogroups and PC [31]. In a recent study, *DYZ1* repeat arrays were found to be drastically reduced from average 4000 copies per haploid genome to 550 copies. This indicates that besides several functional Y-linked and autosomal genes, *DYZ1* repeat arrays are equally prone to alteration in cases of PC [32]. Whether this would also have an ethnic bearing is not clear at this stage. However, work along this line would surely generate much needed data related to genetical background and cancer. Likewise, testicular cancer is the common one in white males aged between 20–40 years [33]. The worldwide incidence is 7.5 per 100,000, but this ratio varies across the

countries. Individuals with the intersex syndrome and a relative reduction of the Y chromosome genetic material carry a high risk of germ cell neoplasias [34]. Perhaps presence of intact Y in all the cells are needed to maintain the critical balance of gene products of which some may act as tumor suppressor.

3. EMERGENCE OF NEWER TECHNOLOGIES AND GENOME ANALYSIS

Given the complexity of cellular systems, techniques have been developed over the years that allow the comprehensive analysis of the genome. The human genome culminated in year 2003 but much of the enthusiasm died down since major challenges related to human health care systems still exist. This surely does not mean that sequencing of the human genome remained unrewarded. Research related to many fundamental biological phenomena has accelerated the advances in genomic technologies. Further, high throughput sequencing technologies have changed the much desired landscape of DNA based diagnosis in human. Many techniques have been introduced during the past half a century. These have now evolved and become much more powerful. Thus, with conventional approaches, new found tools have synergized the research. For genetic testing of disorders scenarios, an appropriate and technically feasible approach needs to be defined. We provide a brief overview on the conventional and associated technologies useful for human genome analyses though not necessarily in order of priority.

3.1. Relevance of Cytogenetical Approaches

Identification of chromosomal aberrations is important for risks assessment [35,36]. In order to explore the integrity of individual chromosomes within the nucleus of intact cells, fluorescence *in situ* hybridization (FISH) is used to map DNA sequences to specific regions of human chromosomes. FISH is used to assess copy number variation in association with real time PCR particularly in the context of prenatal diagnosis and tumor characterization. It is particularly useful for detecting submicroscopic chromosomal deletions associated with specific malformation syndromes. Interphase FISH is highly sensitive in detecting the *BCR/ABL* fusion, and therefore is very useful for following patient's response to therapy [37,38]. Many specific chromosomal and gene rearrangements have been characterized in solid tumors. These rearrangements, for example, the translocation t(X;18)(p11.2;q11.2) in synovial sarcoma and the *EWS/FLI1* fusion in Ewing sarcoma/ peripheral primitive neuroectodermal tumor, can be detected by dual color interphase FISH in formalin-fixed, paraffin-embedded tumor tissues [39]. Microsomal syndromes and disorders are characterized by small deletions in specific chromosomal segments and can be reliably detected by FISH in Prader-Willi, Angelman, Williams, Miller-Dieker, Smith-Magenis and velocardiofacial syndromes [40,41]. With the improvement of the quality of chromosome preparation, the deletion of 17p11.2 in Smith-Magenis syndrome can be detected by G-banding analysis [42].

Spectral Karyotyping (SKY) and Multiplex–FISH (M-FISH) enables the simultaneous tracking of all human chromosomes. SKY analysis has been used to detect inter chromosomal rearrangements and aneuploidy [43,44]. Comparative genomic hybridization (CGH) or reverse *in situ* hybridization is a FISH method for genome-wide screening to uncover the differences in copy number of any DNA sequence in an individual. The size of the DNA segments that CGH can detect is estimated to be in range of 10-20Mb. CGH is useful in the characterization of *de novo* unbalanced constitutional anomalies [45]. Since the entire genome can be scanned for gains or losses without preparing metaphase chromosomes of the cells or tissues tested, CGH has been widely used in investigations of solid tumors [46]. Further, Primed *in situ* labeling (PRINS) refers to a process of reannealing short oligonucleotide primers to target sequences *in situ*, followed by elongation of the sequences with a *Taq* polymerase and simultaneous labeling of the target sequences with a fluorochrome. This technique has been used as an efficient alternative tool to detect aneuploidies [47,48]. Thus, cytogenetical parameters are still the reliable tools to address a large number of key questions related to normal/abnormal chromosomes.

3.2. Genomics and Recombinant DNA Technology

Most commonly used recombinant DNA techniques for genome analysis are Southern, Northern, Western Blotting, Restriction Fragment Length Polymorphism (RFLP), Reverse Transcription Polymerase Chain Reaction (RT-PCR), cloning, sequencing and Real Time PCR. Conventionally, DNA synthesis was done using tritiated (H^3) thymidine autoradiography. Likewise, RFLP akin to DNA typing facilitated the identification of altered alleles. The restriction enzymes used for RFLP are usually based on the empirical optimization unless the gene sequences are known. Information on gene sequence may only be obtained if

the gene has been cloned and sequenced. Earlier, gene isolation used to be a daunting task since it involved construction and screening of the genomic or cDNA libraries. With the availability of PCR, gene isolation has now become feasible. In many instances, instead of employing RFPL approach, direct PCR followed by sequencing is used to uncover alleles or allelic variation which is more accurate.

In the process of studying genetic variation, arbitrarily primed polymerase chain reaction (AP-PCR) also known as rapid amplification of polymorphic DNA (RAPD) was developed [49,50]. AP-PCR/RAPD was found to have limited applications and even more limited implications. With the passage of time, analogous to multilocus polymorphic band profiles, minisatellite associated sequence amplification (MASA) was developed. MASA reaction is conducted using cDNA template and primers based on VNTR loci. However, such reactions would uncover only genes that are tagged with VNTR loci [For technical details, see 51-57]. Thus, to uncover the entire pool of genetic complexity, awareness of associated technology is warranted.

Many methods to detect genetic diseases exist, ranging from microscopic examination of intact chromosomes to analyses of the expression of gene/ gene products at the mRNA and protein level. A brief overview focusing on the upcoming high throughput techniques for analysis of complex genetic disorders is provided hereunder.

3.3. Impact of Sequencing Technologies

Continuous technological improvements in DNA sequencing has created an ambiance par excellence that a large number of disease causing microbe and viral genomes are sequenced on regular basis. This in turn has fueled the research related to metagenomics on a large scale, development of drug molecules against defined targets including functional and comparative genomics. Arguably, the strongest rationale for ongoing sequencing is the quest for identification and interpretation of DNA sequence not only from the human genome but also across the other species.

3.3.1. First Generation Sequencing

The automated Sanger method is considered to be the 'first–generation' technology. The "original" sequencing based on the Sanger's chemistry uses specifically labeled nucleotides to polymerize through the templates [58]. In the chain termination reaction, the replicated section of DNA is synthesized in a series of small fragments rather than in one strand. The nucleotide that ends each fragment is tagged with a radioactive or fluorescent marker. After a series of technical innovations, the Sanger method has now evolved to read 1000–1200 base pair (bp) in one cycle [59,60]. Now, several newer machines and sequencing systems have emerged. These newer systems operate on varying principle employing very different chemistry, algorithm and softwares to handle large chunk of data ensuring less or negligible errors.

3.3.2. Next Generation Sequencing (NGS)

One concept that has gained momentum is to undertake sequencing employing approaches where reads are more, errors are less and data output is enormous. To achieve this, massive parallel sequencing, next-generation platforms have been developed. This uses clonal amplification of the DNA templates on a solid support matrix followed by cyclic sequencing. Next-generation sequencing (NGS) has synergized the genomic-scale biological research, and its effects have started percolating down through the ladder in the form of translation research, system and synthetic biology. In this sense, the comprehensive sequencing of the genome, epigenome and transcriptomes of cancers and corresponding "normal" (germ-line) DNA are heralding the start of personalized medical genomics. The first commercial NGS platform was based on pyrosequencing techniques; it was soon surpassed in output by reversible dye termination and sequencing by ligation reaction [61]. In a recent study, two Y chromosomes separated by 13 generations were sequenced using NGS technology [62,63]. Approximately one mutation per generation was identified, suggesting that every individual Y chromosome can be distinguished by sequencing. This approach is equally useful in the context of forensic science [64] to resolve between the closely related men. With the availability of NGS, entire cancer genome has been sequenced. In a systematic study employing, Sanger, NSG and pyrosequencing approaches, *EGFR* and *KRAS* mutations were assessed in lung cancer specimens [65]. NGS provided sensitivity superior to the other two methods. The number of genetic mutations associated with cancer seems to be growing. In this context, epidemiological studies have elucidated molecular basis of cancer in a number of cases including *BRCA1* and *BRCA2* genes [66]. Similarly, mutations in *MLH1* and *MSH2* genes have been shown to be associated with a higher risk of

colon cancer [67]. To uncover structural variations in the genome, a very useful program BreakDancer [68] is available. NGS has facilitated studies on genetic variations in lung cancer, melanomas and breast cancer showing single nucleotide polymorphism [69-72].

Besides NGS, several newer platforms such as Roche GS-FLX 454 Genome Sequencer (originally 454 sequencing), the Illumina Genome Analyzer (originally Solexa technology), the ABI SOLiD analyzer, Polonator G.007, and the Helicos BioSciences HeliScope are available (For details, refer to Table **4**). These platforms have provided unprecedented opportunities for high-throughput functional genomic research. The Roche GS-FLX 454 Genome Sequencer was the first commercial platform introduced in 2004 as the 454 Sequencer. The second complete genome of an individual (James D. Watson) was sequenced with this platform [73]. Importantly, small RNA sequencing studies with the 454 technology contributed to the discovery of a novel class of small RNAs, termed Piwi-interacting RNAs that are expressed in mammalian testes [74,75]. Applications of NGS extends beyond DNA sequencing because the core genome biotechnology also offers the opportunity to sequence and analyze the whole transcriptome (RNA-Seq), epigenetic modifications (Methyl-Seq) and transcription factor binding sites as well as histone modifications (Chip-Seq) [76,77]. Further, ChIP–Seq data was used to map the positions of two types of nucleosomes, as well as RNA Pol II transcription preinitiation complexes in human CD4+ T cells [78].

Another dimension of NGS technology is the modest cost with which hundreds of gigabases are sequenced. NGS-based systems have enhanced cancer biology research at an unprecedented scale resulting in rapid sequencing of tumor genomes. Similarly, sequencing of the pathogens has been done to determine drug resistance and to identify chromosomal abnormalities during pregnancy.

3.3.3. Third Generation Sequencing (TGS)

Despite NGS, a new generation of single-molecule sequencing (SMS) technologies is emerging with promise of much longer read with reduced cost in shorter time [79,80]. TGS is still passing through the phases of technically demanding innovations before the same is made available to be used for large scale analysis of clinical samples. It would be interesting to examine its performance compared to that of other existing systems. TGS is used to sequence DNA, but the DNA polymerase can be replaced with a reverse transcriptase enzyme to directly sequence RNA [81]. Recently Pacific Biosciences have introduced a TGS platform, single-molecule real-time (SMRT) (see Table **4**). A recently published application of the SMRT technology demonstrated direct, real-time observation of the ribosome as it translated mRNA [82]. TGS therefore, stands ready to provide unprecedented snapshots of complex and demanding systems including accurate network view of the diseased genome.

Other technologies available between NGS and TGS are Ion Torrent's semiconductor sequencer and Helicos Genetic Analyses Platform, the first commercially available sequencing instruments to carry out single molecule sequencing (SMS) [83,84] (Table **4**). Ion Torrent has recently released 400 base sequencing on the Ion PGM™ System for improved assembly *de novo* of microbial sequencing, making it the only benchtop sequencer to offer long read sequencing as a cost effective option for routine use. The launch of the Ion AmpliSeq™ technology for targeted RNA sequencing includes panels for cancer and apoptosis, along with the ability to customize human RNA panels.

Thus, it would be even more difficult to handle, annotate and interpret such a large quantum of data generated by these robust and ever growing sophisticated systems. It would surely warrants correspondingly equally powerful bioinformatic tools and large scale automated data processing systems to make the judicious use of these minds, materials and machine.

3.3.4. Exome Sequencing

The development of methods for coupling targeted capture and massively parallel DNA sequencing has made it possible to economically determine nearly all the coding sequence variation present in a human genome. This is called 'exome sequencing' where all the exons are sequenced. This smart approach is used as a reliable tool for dissecting the genetic basis of diseases that have been intractable to conventional diagnosis. Steps involved in exome sequencing include library preparation, target capture, target enrichment and sequencing. Over the past 2 years, exome sequencing have paved the path for identifying genes that underlie the cause of known or suspected Mendelian disorders for which conventional

Table 4: Characteristics of Next and Third Generation Sequencing Platforms

S.No.	Company/Platform	Amplification approach	Chemistry	Average Read Length (bp)	Run time	Pros/Cons
1	Roche GS-FLX 454 Genome Titanium	In this, single-stranded DNA binding beads are encapsulated by vigorous vortexing into aqueous micelles containing PCR reactants surrounded by oil for emulsion PCR amplification. During the pyrosequencing process, light emitted from phosphate molecules during nucleotide incorporation is recorded as the polymerase synthesizes the DNA strand.	Pyrosequencing	330bp	10-23 hrs	Pro: Longer reads. Short time Con: High reagent cost, Homopolymer errors
2	Illumina/Solexa's Genome Analyzer	In this, all four nucleotides are added simultaneously into oligo-primed cluster fragments in flow-cell channels along with DNA polymerase. Bridge amplification extends cluster strands with all four fluorescently labeled nucleotides for sequencing.	Sequencing by synthesis Reverse dye terminator chemistry (SBS RDT)	75-100bp	7-40 hrs	Pro: One of the most widely used platforms Con: Low multiplexing of samples
3	ABI/ SOLiD (Solid oligonucleotide ligation detection) 5500xL	It uses an emulsion PCR approach with small magnetic beads to amplify the DNA fragments for parallel sequencing	Sequencing by ligation (SBL)	50-100bp	2-7 days	Pro: Ultra high output, scalable runs allow sequencing on part flow cell Con: Shorter reads than other platforms, long time for clonal template preparation
4	Danaher/Dover/AzcoPolonator G.007	It is a new platform in the market with emphasis on competitive pricing. The Polonator platform employs a sequencing-by-ligation approach using a randomly arrayed, bead-based, emulsion PCR to amplify DNA fragments for parallel sequencing.	Non cleavable probe (SBL)	26bp	8-10Gbp /Run	Pro: One of the least expensive platforms Con: Users are required to maintain and quality control reagents, Short read lengths
5	Illumina Mi-seq	The MiSeq instrument, in fact, requires no user intervention from cluster generation to data analysis. Cluster generation is typically quite robust provided the sequencing libraries are of high quality and the concentration of the library is accurately measured by quantitative PCR.	SBS RDT	36-150bp	4-27 hrs	Pro: Proven chemistry. Fully automated workflow Con: expensive per base
6	Helicos BioSciences/ RNA-Seq	It sequences RNA templates directly without the need to convert them into cDNAs.	High throughput sequencing	150bp	3-4 days	Pro: Fast and generate high read lengths enables the discovery of novel splice forms, transcripts and RNA-editing
7	Illumina/ChIP-Seq	This technique couples the commonly used chromatin immunoprecipitation procedure, in which DNA–protein complexes are cross-linked and precipitated using an antibody, to next-generation sequencing of DNA fragments bound to the precipitated protein.	SBS RDT	50bp	-	Pro: Low sample input, enables comprehensive binding of *in vivo* binding sites across an entire genome Con: Short read lengths

(Table 4). Continued.

S.No.	Company/Platform	Amplification approach	Chemistry	Average Read Length (bp)	Run time	Pros/Cons
8	PacBio® RS High Resolution Genetic Analyzer	It incorporates novel, single molecule sequencing techniques (SMRT) and advanced real time analytics.	SMRT	200-250bp	-	Pro: Long read length, high accuracy, ability to sequence large repeat regions and complex genomes Con: N/A
9	Life technologies/ Ion AmpliSeq	It detects signal by the release of hydrogen ions resulting from the activity of DNA polymerase during nucleotide incorporation. In essence, the Ion Torrent chip is a very sensitive pH meter.	Sequencing by synthesis Hydrogen ion detection (SBS H^+)	35-75bp	2hrs	Pro: Label free chemistry-cheap and fast, high scalable, potential Con: Homoplymer errors, short reads, laborious template preparation but semi-automatable
10	Helicos BioSciences/ HeliScope	The technology has the ability to sequence single DNA molecules without amplification, defined as Single-Molecule Real Time (SMRT) DNA sequencing. It uses highly sensitive fluorescence detection system to directly interrogate single DNA molecules *via* sequencing by synthesis. Template libraries, prepared by random fragmentation and poly-A tailing (that is, no PCR amplification), are captured by hybridization to surface-tethered poly-T oligomers to yield a disordered array of primed single-molecule sequencing templates.	Single molecular sequencing (SMS)	32bp	8 days	Pro: No bias representation of template for genome and Seq- based applications Con: High error rates, short reads

approaches have failed. The hypothesis behind exome-sequencing related to complex diseases propelled by early sequencing results [85,86], is that multiple rare variants in protein-coding genes responsible for disease of interest may be accurately identified.

Exome sequencing is envisaged to facilitate improved diagnostics, prevention strategies and targeted therapeutics. This has already been found to be useful for disorders such as Freeman-Sheldon syndrome, congenital chloride-loosing enteropathy, Kabuki syndromes, Clericuzio-type poilioloderma with neutropenia, familial exudative vitreoretinopathy to name a few [87-89]. Several families with dominantly inherited adult-onset arterial calcifications were found to show mutations in *NT5E* gene that encodes a protein involved in adenosine metabolism. Specific therapeutic interventions became possible owing to information obtained from the exome sequencing that would have not been feasible otherwise [90].

The availability of commercial capture reagents from both NimbleGen and Agilent that target human exons have greatly accelerated the utilization of exome sequencing strategy [91]. The solution-based exome capture kits are easily adaptable to a high-throughput workflow which does not require any further material and manpower. Exome sequencing on *de novo* variants in children with idiopathic intellectual disabilities and sporadic autism suggest that such phe-notypes are tractable. Exome sequencing has been successfully used to discover a novel Cys203Tyr variant in X-linked inhibitor of apoptosis (*XIAP*) in a young boy suffering from severe inflammatory bowel disease where reliable diagnosis was elusive [92].

The growing number of exome sequencing demonstrates the power of this approach in mapping genes involved in Mendelian phenotypes. Despite our high expectations from this approach, the success is not always guaranteed. Non-allelic heterogeneity,

regulatory and structural variations underlying phenotypes all pose challenges for sequencing-based discovery of Mendelian genes. New statistical and computational methods are envisaged to enhance the success rate of exome sequencing in the context of Mendelian disorders.

3.4. Microarray and Gene Expression

DNA microarray is another latest breakthrough in the experimental molecular biology. Arrays have added an additional dimension to ever growing pool of knowledge. Cancer research coupled with diagnostic DNA microarrays is playing a dominant role to address issues related to identification of gene(s) in such anomalies. Based on gene expression profiles of Acute Lymphoblastic Leukemia (ALL), a novel subtype was identified [93].

Microarrays' ability to identify key markers for prognosis and treatment response by profiling thousands of genes expressed in a single cancer is significant in a specific cancer types. These include breast cancer in women and prostate cancer in men [94,95]. Coupled with biochemical analysis such as immunohistochemistry (IHC) and enzyme linked immunesorbent assay (ELISA), microarrays may be used for diagnostic and prognostic purposes in the context of translational research [96]. Employing cDNA microarray approach, up-regulation of osteopontin gene encoding calcium binding glycophosphoprotein was identified in ovarian cancer [97].

Microarrays have also been used for studying single nucleotide polymorphisms (SNPs) which in turn is useful to identify disease markers, loss of heterozygosity (LOH), tumor suppressor genes, and drug responses to the patients. LOH were identified in bladder cancer, prostate cancer and small-cell lung carcinomas samples [98,99] with HuSNP arrays bearing probes representing 1500 SNPs. Bignell *et al.* (2004) used an Affymetrix SNP research array bearing oligonucleotides representing approximately 8500 SNPs to identify genotype information as well as changes in DNA copy number in 20 different cancer cell lines (see Table **5**) [100].

Microarray technology may improve the identification of both genetic and molecular causes of susceptibility to certain diseases [101,102]. This has

Table 5: Comparison of Commercially Available Microarray Platforms

S.No.	Applicable gene chip products/ services	Company's name	Main features
1	Infiniti RVP (Solid chip)	AutoGenomics, Inc.	The detection step by the analyzer is completely automatic
2	ResPlex II assay (Liquid chip)	Qiagen	A unique target enriched multiplex PCR (TEM-PCR) allows large numbers of targets included in one reaction without significant loss of sensitivity
3	U133 Plus 2.0 GeneChip	Affymetrix	Genome- level alterations of zinc homeostasis may be prevalent in clinical pediatric septic shock
4	Hu 133A and 133b GeneChip	Affymetrix	Human blood leukocytes response to acute systemic inflammation includes transient dysregulation of leukocyte bioenergetics and modulation of translational machinery.
5	Atlas array	Clonetech Laboratories	Microarray technology provides a powerful new tool for rapidly analyzing tissue specific changes in gene expression induced by sepsis
6	GeneChip Human Tiling 2.0R Array Set	Affymetrix	Most comprehensive whole genome array set for studying protein/DNA interactions in chromatin immunoprecipitation (Chip) experiments
7	GeneChip CustomSeq Resequencing Arrays	Affymetrix	Flexible custom arrays containing up to 300Kb of unique, high quality, double stranded sequence for less than a penny per base
8	Image consortium libraries	Livermore National Laboratory	Both gram positive and gram negative sepsis share a final common pathway involved in pathogenesis of sepsis, but certain gene are differentially expressed under distinct regulation
9	Genome Wide Human SNP Array 6.0	Affymetrix	Highest coverage for combined copy number and Loss of heterozygosity detection on a single array

indeed facilitated finding of drug targets and augmented molecular diagnostics [103,104]. Microarrays are entirely dependent on the state of knowledge of the genome under investigation and do not measure posttranslational modifications (e.g., phosphorylation) [105].

3.5. Implications of RNAi Technology

Following human genome studies, a shift has occurred from mRNAs to noncoding RNAs as a main regulator of the human genes. Putatively, mammalian miRNAs have originated from transposons and repeats [106]. The discovery of RNAi has led to the realization that RNAi machinery is also involved in normal gene expression employing a class of small RNAs known as microRNAs.

Micro ribonucleic acids (miRNAs) are a large class of endogenously expressed single stranded, evolutionarily conserved small non-coding RNAs, typically 17-25 ribonucleotides in length that are found in plants, animals and DNA viruses. miRNAs regulate cell division, differentiation, cell fate decisions, development, oncogenesis, apoptosis, gene expression and are involved in a number of genetic diseases [107-109]. Oligonucleotide miRNA microarray analysis has been used extensively as high-throughput method for the evaluation of global expression in a large number of samples [110]. Significantly, miRNA levels are dramatically shifted in various cancers, and this also acts as oncogenes [111-112]. The expression profile of miRNAs is tightly regulated for a particular type of tissue and stage of cell differentiation [113]. Impaired miRNAs functioning which occurs during tumor transformation can be evaluated as a consequence rather than the cause of loss of cell identity. In designing a particular RNAi, it is important to identify the sense/antisense combination that holds the key for suppression of the target mRNA. In this context, rules have been established to ensure >90% gene expression inhibition. The high sequence conservation of many miRNAs among distantly related organisms suggests strong evolutionary pressure and involvement in main physiological processes. In fact, a subset of miRNAs are correlated with a range of clinically imperative diseases including myocardial infarction, virus infection, Alzheimer's disease, metabolic diseases and several types of cancers (Table **6**) [114-116].

Table 6: Micro RNAs and their Association with Genetic Diseases

S.No.	miRNAs	Diseases/Phenotypes affected	Gene Targets
1	miR-21	Cancer: Breast, Colon, Lung, Pancreas, Brain, Liver and cardiac hypertrophy	*CDK6,PDCD4,FAS,IL6R TPM1, CDKN1A,SOCS5*
2	miR-15B	Cell proliferation, cell cycle	*CCNE1*
3	miR-26a	*In vivo* transformation of NIH3T3 cells	*PTEN,Rb1, MAP3K2, MEKK2*
4	miR-34a	Cell proliferation, cell cycle, apoptosis, invasion	*c-Met, Notch 1, Notch 2*
5	miR-9	Spinal motor neuron disease	
6	miR-25/92	Cancer: Leukemia, lung, stomach, colon, prostate and thyroid	*CDKN1C, BCL2L11*
7	miR-19, miR-101, miR-100	Spinocerebellar ataxia type 1	
8	miR-142	Aggressive B cell leukemia	*Translocated c-MYC gene*
9	miR-196	Crohn's disease	
10	miR-155, miR-186	Chronic lymphocytic leukemia	*BIC RNA, IgVh gene*
11	miR-675	Silver Russel Syndrome	
12	miR-296	Endothelial tubule formation, migration, *in vivo*, tumor neovascularization	*HGS*
13	miR-221	Cancer: Thyroid, Stomach, pancreas, prostate, melanoma	*CDK1B, CDKN1C, KIT*
14	miR-7, miR-184	Parkinson's disease	
15	miR-155, miR-802	Down's syndrome	
16	miR-372/3	Breast, testicular germ cells	*LATS2,CD44,CD24*
17	miR-146a	Rheumatoid arthritis	
18	miR-17/20/93/106	Lung, colon, stomach, pancreas, prostate, leukemia, thyroid	*E2F1, CDKN1A, RUNX1, NCOA3*

RNAi may or may not have the potential to ameliorate all the diseases but is expected to provide much needed additional dimension towards the same. Pharmacological manipulation of miRNA is still in its infancy. However, correlation between the expression of miRNAs and their effects on target oncogenes, tumorigenesis and proliferation of cancer cells has been experimentally established. Thus, discovery and characterization of RNAi is not only a powerful molecular biological tool to suppress the expression of a target gene but also an emerging therapeutic strategy to silence diseased genes. This is particularly tempting in the context of cancer research.

3.6. Impact of Proteomics on Genomics

Advances have been made in the arena of proteomics based on identification of a large number of peptides [117]. Proteomics has now become an integral part of the genome analysis and continue to play important roles in dissecting protein functions. In general, proteomic approaches can be used for proteome profiling, comparative expression analysis of two or more protein samples, localization, identification of posttranslational modifications, and study of protein–protein interactions. A wide range of proteomic approaches are available including gel-based, one - and two dimensional electrophoresis [118,119]. In addition, gel-free high throughput screening which requires isotope-coded affinity tag ICAT [120], stable isotope labeling with amino acids in cell culture (SILAC) [121] and isobaric tagging for relative and absolute quantitation (iTRAQ) are available [122]. Shotgun proteomics [123] and two-dimensional fluorescence difference gel electrophoresis (2DE DIGE) [124] as well as protein microarrays [125] are applied to obtain overviews of protein expression in tissues, cells, and organelles. Large-scale immunological assays [126], multiple reaction monitoring assay (MRM) [127], and label-free quantification of high mass resolution Liquid chromatography–mass spectrometry (LC-MS) [128] are being explored for high throughput analysis. Proteomics based studies have been used to investigate human diseases using animal models to gain insight into disease mechanisms. These include studies in animal models of cancer, Alzheimer's disease, stroke, dilated cardiomyopathy and others [129,130]. With the enormous data available, it is essential to develop an equally effective computational framework for the integration of proteomics data with phenomic and functional genomic ones for more accurate genotype-phenotype correlation.

4. CONCLUDING REMARKS

Genome analysis has progressed from the completion of the human genome to functional genomics and finally to translation research. Further, biological system and synthetic biology both have synergized the process. These developments indeed have augmented our understanding about the genetic diseases. However, finding therapeutic measures is still a major challenge. In the absence of epidemiological data, setting up the research priority remains insipid. Thus, large scale generation of epidemiological data particularly in the developing countries would go a long way to prioritize the focus of research. Similarly, bringing clinicians and researchers on one platform would be yet another step in the right direction. We believe that the quest to fight with the diseases would continue so also the emergence of the newer and far more despicable diseases. However, in view of the overall development during the last five decades or so, it is envisaged that our innate optimism would enable us to conquer most of the diseases ameliorating human sufferings.

ACKNOWLEDGEMENT

We thank Department of Biotechnology (DBT) and Department of Science and Technology (DST), New Delhi, for the award of research grants BT/PR11805/ MED/12/424/2009, BT/PR14102/AAQ/01/438/2010 and SR/SO/AS-115/2012, respectively to SA and core grant to the National Institute of Immunology, New Delhi. SA acknowledges award of the J.C. Bose National Fellowship by DST, New Delhi.

ABBREVIATIONS

CGH = Comparative genomic hybridization

FISH = Fluorescence *in situ* hybridization

GWAS = Genome wide association studies

HGP = Human Genome Project

NGS = Next generations sequencing

PC = Prostate Cancer

SNP = Single nucleotide polymorphism

SWY = Swyer syndrome

TGS = Third generation sequencing

REFERENCES

[1] Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 1953; 171(4356): 737-738. http://dx.doi.org/10.1038/171737a0

[2] Online Mendelian Inheritance in Man (OMIN™). (2003). Baltimore, MD: McKusick-Nathans Institute of Genetic Medicine (http://www.ncbi.nlm.nih.gov/omim).

[3] Wright AF, Hastie ND. Complex genetic diseases: controversy over the Croesus code. Genome Biol 2001; 2(8): COMMENT 2007.

[4] Nussbaum RL, McInnes RR, Willard HF. Thompson and Thompson genetics in medicine. 6th ed. Philadelphia, W. B. Saunders; 2001 pp. 51-388.

[5] Strittmatter WJ, Roses AD. Apolipoprotein E and Alzheimer's disease. Annu Rev Neurosci 1996; 19: 53-77. http://dx.doi.org/10.1146/annurev.ne.19.030196.000413

[6] Peyser PA. Genetic epidemiology of coronary artery disease. Epidemiol Rev 1997; 19(1): 80-90. http://dx.doi.org/10.1093/oxfordjournals.epirev.a017949

[7] Mein CA, Esposito L, Dunn MG. A search for type 1 diabetes susceptibility genes in families from the United Kingdom. Nat Genet 1998; 19(3): 297-300. http://dx.doi.org/10.1038/991

[8] Harris P, Boyd E, Young BD, Ferguson-Smith, MA. Determination of the DNA content of human chromosomes by flow cytometry. Cytogenet Cell Genet 1968; 41(1): 14-21. http://dx.doi.org/10.1159/000132190

[9] Rozen S, Skaletsky H, Marszalek JD, Minx PJ, Cordum HS, Waterston RH *et al*. Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. Nature 2003; 423(6942): 873- 876. http://dx.doi.org/10.1038/nature01723

[10] Nakagome Y, Seki S, Fukutani K, Nagafuchi S, Nakahori Y, Tamura T. PCR detection of distal Yp sequences in an XX true hermaphrodite. Am J Med Genet 1991; 41(1): 112-114. http://dx.doi.org/10.1002/ajmg.1320410127

[11] Vogt PH. Human chromosome deletions in Yq11, AZF candidate genes and male infertility: history and update. Mol Hum Reprod 1998; 4(8): 739-744. http://dx.doi.org/10.1093/molehr/4.8.739

[12] Kuroda-Kawaguchi T, Skaletsky H, Brown LG, Minx PJ, Cordum HS, Waterston RH, *et al*. The AZFc region of the Y chromosome features massive palindromes and uniform recurrent deletions in infertile men. Nat Genet 2001; 29(3): 279-86. http://dx.doi.org/10.1038/ng757

[13] Sargent CA, Boucher CA, Kirsch S, Brown G, Weiss B, Trundley A, *et al*. The critical region of overlap defining the AZFa male infertility interval of proximal Yq contains three transcribed sequences. J Med Genet 1999; 36(9): 670-677.

[14] Premi S, Srivastava J, Panneer G, Ali S. Startling mosaicism of the Y-Chromosome and tandem duplication of the *SRY* and *DAZ* genes in patients with Turner syndrome. PLoS ONE 2008; 3(11): e3796. http://dx.doi.org/10.1371/journal.pone.0003796

[15] Quintana-Murci L, Fellous M. The human Y chromosome: the biological role of a functional wasteland. J Biomed Biotechnol 2001; 1(1): 18- 24. http://dx.doi.org/10.1155/S1110724301000080

[16] Sauter G, Moch H, Wagner U. Y chromosome loss detected by FISH in bladder cancer. Cancer Genet Cytogenet 1995; 82(2): 163-169. http://dx.doi.org/10.1016/0165-4608(95)00030-S

[17] de Graaff WE, van Echten J, van der Veen AY. Loss of the Y-chromosome in the primary metastasis of a male sex cord stromal tumor: pathogenetic implications. Cancer Genet Cytogenet 1999; 112(1): 21-25. http://dx.doi.org/10.1016/S0165-4608(98)00245-3

[18] Center R, Lukeis R, Vrazas V. Y chromosome loss and rearrangement in non-small lung cancer. Int J Cancer 1993; 55(3): 390-393. http://dx.doi.org/10.1002/ijc.2910550309

[19] Hunter S, Gramlich T, Abbott K. Y chromosome loss in esophageal carcinoma: an in situ hybridization study. Genes Chromosomes Cancer 1993; 8(3): 172-177. http://dx.doi.org/10.1002/gcc.2870080306

[20] Richard SM, Knuutila S, Peltomäki P, Bianchi MS, Bianchi NO. Y chromosome instability in lymphoproliferative disorders. Mutat Res 2003; 525(1-2): 103-107. http://dx.doi.org/10.1016/S0027-5107(03)00007-1

[21] Kirk JA, VanDevanter DR, Biberman J, Bryant EM. Y chromosome loss in chronic myeloid leukemia detected in both normal and malignant cells by interphase fluorescence in situ hybridization. Genes Chromosomes Cancer 1994; 11(3): 141-145. http://dx.doi.org/10.1002/gcc.2870110302

[22] Page DC. Hypothesis: A Y-chromosomal gene causes gonadoblastoma in dysgenetic gonads. Development 1987; 101(Suppl): 151-155.

[23] Tsuchiya K, Reijo R, Page DC, Disteche CM. Gonadoblastoma, molecular definition of the susceptibility region on the Y chromosome. Am J Hum Genet 1995; 57(6): 1400- 1407.

[24] Repping S, Skaletsky H, Lange J, Silber S, Van der Veen F, Oates RD *et al*. Recombination between palindromes P5 and P1 on the human Y chromosome causes massive deletions and spermatogenic failure. Am J Hum Genet 2002; 71(4): 906- 922. http://dx.doi.org/10.1086/342928

[25] Pan Y, Kytölä S, Farnebo F, Wang N, Lui WO, Nupponen N *et al*. Characterization of chromosomal abnormalities in prostate cancer cell lines by Spectral Karyotyping. Cytogenet Cell Genet 1999; 87(3-4): 225-232. http://dx.doi.org/10.1159/000015432

[26] Vijayakumar S, Garcia D, Hensel CH, Banerjee M, Bracht T, Xiang R *et al*. The human Y chromosome suppresses the tumorigenicity of PC-3, a human prostate cancer cell line, in athymic nude mice. Genes Chromosomes Cancer 2005; 44(4): 365-372. http://dx.doi.org/10.1002/gcc.20250

[27] Jordan JJ, Hanlon AL, Al-Saleem TI, Greenberg RE, Tricoli JV. Loss of the short arm of the Y chromosome in human prostate carcinoma. Cancer Genet Cytogenet 2001; 124(2): 122- 6. http://dx.doi.org/10.1016/S0165-4608(00)00340-X

[28] Yuan X, LuML, LiT, Balk SP. SRY interacts with and negatively regulates androgen receptor transcriptional activity. J Biol Chem 2001; 276(49): 46647-54. http://dx.doi.org/10.1074/jbc.M108404200

[29] Ewis AA, Lee J, Naroda T, Sano T, Kagawa S, Iwamoto T, *et al*. Prostate cancer incidence varies among males from different Y-chromosome lineages. Prostate Cancer Prostatic Dis 2006; 9(3): 303-309. http://dx.doi.org/10.1038/sj.pcan.4500876

[30] Paracchini S, Pearce CL, Kolonel LN, Altshuler D, Henderson BE, Tyler-Smith C. A Y chromosomal influence on prostate cancer risk: the multi-ethnic cohort study. J Med Genet 2003; 40(11): 815-819. http://dx.doi.org/10.1136/jmg.40.11.815

[31] Kim W, Yoo TK, Kim SJ, Shin DJ, Tyler-Smith C, Jin HJ, *et al*. Lack of association between Y-chromosomal haplogroups and prostate cancer in the Korean population. PLoS ONE 2007; 24; 2(1): e172.

[32] Pathak D, Premi S, Srivastava J, Chandy SP, Ali S. Genomic instability of the DYZ1 repeat in patients with Y chromosome anomalies and males exposed to natural background radiation. DNA Res 2006; 13(3): 103-109. http://dx.doi.org/10.1093/dnares/dsl002

[33] Speltra FE, Garolla A, Selice R, Zuccarello D, Foresta C. Y chromosome haplogroups and susceptibility to testicular cancer. Mol Hum Reprod 2007; 13(9): 615-619. http://dx.doi.org/10.1093/molehr/gam052

[34] Peltomäki P, Lothe R, Borresen AL, Fosså SD, Brogger A, de la Chapelle A. Altered dosage of the sex chromosomes in human testicular cancer: a molecular genetic study. Int J Cancer 1991; 47(4): 518-22. http://dx.doi.org/10.1002/ijc.2910470408

[35] Speicher MR, Carter NP. The new cytogenetics: Blurring the boundaries with molecular biology. Nat Rev Genet 2005; 6(10): 782-792. http://dx.doi.org/10.1038/nrg1692

[36] Trask BJ. Human cytogenetics: 46 chromosomes, 46 years and counting. Nat Rev Genet 2002; 3(10): 769-778. http://dx.doi.org/10.1038/nrg905

[37] Buño I, Wyatt WA, Zinsmeister AR, Dietz-Band J, Silver RT, Dewald GW. A special fluorescence *in situ* hybridization technique to study peripheral blood and assess the effectiveness of interferon therapy in chronic myeloid leukemia. Blood 1998; 92(7): 2315-2321.

[38] Dewald G, Stallard R, Alsaadi A. A multicenter investigation with D-FISH BCR/ABL1 probes. Cancer Genet Cytogenet 2000; 116(2): 97-104. http://dx.doi.org/10.1016/S0165-4608(99)00120-X

[39] Kumar S, Pack S, Kumar D. Detection of EWS-FLI-1 fusion in Ewing's sarcoma/peripheral primitive neuroectodermal tumor by fluorescence *in situ* hybridization using formalin-fixed paraffin-embedded tissue. Hum Pathol 1999; 30(3): 324-330. http://dx.doi.org/10.1016/S0046-8177(99)90012-6

[40] Devriendt K, Fryns JP, Mortier G, van Thienen MN, Keymolen K. The annual incidence of DiGeorge/ velocardiofacial syndrome. J Med Genet 1998; 35(9): 789-790. http://dx.doi.org/10.1136/jmg.35.9.789-a

[41] Lewin P, Kleinfinger P, Bazin A, Mossafa H, Szpiro-Tapia S. Defining the efficiency of fluorescence *in situ* hybridization on uncultured amniocytes on a retrospective cohort of 27,407 prenatal diagnosis. Prenat Diagn 2000; 20(1): 1-6. http://dx.doi.org/10.1002/(SICI)1097-0223(200001)20:1<1::AID-PD739>3.0.CO;2-6

[42] Fan YS, Farrell SA. Prenatal diagnosis of interstitial deletion of 17(p11.2p11.2) (Smith-Magenis syndrome). Am J Med Genet 1994; 49(2): 53-254. http://dx.doi.org/10.1002/ajmg.1320490220

[43] Schrock E. Multicolor spectral karyotyping of human chromosomes. Science 1996; 273(5274): 494-497. http://dx.doi.org/10.1126/science.273.5274.494

[44] Speicher MR. Karyotyping human chromosomes by combinatorial multi-fluor FISH. Nat Genet 1996; 12(4): 368-375. http://dx.doi.org/10.1038/ng0496-368

[45] Levy B, Dunn TM, Kaffe S, Kardon N, Hirschhorn K. Clinical applications of comparative genomic hybridization. Genet Med 1998; 1(1): 4-12. http://dx.doi.org/10.1097/00125817-199811000-00004

[46] Weiss MM, Hermsen, MAJA, Meijer GA. Comparative genomic hybridization. Mol Pathol 1999; 52(5): 243-251. http://dx.doi.org/10.1136/mp.52.5.243

[47] Velagaleti GVN, Tharapel SA, Tharapel AT. Validation of primed *in situ* labeling (PRINS) for interphase analysis: Comparative studies with conventional fluorescence *in situ* hybridization and chromosome analysis. Cancer Genet Cytogenet 1999; 108(2): 100-106. http://dx.doi.org/10.1016/S0165-4608(98)00124-1

[48] Pellestor F, Andréo B, Coullin P. Interphase analysis of aneuploidy in cancer cell lines using primed *in situ* labeling. Cancer Genet Cytogenet 1999; 111(2): 111-118. http://dx.doi.org/10.1016/S0165-4608(98)00224-6

[49] Wang G, Whittam TS, Berg CM, Berg DE. RAPD (arbitrary primer) PCR is more sensitive than multilocus enzyme electrophoresis for distinguishing related bacterial strains. Nucl Acids Res 1993; 21(25): 5930-5933. http://dx.doi.org/10.1093/nar/21.25.5930

[50] Williams DL, Kline BC. Rapid identification of a point mutation of the Mycobacterium tuberculosis catalase-peroxidase (*katG*) gene associated with isoniazid resistance. J Infect Dis 1995; 171(1): 240-245. http://dx.doi.org/10.1093/infdis/171.1.240

[51] Kapur V, Prasantha SG, Ryan CO, Azfera, MA, Ali S. Development of a DNA marker by minisatellite associated sequence amplification (MASA) from the endangered Indian rhino (*Rhinoceros unicornis*). Mol Cell Probes 2003; 17(1): 1-4. http://dx.doi.org/10.1016/S0890-8508(02)00116-0

[52] Srivastava J, Premi S, Pathak D, Ahsan Z, Tiwari M, Garg LC, Ali S. Transcriptional status of known and novel genes tagged with consensus of 33.15 repeat loci employing minisatellite associated sequence amplification (MASA) and real time PCR in water buffalo *Bubalus bubalis*. DNA Cell Biol 2006; 25(1): 31-48. http://dx.doi.org/10.1089/dna.2006.25.31

[53] Srivastava J, Premi S, Kumar S, Ali S. Organization and differential expression of the GACA/GATA tagged somatic and spermatozoal transcriptomes in buffalo *Bubalus bubalis*. BMC Genomics 2008; 9: 132. http://dx.doi.org/10.1186/1471-2164-9-132

[54] Srivastava J, Premi S, Kumar S, Ali S. Expressional dynamics of minisatellite 33.15 tagged spermatozoal transcriptome in *Bubalus bubalis*. BMC Genomics 2009; 10: 303. http://dx.doi.org/10.1186/1471-2164-10-303

[55] Pathak D, Srivastava J, Samad R, Parwez I, Kumar S, Ali S. Genome wide search of the genes tagged with the consensus of 33.6 repeat loci in buffalo *Bubalus bubalis* employing minisatellite associated sequence amplification. Chromosome Res 2010; 18(4): 441-58. http://dx.doi.org/10.1007/s10577-010-9132-0

[56] Kumar S, Gupta R, Ali S. Molecular mining of alleles in water buffalo *Bubalus bubalis* and characterization of the *TSPY1* and *COL6A1* genes. PLoS ONE 2011; 6(9): e24958. http://dx.doi.org/10.1371/journal.pone.0024958

[57] Rawal L, Ali S, Ali S. Molecular mining of GGAA tagged transcripts and their expression in water buffalo *Bubalus bubalis*. Gene 2012; 492(1): 290-295. http://dx.doi.org/10.1016/j.gene.2011.10.004

[58] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 1977; 74(12): 5463-5467. http://dx.doi.org/10.1073/pnas.74.12.5463

[59] Hert DG, Fredlake CP, Barron AE. Advantages and limitations of next-generation sequencing technologies: a comparison of electrophoresis and non-electrophoresis methods. Electrophoresis 2008; 29(23): 4618-4626. http://dx.doi.org/10.1002/elps.200800456

[60] Schloss JA. How to get genomes at one ten-thousandth the cost. Nat. Biotechnol 2008; 26(10): 1113-1115. http://dx.doi.org/10.1038/nbt1008-1113

[61] Fuller CW, Middendorf LR, Benner SA, Church GM, Harris T, Huang X, *et al*. The challenges of sequencing by synthesis. Nat Biotechnol 2009; 27(11): 1013-1023. http://dx.doi.org/10.1038/nbt.1585

[62] Xue Y, Tyler-Smith C. The hare and the tortoise: one small step for four SNPs, one giant leap for SNP-kind. Forensic Scilnt Genet 2010; 4(2): 59-61. http://dx.doi.org/10.1016/j.fsigen.2009.08.005

[63] Xue Y, Wang Q, Long Q, Ng BL, Swerdlow H, Burton J, *et al*. Human Y chromosome base-substitution mutation rate measured by direct sequencing in a deep-rooting pedigree. Curr Biol 2009; 19(17): 1453-1457. http://dx.doi.org/10.1016/j.cub.2009.07.032

[64] Ballantyne KN, Goedbloed M, Fang R, Schaap O, Lao O, Wollstein A, *et al*. Mutability of Y-chromosomal microsatellites: rates, characteristics, molecular bases, and forensic implications. Am J Hum Genet 2010; 87(3): 341-353. http://dx.doi.org/10.1016/j.ajhg.2010.08.006

[65] Ansén S, Bangard C, Querings S, Gabler F, Scheffler M, Seidel D, *et al*. Osteoblastic response in patients with non-small cell lung cancer with activating EGFR Mutations and bone metastases during treatment with EGFR kinase inhibitors. J Thorac Oncol 2010; 5(3): 407-9. http://dx.doi.org/10.1097/JTO.0b013e3181cf32aa

[66] Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, *et al*. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. Am J Hum Genet 1998; 62(3): 676-89. http://dx.doi.org/10.1086/301749

[67] Gille JJ, Hogervorst FB, Pals G, Wijnen JT, van Schooten RJ, Dommering CJ, *et al*. Genomic deletions of MSH2 and MLH1 in colorectal cancer families detected by a novel mutation detection approach. Br J Cancer 2002; 87(8): 892-7. http://dx.doi.org/10.1038/sj.bjc.6600565

[68] Chen K, Wallis JW, McLellan MD, Larson DE, Kalicki JM, Pohl CS, *et al*. BreakDancer: An algorithm for high-resolution mapping of genomic structural variation. Nat Methods 2009; 6(9): 677-81. http://dx.doi.org/10.1038/nmeth.1363

[69] Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, *et al*. The mutational landscape of head and neck squamous cell carcinoma. Science 2011; 333(6046): 1157-60. http://dx.doi.org/10.1126/science.1208130

[70] Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, *et al*. The landscape of somatic copy-number alteration across human cancers. Nature 2010; 463(7283): 899-905. http://dx.doi.org/10.1038/nature08822

[71] Shah SP, Morin RD, Khattra J, Prentice L, Pugh T, Burleigh A, *et al*. Mutational evolution in a lobular breast tumor profiled at single nucleotide resolution. Nature 2009; 461(7265): 809-13. http://dx.doi.org/10.1038/nature08489

[72] Berger MF, Levin JZ, Vijayendran K, Sivachenko A, Adiconis X, Maguire J, *et al*. Integrative analysis of the melanoma transcriptome. Genome Res 2010; 20(4): 413-27. http://dx.doi.org/10.1101/gr.103697.109

[73] Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, *et al*. The complete genome of an individual by massively parallel DNA sequencing. Nature 2008; 452(7189): 872-876. http://dx.doi.org/10.1038/nature06884

[74] Girard A, Sachidanandam R, Hannon GJ, Carmell MA. A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature 2006; 442(7079): 199-202.

[75] Houwing S. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in zebrafish. Cell 2007; 129(1): 69-82. http://dx.doi.org/10.1016/j.cell.2007.03.026

[76] Meldrum C, Doyle MA, Tothill RW. Next-Generation Sequencing for Cancer Diagnostics: a Practical Perspective. Clin Biochem Rev 2011; 32(4): 177-195.

[77] Zhang J, Chiodini R, Badr A, Zhang G. The impact of next-generation sequencing on genomics. J Genet Genomics 2011; 38(3): 95-109. http://dx.doi.org/10.1016/j.jgg.2011.02.003

[78] Schmid CD, Bucher P. ChIP-Seq data reveal nucleosome architecture of human promoters. Cell 2007; 131(5): 831-832. http://dx.doi.org/10.1016/j.cell.2007.11.017

[79] Derrington IM, Butler TZ, Collins MD, Manrao E, Pavlenok M, Niederweis M, *et al*. Nanopore DNA sequencing with MspA. Proc Natl Acad Sci USA 2010; 107(37): 16060-16065. http://dx.doi.org/10.1073/pnas.1001831107

[80] Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, *et al*. Real-time DNA sequencing from single polymerase molecules. Science 2009; 323(5910): 133-138. http://dx.doi.org/10.1126/science.1162986

[81] Ozsolak F, Platt AR, Jones DR, Reifenberger JG, Sass LE, McInerney P, *et al*. Direct RNA sequencing. Nature 2009; 461(7265): 814-818. http://dx.doi.org/10.1038/nature08390

[82] Travers KJ, Chin CS, Rank DR, Eid JS, Turner SW. A flexible and efficient template format for circular consensus sequencing and SNP detection. Nucl Acids Res 2010; 38(15): e159. http://dx.doi.org/10.1093/nar/gkq543

[83] Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, Braslavsky I, *et al*. Single-molecule DNA sequencing of a viral genome. Science 2008; 320(5872): 106- 109. http://dx.doi.org/10.1126/science.1150427

[84] Tessler LA, Reifenberger JG, Mitra RD. Protein quantification in complex mixtures by solid phase single-molecule counting. Anal Chem 2009; 81(17): 7141-7148. http://dx.doi.org/10.1021/ac901068x

[85] Cohen JC, Kiss RS, Pertsemlidis A, Marcel YL, McPherson R, Hobbs HH. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. Science 2004; 305(5685): 869-872. http://dx.doi.org/10.1126/science.1099870

[86] Romeo S, Yin W, Kozlitina J, Pennacchio LA, Boerwinkle E, Hobbs HH, *et al*. Rare loss-of-function mutations in ANGPTL family members contribute to plasma triglyceride levels in humans. J Clin Invest 2009; 119(1): 70-79.

[87] Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, McMillin MJ, Gildersleeve HI, *et al*. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. Nat Genet 2010; 42(2): 790-793. http://dx.doi.org/10.1038/ng.646

[88] Volpi L, Roversi G, Colombo EA, Leijsten N, Concolino D, Calabria A, *et al*. Targeted next-generation sequencing appoints c16orf57 as clericuzio-type poikiloderma with neutropenia gene. Am J Hum Genet 2010; 86(1): 72-76. http://dx.doi.org/10.1016/j.ajhg.2009.11.014

[89] Nikopoulos K, Gilissen C, Hoischen A, vanNouhuys CE, Boonstra FN, Blokland EA, t*et al*. Next-generation sequencing of a 40 Mb linkage interval reveals TSPAN12 mutations in patients with familial exudative vitreoretinopathy. Am J Hum Genet 2010; 86(2): 240-247. http://dx.doi.org/10.1016/j.ajhg.2009.12.016

[90] St. Hilaire C. *NT5E* mutations and arterial calcifications. N Engl J Med 2011; 364(5): 432-42. http://dx.doi.org/10.1056/NEJMoa0912923

[91] Parla JS, Iossifov I, Grabill I, Spector MS, Kramer M, McCombie WR. A comparative analysis of exome capture Genome Biol 2011; 12(9): R97. http://dx.doi.org/10.1186/gb-2011-12-9-r97

[92] Worthey EA. Making a definitive diagnosis: successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. Genet Med 2011; 13(3): 255-262. http://dx.doi.org/10.1097/GIM.0b013e3182088158

[93] Yeoh EJ, Ross ME, Shurtleff SA, Williams WK, Patel D, Mahfouz R, *et al*. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell 2002; 1(2):133-43. http://dx.doi.org/10.1016/S1535-6108(02)00032-6

[94] Grouse LH, Munson PJ, Nelson PS. Sequence databases and microarrays as tools for identifying prostate cancer biomarkers. Urology 2001; 57(4 Suppl 1): 154-9. http://dx.doi.org/10.1016/S0090-4295(00)00963-8

[95] Brenton JD, Aparicio SA, Caldas C. Molecular profiling of breast cancer: portraits but not physiognomy. Breast Cancer Res 2001; 3(2): 77-80. http://dx.doi.org/10.1186/bcr274

[96] Kim JH, Skates SJ, Uede T, Wong KK, Schorge JO, Feltmate CM, *et al*. Osteopontin as a potential diagnostic biomarker for ovarian cancer. JAMA 2002; 287(13): 1671-9. http://dx.doi.org/10.1001/jama.287.13.1671

[97] Wong KK, Cheng RS, Mok SC. Identification of differentially expressed genes from ovarian cancer cells by MICROMAX cDNA microarray system. Biotechniques 2001; 30(3): 670-5.

[98] Hoque MO, Lee CC, Cairns P, Schoenberg M, Sidransky D. Genome-wide genetic characterization of bladder cancer: a comparison of high-density single-nucleotide polymorphism arrays and PCR-based microsatellite analysis. Cancer Res 2003; 63(9): 2216-22.

[99] Dumur CI, Dechsukhum C, Ware JL, Cofield SS, Best AM, Wilkinson DS, *et al*. Genome-wide detection of LOH in prostate cancer using human SNP microarray technology. Genomics 2003; 81(3): 260-9. http://dx.doi.org/10.1016/S0888-7543(03)00020-X

[100] Bignell GR, Huang J, Greshock J, Watt S, Butler A, West S *et al*. High-resolution analysis of DNA copy number using oligonucleotide microarrays. Genome Res. 2004; 14(2): 287-95. http://dx.doi.org/10.1101/gr.2012304

[101] Haugen A. Progress and potential of genetic susceptibility to environmental toxicants. Scand J Work Environ Health 1999; 25(6): 537-40. http://dx.doi.org/10.5271/sjweh.477

[102] Eyster KM, Lindahl R. Molecular medicine: a primer for clinicians. Part XII: DNA microarrays and their application to clinical medicine. S D J Med 2001; 54(2): 57-61.

[103] Ivanov I. DNA microarray technology and antimicrobial drug discovery. Pharmacogenomics 2000; 1(2): 169-78. http://dx.doi.org/10.1517/14622416.1.2.169

[104] Marton MJ. Drug target validation and identification of secondary drug target effects using DNA microarrays. Nat Med 1998; 4(11): 1293-301. http://dx.doi.org/10.1038/3282

[105] Luo Z, Geschwind DH. Microarray applications in neuroscience. Neurobiol Dis 2001; 8(2): 183-193. http://dx.doi.org/10.1006/nbdi.2001.0387

[106] Sevignani C, Calin GA, Siracusa LD, Croce CM. Mammalian micro RNAs: a small world for fine-tuning gene expression. Mamm Genome 2006; 17(3): 189-202. http://dx.doi.org/10.1007/s00335-005-0066-3

[107] Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP. Prediction of plant microRNA targets. Cell 2002; 110(4): 513-520. http://dx.doi.org/10.1016/S0092-8674(02)00863-2

[108] Wiemer EA. The role of microRNAs in cancer: no small matter. Eur J Cancer 2007; 43(10): 1529-1544. http://dx.doi.org/10.1016/j.ejca.2007.04.002

[109] Nelson PT, Keller JN. RNA in brain disease: no longer just "the messenger in the middle". J Europathol Exp Neurol 2007; 66(6): 461-468. http://dx.doi.org/10.1097/01.jnen.0000240474.27791.f3

[110] Kim VN, Nam JW. Genomics of microRNA. Trends Genet 2006; 22(3): 165-173. http://dx.doi.org/10.1016/j.tig.2006.01.003

[111] Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D. MicroRNA expression profiles classify human cancers. Nature 2005; 435(7043): 834-838. http://dx.doi.org/10.1038/nature03702

[112] Esquela-Kerscher A, Slack FJ. Oncomirs - microRNAs with a role in cancer. Nat Rev Cancer 2006; 6(4): 259-269. http://dx.doi.org/10.1038/nrc1840

[113] Griffiths PE and Stotz K. genes in the postgenomic era Theoretical Medicine and Bioethics. Theor Med Bioeth 2006; 27(6): 499-521. http://dx.doi.org/10.1007/s11017-006-9020-y

[114] Lukiw WJ. Micro-RNA speciation in fetal, adult and Alzheimer's disease hippocampus. Neuroreport 2007; 18(3): 297-300. http://dx.doi.org/10.1097/WNR.0b013e3280148e8b

[115] Nelson PT, Keller JN. RNA in brain disease: no longer just "the messenger in the middle". J Europathol Exp Neurol 2007; 66(6): 461-468. http://dx.doi.org/10.1097/01.jnen.0000240474.27791.f3

[116] Krutzfeldt J, Stoffel M. MicroRNAs: a new class of regulatory genes affecting metabolism. Cell Metab 2006; 4(1): 9-12. http://dx.doi.org/10.1016/j.cmet.2006.05.009

[117] Ferguson PL, Smith RD. Proteome analysis by mass spectroscopy. Annu Rev Biophys Biomol Struct 2003; 32: 399-424. http://dx.doi.org/10.1146/annurev.biophys.32.110601.141854

[118] Wolters DA, Washburn MP, Yates JR 3rd. An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem 2001; 73(23): 5683-5690. http://dx.doi.org/10.1021/ac010617e

[119] Foss EJ. Genetic basis of proteome variation in yeast. Nature Genet 2007; 39(11): 1369-1375. http://dx.doi.org/10.1038/ng.2007.22

[120] Jenkins LM. Quantitative proteomics analysis of the effects of ionizing radiation in wild type and p53 K317R knock-in mouse thymocytes. Mol Cell Proteomics 2008; 7(4): 716-727. http://dx.doi.org/10.1074/mcp.M700482-MCP200

[121] Martin B. iTRAQ analysis of complex proteome alterations in 3xTgAD Alzheimer's mice: understanding the interface between physiology and disease. PLoS ONE 2008; 3(7): e2750.

[122] Chiang MC. Systematic uncovering of multiple pathways underlying the pathology of Huntington disease by an acid-cleavable isotope-coded affinity tag approach. Mol Cell Proteomics 2007; 6(5): 781-797. http://dx.doi.org/10.1074/mcp.M600356-MCP200

[123] Liao L, Park SK, Xu T, Vanderklish P, Yates JR 3rd. Quantitative proteomic analysis of primary neurons reveals diverse changes in synaptic protein content in *fmr1* knockout mice. Proc Natl Acad Sci USA 2008; 105(40): 15281-15286. http://dx.doi.org/10.1073/pnas.0804678105

[124] Kruger M. SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. Cell 2008; 134(2): 353-364. http://dx.doi.org/10.1016/j.cell.2008.05.033

[125] McClatchy DB, Liao L, Park SK, Venable JD, Yates JR. Quantification of the synaptosomal proteome of the rat cerebellum during post-natal development. Genome Res 2007; 17(9): 1378-1388. http://dx.doi.org/10.1101/gr.6375007

[126] Alagaratnam S. Serum protein profiling in mice: identification of factor XIIIa as a potential biomarker for muscular dystrophy. Proteomics 2008; 8(8): 1552-1563. http://dx.doi.org/10.1002/pmic.200700857

[127] Faca VM. A mouse to human search for plasma proteome changes associated with pancreatic tumor development. PLoS Med 2008; 5(6): e123.

[128] Hung KE. Comprehensive proteome analysis of an Apc mouse model uncovers proteins associated with intestinal tumorigenesis. Cancer Prev Res 2009; 2(3): 224-233. http://dx.doi.org/10.1158/1940-6207.CAPR-08-0153

[129] Ruddat VC, Whitman S, Klein RD, Fischer SM. Evidence for downregulation of calcium signaling proteins in advanced mouse adenocarcinoma. Prostate 2005; 64(2): 128-138. http://dx.doi.org/10.1002/pros.20207

[130] Gao BB, Clermont A, Rook S, Fonda SJ. Extracellular carbonic anhydrase mediates hemorrhagic retinal and cerebral vascular permeability through prekallikrein activation. Nat Med 2007; 13(2): 181-188. http://dx.doi.org/10.1038/nm1534

Received on 24-09-2013 Accepted on 11-10-2013 Published on 30-11-2013