Journal of King Saud University – Science 34 (2022) 101961

Contents lists available at ScienceDirect

Journal of King Saud University – Science

journal homepage: www.sciencedirect.com

Review

Biological and genetic basis of various human genetic disorders and the application of biological and genetic markers

Atif Abdulwahab A. Oyouni [a,b]

[a] Department of Biology, Faculty of Sciences, University of Tabuk, Tabuk, Saudi Arabia
[b] Genome and Biotechnology Unit, Faculty of Sciences, University of Tabuk, Tabuk, Saudi Arabia

ARTICLE INFO

Article history:
Received 19 January 2022
Revised 20 February 2022
Accepted 7 March 2022
Available online 11 March 2022

Keywords:
Antigenic marker
Non-infectious diseases
Genetic diseases
Aberrant DNA
Chromosomal instability

ABSTRACT

The antigens of the ABO blood group systems are expressed on the membrane of the red blood cells and on the surface of several pathological and normal cells and tissues. Following this earlier investigations, the pathological state of genetic diseases and disorders were determined from the blood fluids and blood cells. Biology of most of the genetic diseases was associated with the development of marks and the ABO antigens were associated with the development of various tumours, namely gastric and pancreatic cancers. ABO antigens are used as the prognostic biomarkers in various types of cancers in human. Moreover, the association of these antigenic effects is uncertain. In human, several epigenetic marks required for the normal development. These include DNA methylation at GpG dinucleotides, noncoding RNAs and covalent modifications of histone. These functions are regulated in organized manner, regulating mitotically heritable changes in the expression of genes without changing the primary sequences of DNA. Any changes in the expression of these proteins due to genetic or environmental factors affect normal function and leads to aberrant epigenetic pattern. These aberrant epigenetic patterns lead to various human disorders, including imprinting and sub-fertility. The present review discusses the biology and genetics of various diseases and highlights the important theoretical and technical problems. Human genome sequences and other technical implications in the line of genetic diseases are discussed in this review.

Contents

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

E-mail address: a.oyouni@ut.edu.sa

https://doi.org/10.1016/j.jksus.2022.101961
1018-3647/

1. Introduction

ABO blood group was discovered by Karl Landsteiner in 1901 and human blood group was classified into various types, including A, B, O, AB on the basis of presence or absence of antigens on the blood cell surface (Karl Landsteiner, 2020). The expression of ABO gene is not only limited with red blood cells but also in tissue surfaces, including, sensory neurons, epithelial cells, kidney, endothelial cells, heart, platelets, pancreas and lungs and revealed the significance of ABO system in clinical practice beyond blood transfusion and immunohematology (Franchini and Bonfanti, 2015). Endoplasmic reticulum (ER) malfunction might have potential impact on various neurological disorders including multiple sclerosis and Alzheimer's disease. ER is one of the important and largest organelles in the human cell and is an important site of protein synthesis, protein folding and transport, carbohydrate metabolism, steroid and lipid biosynthe sis and calcium storage (Westrate et al., 2015). Consequently, variations in the structure and composition of nuclear structure are reported in various genetic diseases. In normal state, nucleus has a unique substructure and it is dynamic in structure. Oncogene expression, viral infection, and inherited human disorders can cause various changes in nuclear organization. Chromosomal instability (CIN) is one of the factors associated with certain diseases including, cancer (Sathyamoorthi et al., 2019; Kumaresan et al., 2019; Sathyamoorthi et al., 2018). CIN is termed as misleading segregation of chromosomes and is the problem during mitosis process. Chromosomal instability is the process of loss and gain of chromosomes than normal state. CIN is one of the important markers of genome instability linked with solid cancer. Werner syndrome (WS) is one of the rare autosomal recessive disorder and this disease is characterized by instability of genomic DNA and by onset of age-related diseases, including cancers. Likewise, cystic fibrosis is an autosomal recessive, leathal disorder among Caucasians. Zellweger syndrome is one of the peroxisome biogenesis disorders and this disease was caused by defects in 13 genes and called PEX genes. Amyotrophic lateral sclerosis (ALS) is a fetal neurodegenerative disease that causes muscle weakness because of weakness in the motor neurons. Epigenetic changes involved in the development of various disease, including, cancer, cardiovascular disease, metabolic disease, autoimmune disease and neurodegenerative disease and most of them related to aging (Jung and Pfeifer, 2015). Whole genome sequencing as well as whole exome sequencing is used for the determination of genetic diseases. Genome sequences are useful to provide valuable information on the status of the genetic diseases and drug response efficacy. The whole genome of a patient with a family history of vascular disease and very early death but no critical clinical record, and determined elevated post-test probability risks of coronary artery disease and myocardial infarction (Pushkarev et al., 2009). In addition to the whole genome sequencing approach transcriptome, epigenome, metabolome and proteome of the human body and omics approach such as microRNA and gut microbiome and various immune receptor may also be significant factor for the monitoring of health and personalized medicine, either in combination with other iPOP omics or alone (Ravichandran et al., 2018; Sathyamoorthi et al., 2017; Ravichandran et al., 2017). The present review summarize the biological and genetic basis of various ABO blood group, endoplasmic reticulum and nucleus and determination of rare human disorders including rare human cancers and perspectives on personalized medicine using whole genome and whole exome sequencing.

2. Biological basis of human ABO blood group

The ABO gene consists of seven exons and occupied on chromosome 9 and encodes an enzyme glycosyltransferase that involved in the formation of antigenic structures on the surface of the blood cells (Clausen et al., 1994). In A and B alleles of the ABO genes, the amino acid coding sequence varied, and they catalyze the transfer of various carbohydrate molecules (galactose or N-acetylgalactosamine) to form the B and A antigen, respectively. Moreover individuals associated with O blood group do not produce B or A antigens due to a deletion of single-base nonsense (Yamamoto et al., 2012). Single-nucleotide polymorphism (SNPs) has been detected in ABO genes and over 100 alleles have been described previously (Fokkema et al., 2011). Amundadottir et al. (2009) have been reported pancreatic cancer cohort variations associated with cancer development and were mapped to the ABO locus. Schunkert et al. (2011) studied the genetic variant of rs579459 in ABO alleles was mainly liked with serious risk of coronary artery diseases. Blood group antigens were determined in body fluids such as saliva, sweat, seminal fluid, urine, amniotic fluid, and gastric secretions (Ewald and Sumner, 2016). However, most of the antigens are the end product of a single gene, and alterations at the genetic level such as, SNPs, alternative splicing, insertions, inversion and deletions leading to variation in antigens and can develop new antigen and may be loss of expression (Denomme, 2011). ABO blood groups have been associated with various infections and involved in non-infectious diseases (Mäkivuokko et al., 2012).

3. Endoplasmic reticulum and functions

Structurally, ER is composed of various different structural domains, and each of which is linked with a specific function. Moreover the roles of various structural domains have not been completely reported. ER is a major site for protein synthesis for integral or secreted membrane proteins as well as biosynthesis of cytosolic proteins (Jan et al., 2014). Protein biosynthesis required identification of ribosomes to the cytosolic region of the ER and canonical pathway that involved co-translational docking of ribosome complex and the mRNA on the membrane of ER. After protein synthesis and translocation of protein into the ER lumen, the secreted proteins undergo various modifications, and folding with the use of various enzymes and chaperones. These modifications include the formation of disulfide bond, N-linked glycosylation and oligomerization process (Braakman and Hebert, 2013). ER is

the major site for the biosynthesis of Ca^{2+} and the concentration of Ca^{2+} in cytosole was ~ 100 nM, whereas the Ca^{2+} concentration was ranged between 100 and 800 μM in the lumen of the endoplasmic reticulum, moreover the extracellular concentration of Ca^{2+} was ~ 2 mM (Samtleben et al., 2013). The ER has various calcium channels, 1,4,5-trisphosphate (IP_3) receptors, and ryanodine receptors involve in the release of Ca^{2+} from the ER into the cytosol and maintains intracellular calcium level. ER is also involved in the biosynthesis of lipid.

4. Nucleus structure and nuceloorganization changes in diseased condition

Nucleus was first invented by Brown in 1831 and the cellular nucleus is one of the important organelles and was least understood. The structure and functional properties of nucleus was not completely elucidated. The nucleus is largely disordered, membrane bound DNA and various other molecules and the structures are dispersed as a result of replication, transcription, and RNA processing activities in different regions of DNA. Analysis of organization principles of the nucleus, including the arrangement of chromosomal DNA, synthesis, assembly, and transport of various synthesized macromolecules in a coordinated manner (Lekkerkerker and Tuinier, 2011). About 500 KD protein would diffuse one side of the nucleus to the other side. The nucleolous is formed around the rDNA repeats which formed cluster at chromosomal loci termed nucleolar organizers and this region is the factory in which 5.8S, 18S and 28S rRNAs are processed, transcribed and assembled into subunits of ribosomes. The formation of nucleolus is cell cycle and transcription dependent. These nucleolus breaks down and during cell cycle. Chromosomes occupy about ~ 10% of the whole nuclear volume and the total volume of macromolecules ranged between ~ 20% and ~ 40%. Nuclei have different functions and the components of nucleus have no specific external membrane except chromosomes. The chromosomes are dynamic, mobile or can fuse and divide. The structural components are essential for RNAs and the mechanism of components formation has been unclear (Wang et al., 2004). DNA is associated with nucleosomes as a polyelectrolyte polymer and involved in a regular helical conformation and the diameter was approximately ~ 30 nm. The irregular structure has also been reported previously (Sinclair et al., 2010).

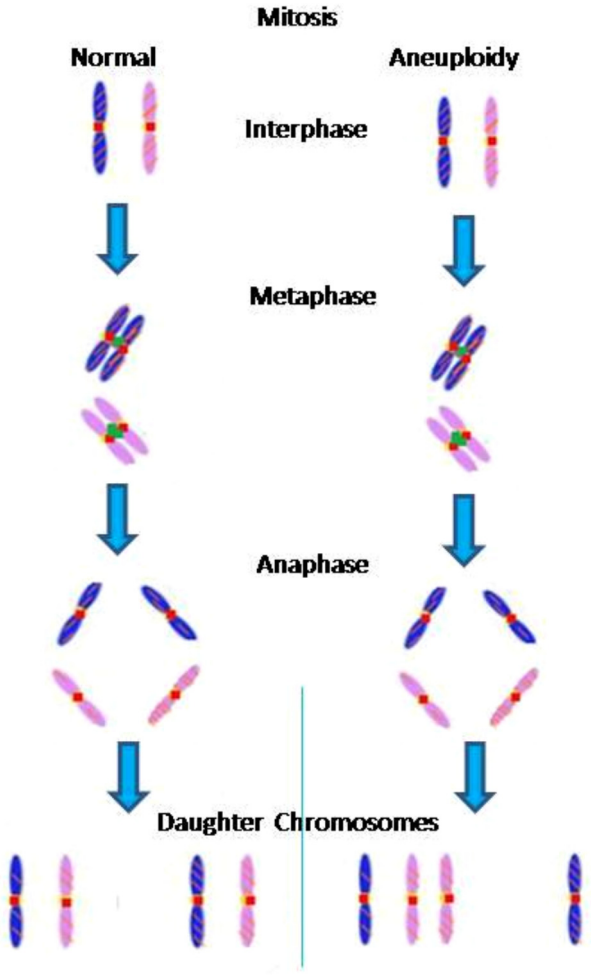

Fig. 1. Normal mitotic cell division and aneuploidy in the development of tumorigenesis.

5. Oncogenes and chromosomal instability

The solid tumours are generally characterized by aneuploidy (abnormal chromosome numbers) and karyotypic analysis revealed that the most of these solid tumours are chromosomally unstable and heterogenous. Chromosomal instability (CIN) is termed as misleading segregation of chromosomes and is the problem during mitosis process. Karyotypic analyses reveal that tumours show both inter- and intra-tumour heterogeneity revealing that most tumours are chromosomally unstable and aneuploid in nature. Chromosomal instability is the process of loss and gain of chromosomes than normal state (Thompson et al., 2010). CIN is one of the important markers of genome instability linked with solid cancer. CIN is caused by defects in chromosome segregation during mitotic division and most of the cancer cells show aberrant chromosome segregation. The scheme of chromosomal abnormality in mitosis is described in Fig. 1. These aberrant chromosomes could arise through either by imbalances in protein levels, or mutation of genes encoding several mitotic proteins or reduced mitotic fidelity and has been extensively analyzed (Nicholson and Cimini, 2013). The important outcome of CIN is to make variation in the karyotype. Moreover, CIN and aneuploidy were different and has been reported previously by Thompson et al. (2010). Aneuploidy is associated with abnormal chromosome number and, some tumours also have abnormal chromosome nature and tumour cells also have defective karyotype. CIN is an increased rate of chromosome mis-separation that improves karyotypic diversity in mammalian cells within the specific tumour, a common feature associated to aggressive behaviour of tumour (Brown and Geiger, 2018). CIN and aneuploidy have been associated with metastasis, poor patient prognosis, and resistance to various chemotherapeutic agents. Schvartzman et al. (2010) reported that CIN and aneuploidy play an aggressive role in relapse and tumorigenesis. Hence, analysis of underlining mechanisms that involve in CIN improved the efficacy of cancer treatment.

6. Molecular mechanism of Werner syndrome

Werner syndrome (WS) is one of the rare autosomal recessive disorder and this disease is characterized by instability of genomic DNA and by onset of age-related diseases, including cancers. This disease has been first described by Otto Werner in 1904 and the principle characters of WS has been studied in late 1940 s (Epstein et al., 1985). The phenotype variability has been reported among the patients, and this syndrome is characterized by age-related diseases such as diabetes mellitus, osteoporosis, ocular cataracts, atherosclerosis, early graying of the hair and various types of neoplasm. WS was also includes characters not associated with

age related diseases such as, short stature, subcutaneous atrophy, hyperkeratosis, telangiectasia, trophic ulcers of the legs, increased hyaluronic acid in the urine, calcification of the blood vessels, reduced fertility and hypogonadism. In some cases, rare type of cancers has been reported (Goto et al., 1996). WS is rare among world population (1/1,000,000) and this incidence is somewhat common amongst Japanese population because of consanguineous marriages (Miki et al., 1997). About 25 various mutations among WS patients have been reported in 32 exons of the WRN gene. These mutations include stop codons, deletions, insertions and exon deletion (Shen and Loeb, 2000). DNA damage is common phenomena and most of the chemicals involved in DNA damage. There are various mechanisms in mammalian cells to repair the damaged DNA molecule and these include, one step reactions, multistep reactions and single- and multistep base excision mechanisms. There are several proteins (helicases or exonucleases) involved in these repairing mechanisms. Bur failure in these repairing mechanism leads to the development of DNA lesions and mutations in genes involved in cell cycle regulation and cellular metabolism. DNA damage can leads to DNA breaks and chromosomal abnormality. *WRN* is the gene encodes a RecQ-type DNA helicase and is responsible for the disease and this enzyme involved in chromosome repair, recombination and maintenance of chromosome integrity during DNA replication process (Rossi et al., 2010).

7. Molecular mechanism of cystic fibrosis

Cystic fibrosis is an autosomal recessive, leathal disorder among Caucasians. In general population the frequency was about 1:2,500 at birth and the carrier frequency was 1:25. Moreover, the frequency of this disease varied based on ethic, and geographical locations. This disease is characterized by pancreatic dysfunction, progressive lung disease, male infertility and elevated sweet electrolytes. In later 1980 s, CF gene was cloned and cystic fibrosis transmembrane conductance regulator (CFTR) was determined (Kerem et al., 1989). The general CFTR mutation was characterized by 3 bp deletion and this leads to the removal of amino acid residue (phenylalanine) at 508 region. This mutation was determined in more than 70% of the chromosome of CF patients, however its frequency varied based on population groups. The frequency of the Δ508 mutation was 90% in Denmark and it decreased as 30% in Turkey. The frequency increased over 70% in the population groups such as, France, England and Netherlands (Shoshani et al., 1992). In recent times, haplotype analysis was performed to study the evolution of the ΔF508 mutation among population in Europe. The findings revealed that the ΔF508 mutation started before 52,000 years and the genomic organization was different from the present European group. The CFTR gene consists of 250 kb and 27 exons in chromosome 7q31. CFTR gene involved in the regulation of other proteins. About 500 various CFTR mutations were determined and most of the mutations were rare. A total of 10 mutations were determined in over 100 patients (Gabriel et al., 1993).

8. Molecular mechanism of Zellweger syndrome

Zellweger syndrome is one of the peroxisome biogenesis disorders and this disease was caused by defects in 13 genes and called PEX genes. These PEX genes are required for the functions of peroxisomes. The disorders are classified into Rhizomelic Chondrodysplasia Punctua spectrum and Zellweger spectrum disorders. Peroxisomes are required for the development of normal brain development and the defects in PEX genes affected brain fuction. These defects affect the biosynthesis of oils and fatty acids necessary for the cell function. Peroxisomes are required for the development of myelin. These are essential for the typical function of bone, kidney, eye and liver functions (Brosius and Gärtner, 2002). Symptoms of these disorders include large forehead, enlarged liver, underdeveloped eyebrow ridges, seizures and mental retardation. Infants unable to suck or swallow or unable to move and in some case, gastrointestinal bleeding, jaundice, impaired hearing, retinal degeneration and glaucoma were reported. It is an autosomal disorder caused by mutations in genes that encode peroxins and lack of this protein impaired normal structural assembly of peroxisomes. ZS patients have mutations in various PEX genes and these genes affect the activity of enzyme (Lucaccioni et al., 2020).

9. Molecular mechanism of motor neuron disease

Amyotrophic lateral sclerosis (ALS) or Motor neuron disease (MND) is a fetal neurodegenerative disease that causes muscle weakness because of weakness in the motor neurons. However, non-motor pathways also contributed this symptom in more than 50% of the diagnosed patients. MND or ALS can be classified as familial (fALS) or sporadic (sALS). Biomarkers are widely used to determine this disease, improved diagnosis and to develop new therapeutic agents. In recent years an attempt has been made to validate the markers, however only few biomarkers have been reported. Blood biomarkers are widely used to determine peripheral and central damage in ALS. The patients associated with MND disease showed increased methylation of various components and has been reported in some clinical specimens. The failure of motor neurons leads to paralysis, which leads to respiratory failure and leads to death. This disease is not curable and factors such as, oxidative stress, mitochondrial dysfunction, oxidative stress and various genetic alterations were determined. ALS related genes have been determined from the patients and some genetic markers are useful for future gene therapeutics. One of the most causes of ALS is the gene mutation encoded with superoxide dismutase 1 (Velde et al., 2008). The mutated SOD1 has poor stability and causes aggregation in the motor neurons in the central nervous system. Mitochondrial dysfunction and glutamate excitotoxicity played significant role in the degeneration of motor neuron (Forsberg et al., 2011).

10. Molecular mechanism of lysosomal storage diseases

Lysosomal storage diseases (LSDs) are genetic diseases and errors of metabolism that critically affect the normal function of the lysosome. LSDs consist of about 70 monogenic disorders and most of these were inherited with X-chromosomes. These disorders are due to mutations in genes coded for lysosomal proteins, such as, protease, lysosomal glycosidases, transporters, enzyme activators, enzyme modifiers and integral membrane proteins. Impaired function of lysosomal genes due to mutation, affect the function of lysosomes and accumulate various substrate in the lysosome, which leads to cell death and in some cases cell dysfunction. About $\sim$ 1,300 genes involved in the regulation of lysosomal function and various monogenic disorders were described previously, these were classified based on storage substance. LDSs are clinically and genetically heterogenous disorders. The classification of this disease was mainly based on the molecular basis. This type of classification helps to evaluate the mechanisms based on lysosomal storage and based on malfunction. Recently, Filocamo and Morrone (2011), reported classification of LSDs based on laboratory observations. Anderson Fabry disease or Fabry disease among one which was first identified in 1898 and this disease was linked with –X caused by the deficiency of galactosidase (Politei et al., 2009). The deficiency of galactosidase induced the accumulation of glycosphingolipids in lysosomes. In Fabry disease, about 600

mutations has been reported, including small insertions, deletions, missense, splicing, and nonsense mutations (Politei et al., 2009).

11. Angiogenesis, prognosis and treatment

Metastasis and tumour growth were mainly based on lymphangiogenesis and angiogenesis triggered by various signals from tumour cells during growth phase. The behaviour and growth of cancer cells varied based on region in the same organ. Cancer cells grow up to 1–2 mm^3 and stopped, however the growth was beyond 2 mm^3 in some places where angiogenesis is possible. Tumours may become apoptotic or necrotic in the absence of vascular support (Hirai et al., 2001). Hence, angiogenesis is one of the important factors in the development of cancer. Angiogenesis process is initiated when tumour tissues require oxygen and nutrient. This process is tightly regulated by angiogenesis inhibitor and regulator molecules. More than 12 proteins have been determined as angiogenic activators, including, basic fibroblast growth factor, vascular endothelial growth factor, transforming growth factor, angiogenin, platelet-derived endothelial growth factor, tumour necrosis factor (TNF)-α, placental growth factor, granulocyte colony-stimulating factor, and platelet-derived endothelial growth factor. The naturally occurring proteins such as, interferon, endostatin, angiostatin, thorombospondin, platelet factor 4, tissue inhibitor of metalloproteinase-1, −2, and prolactin 16 kd fragment. Angiogentic activators played a key role in the spread of tumours. Immunohistochemical analysis revealed that, the VEGF family and their receptors have expressed in 50% of human cancers. These factors critically affect the prognosis of adenocarcinomas that developed in the endometrium, uterine cervix (Hashimoto et al., 2001), stomach, and ovary and colorectal cancer (André et al., 2000). Also, significant correlation has been detected between the expression of VEGF and prognosis in colorectal cancerm lung cancer, Kaposi sarcoma, neck and head squamous cell carcinoma (O-Charoenrat et al., 2001). The aggressiveness of tumour growth and spread varied based on the levels of angiogenic factors in tissues. Cancer cells require locating blood vessels for routine metastasis and growth. Anginogenic inhibitors have lot of potential to treat cancer disease. These anti-angiogenic molecules target the VEGF pathways, including, decoy receptor or VEGF trap, small-molecule tyrosine kinase inhibitors—TKIs, and VEGFR2 inhibitors. These drug types have several advantages than tumour cell targeting drugs. Bevacizumab, is one of the monoclonal antibodies, involved in anti-angiogenesis process. It binds with circulating VEGF and leads to affect the blood and oxygen supply to the tumour cells (Kazazi-Hyseni et al., 2010).

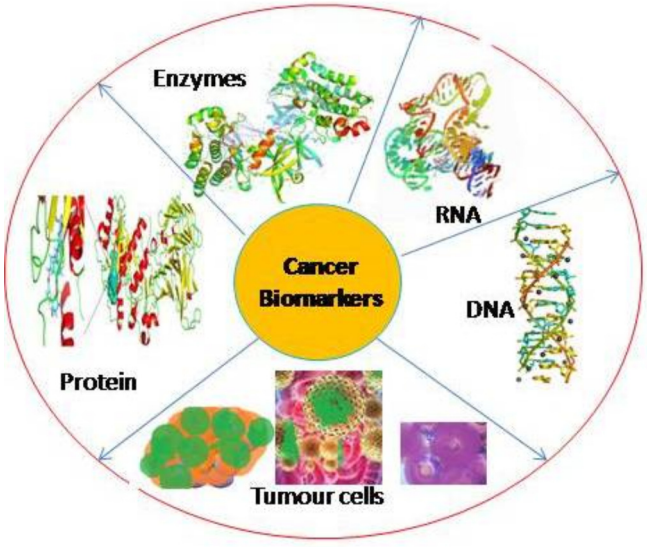

Fig. 2. Molecular makers used for the determination of cancers.

12. Molecular basis of tumour makers

Gene level changes in a tumour cell affects gene expression pattern directly or indirectly in the tumour cells and the surrounding tissues. These gene level variations reflected genetic defects and forming various tumour markers (Johnson, 2001). Generally tumour markers can be identified either in body fluids or tissue. The identification of solid tumour maker is based on the determination of tumour cells in tissues. In recent years, attempts have been made to use tumour maker along with other tools for the determination of cancer, these include carbohydrate antigen for the detection of ovarian malignancy. Moreover, these methods have not approved by various clinical organization (Harris et al., 2007). Tumour markers are useful for the monitoring the success or the failure of the disease. Many molecular markers are used for the determination of types of cancers and described in Fig. 2. The available quantity of the tumour marker reflects the level of tumour in patients and is used to determine the prognosis. Moreover, the level of tumour markers may vary based on the laboratory conditions which alter the diagnostic results (Garg et al., 2015). The tumour makers have several advantages; however determination of cancer in early stage is not possible in most of the cases. Tumour markers also used to diagnose the origin of cancer. Tumour markers are widely used to monitor the conditions of the patients being treated for cancer. The decreased level cancer markers indicated effective treatment, whereas increasing level of marker suggesting alternate treatment. Immunomarkers also used to detect cancers that recur after primary treatment (Nagpal et al., 2016).

13. Molecular and technical prospects of human genome

In human, the genetic information is stored in two different organelles, namely nucleus and mitochondria. The human chromosomes are unequal sizes; the largest chromosome 1 is 249 mln bp, while the smallest chromosome 21, is 54 mln bp size. Genomic sequences can be classified into various ways. From the functional perspective point of view, it is classified as genes, non-coding DNA and pseudogenes. About 3% of human genomes codes for various proteins transposable elements and with time some genes have mutated gradually beyond recognition. In human genome, about 0.5% genome consists of intergenic DNA and introns. The human genome consists of DNA and estimates that about 60% of the DNA is either low number of copies or singly copy and 30% of the estimated DNA is moderately repetitive. And, about 10% of the DNA was considered as highly repetitive characters. Many staining methods have been used to determine the functional properties of these DNA types and determined alternative banding patterns of mitotic chromosomes, karyograms. These three major classes of DNA are scattered throughout the analysed chromosome and these reflected the levels of compartmentalization of the analyzed DNA. C-banding (or chromosome banding) techniques revealed very dark strained region (C-bands) in the chromosome and is called as heterochromatin. The heterochromatin region is highly coiled, and is mainly found at the region of telomeres, centromeres and on the Y chromosome. Moreover the DNA composed of arrays of tandem repeats and hence different nucleotide composition that varies widely from the genome and estimated about 40–42% GC in the DNA. These tandem differences can be useful for the separation of genome based on density gradient centrifugation.

Gradient density centrifugation revealed the presence of three minor band and a major band and the minor bands are called as satellite bands or satellite DNA (Makałowski and Boguski, 1998).

The G-banding method shows a pattern of alternate dark and light bands revealing differences in base composition, chromatin conformation, time of replication, repetitive sequences and the density of genes. Hence, the karyograms define analysis of chromosomal organization and allow for characterization of the human chromosomes. The G-bands are darker and are highly dense bands, rich AT, and less gene rich regions. The human genome consists of more than 300 kb size fragment of DNA that is referred to as isochores and is homogenous in nature based on previous studies (Bernardi, 2000). GC-less isochore families are classified as L1 and L2 (light) and representing about 62% of total human genome. The GC-rich region isochore was classified as H1, H2 and H3 (heavy). The correlation between chromosomal bands and isochors was determined. In human chromosomes, the chromosomes 13, 14, 15, 21, and 22 were varied at their terminus region by a very thin bridge with rounded ends and referred as chromosomal satellites. These satellite regions contain coded regions for ribosomal proteins and rRNA that involved in the formation of nucleolus. In, human, the number of human genes is not clearly reported earlier. It was estimated between 28,000 and 80,000 genes in human haploid. In most of the human genes, the non-coding sequences interrupt the coding sequence and the non-coding regions are spliced out while mRNA maturation. The human genes are mosaic nature, consisting *introns* and *exons*. Introns carry potential information and also coded for various other complete genes. The availability of introns was determined in almost all domain of mRNA molecule. A total of 4731 human exons have been determined by Cold Spring Harbor Laboratory (Zhang, 1998). The size of the human exons varied widely. The longest exon was 6609 bp length and the short exon was 167 bp long. About 2000 human mRNA sequences have been determined earlier (Makałowski and Boguski, 1998; Han et al., 2020).

14. Gene clusters and families

In human genome, many genes clustered in groups of different molecular sizes based on sequence similarity. Gene families developed during the evolutionary period by duplicated genes over the periods of times and reflected in sequence similarity. The more similar genes generally shared a common ancestor, and the other one shares with a weak similarity. Gene duplication process involved various mechanisms, and these include retroposition and recombination. Some of the duplicated genes were inactive and are called pseudogenes (Cheetham et al., 2020). The histone gene family is a typical example of similar genes. The histone family consists of five typical genes, linked, although in varying arrays of different copy numbers dispersed in the human genome. The genes in the histone family (H2 genes) encode particularly identical H4 protein. Individual human genomic clones analysis revealed the presence of H4, clusters of more or two histone genes, or clusters of histone genes such as, H3-H4-H1-H3-H2A-H2B (Hentschel and Birnstiel, 1981). The histone genes form a cluster on chromosome 6 of humans and forms small cluster at chromosome 1. Histone genes lack introns and this is one of the rare features for most of the eukaryotic genes. The ribiosomal rRNA (rRNA) is coded with about 0.4% of the DNA in the whole genome of humans. The 5.8S, 28S and 18S rRNA genes are clustered and approximately 60 copies showing more than 2 million bp of DNA. These gene clusters are available on the short arms region of five acrocentric chromosomes of humans and form the nucleolar organizing regions and about 300 copies were determined. In human genome, the highly conserved amino-acid domains were determined in some genes and these showed weak similarity (Strunz et al., 2021).

Pseudogenes were formed by the results of retroposition. These pseudogens lack the flanking DNA sequences of the various functional loci and lack of introns and therefore are not products of duplication of gene. The development of these types of pseudogenes is mainly based on the activity of reverse transcriptase. In human genome repetitive sequences have been reported. These repetitive sequences may be dispersed throughout the genome or they may be in a tandem orientation. These repetitive sequences were classified based on the functions, sequence relatedness and dispersed pattern. The repetitive sequences are otherwise called as satellite DNA, with unknown function and are the products of transposable element integration, including retropseudogenes of a functional gene and retrogenes. Microsatellites are very small arrays of tandem repeats, and very small base pairs (4 bp or less). Microsatellites are dispersed throughout the genome of humans, moreover TG/CA repeats are highly common, and estimated about 0.5% in the whole genome. Generally, microsatellite have no specific gene functions, however, TG/CA dinucleotide pairs can form Z-DNA, which involved in various functions. In human genes, repeated trinucleotide within genes was associated with various genetic disorders such as, fragile-X syndrome. Telomeric DNA sequences are one of the microsatellites and the lengths ranged between 1 kbp and 15 kbp (Dabir et al., 2020).

15. DNA sequence changes and alterations in the gene expression and epigenetic markers

Epigenetics is termed as the alterations in the gene expression profile of a cell that are not mainly caused by DNA sequence changes (Peschansky and Wahlestedt, 2014). Hence, epigenetic inheritance refers to the transformation of some markers to offspring. These transformations may due to their inheritability, environmental sensitivity and stability over time. Intergenerational epigenetic inheritance reveals the transmission of epigenetic markers from one generation to the next generation from grandparents to grandchild (Pang et al., 2017). Epigenetics can be classified into direct epigenetics and indirect epigenetics. The direct epigenetics occur during an individual's lifespan. This phenomenon implies short-term regulation of gene expression, and is mediated by the action of transcription factors such as c-jun, c-fos, CREB and ZENK. These genes are also called immediate-early genes, because it's adaptive events on other genes (Johnson, 2010). These factors significantly regulate the expression of various genes that encode for many functional proteins, as well as non-coding RNAs (ncRNAs), that can play a significant role in the mediation of epigenetic processes. In recent years, the role of ncRNAs has been described and these ncRNAs involved in direct or indirect regulation of various epigenetic processes in the development of individual (Bohacek and Mansuy, 2015). Transcription factors are the examples of the epigenetic mechanism or direct epigenetics. In indirect epigenetics, the direct effect is transmitted to the offspring and some effects become an environmental trigger for the development of ontogenetic effect of the offspring. The phylogenetic adaption (environmental changes) can be classified into within indirect epigenetics and across indirect epigenetics (van Otterdijk and Michels, 2016). In the case of within indirect epigenetics, all changes related to epigentics act synchronously on the individual. The changes even taken place in the zygote and various environmental factors affect the development of individual (from zygote to gestation). In the case of indirect epigenetics asynchronous mode of development has been reported in the germ cells. It has been reported that the developed epigenetic changes transmitted

Table 1
Whole genome sequence analysis and determination of genes involved in various diseases.

Year	Milestones	References
1986	Four-colour fluorescence-based sequence detection	Smith et al., 1986
1995	Identification of the breast cancer susceptibility gene	Wooster et al., 1995
1997	Human whole-genome shotgun sequencing	Weber and Myers, 1997
1998	Shotgun sequencing of the human genome	Venter et al., 1998
1995	Breast cancer susceptibility gene BRCA2	Wooster et al., 1995
1997	Positional cloning of the APECED gene	Nagamine et al., 1997
1997	Mutations in PEX1	Reuber et al., 1997
1997	Human PEX1 gene	Portsteffen et al., 1997
1997	Pendred syndrome	Everett et al., 1997
1998	X-linked lymphoproliferative disease	Coffey et al., 1998
1998	Nonsyndromic hearing impairment	Van Laer et al., 1998
1999	Darier disease	Sakuntabhai et al., 1999
1999	X-linked spondyloepiphyseal dysplasia tarda	Gedeon et al., 1999
1999	Mutations in the CCN gene family member WISP3	Hurvitz et al., 1999

one generation to the other generations as epimutations, than other genetic mutations (Johnson, 2010).

16. Genetics and biology of indirect epigenetics

The experimental evidences on epigenetic changes have been described previously. Epigenetic modifications can be generally triggered by various non-genetic factors such as, smoking, pollution, diet (Mathers et al., 2010) and these are termed as "stressors". These epigenetic changes involved in the development of various disease, including, cancer, cardiovascular disease, metabolic disease, autoimmune disease and neurodegenerative disease and most of them related to aging. Many research groups have been reported the functions of epigenetic mechanisms in mediating the risk and development of psychopathologies, mainly depression and anxiety (Maccari et al., 2017). Adverse early experiences such as abuse or low maternal care can affect the expression of psychopathologies glucocorticoid receptor gene, which played critical role in depression and anxiety (Smart et al., 2015).

In recent years, the expression levels of microRNAs have been analyzed in relation with psycho- and neuropathologies such as, Huntington's disease, Parkinson's disease, Alzheimer's disease, schizophrenia, multiple sclerosis, autism, addiction, depression, anxiety, and bipolar disorder. Stress is one of the factors studied extensively trigger alterations to miRNAs, mainly in the brain region (Hollins and Cairns, 2016). In recent years, the advances on indirect epigenetics have been described. During pregnancy, various factors affecting offspring epigenetically and this process is termed as fetal programming (Maccari et al., 2017). Swanson et al. (2009) reported the role of maternal and intrauterine environment that shape the developing offspring functionally and structurally. These processes can affect, brain development, leading to various neuropsychiatric disorders, such as, Alzheimer's disease, Parkinson's disease, schizophrenia, attention-deficit hyperactivity disorder, anxiety and depressive disorder (Faa et al., 2016). Maternal epigenetic factors are highly active at the time of gestation and interfere with the development of neurons have been classified into two groups, maternal and fetal. Maternal hormones, nutrient, odours, nutrients and immune factors can also influence during gestation was also considered as maternal factors (Todrank et al., 2011).

17. Whole genome sequencing and determination of genetic diseases

WGS is useful for the analysis of genetic diseases (Gibbs, 2020). Whole genome sequencing as well as whole exome sequencing is used for the determination of genetic diseases. WGS enables to determine single-base variation in the genome of interest. In recent years, whole genomes and exomes for both diseased and health individuals have been deposited and this information is used for personalized health monitoring and development of preventative medicine (Snyder et al., 2010). Whole genome of ovarian cancer, breast cancer, melanoma, hepatocellular carcinoma, small-cell lung cancer, pediatric glioblastoma (Schwartzentruber et al., 2012), etc were characterized. The WGS of the cancer genomes are useful for the development of personalized cancer treatment. WGS also used detect spontaneous mutations in the normal genome of patients that may lead to cancer disease. WGS also applied to analyze the normal genes for other diseases and variations at the personalized level (Baranzini et al., 2010) (Table 1).

18. Personalized disease risk assessment and health monitoring with integrative approach

Genome sequences are useful to provide valuable information on the status of the genetic diseases and drug response efficacy. The whole genome of a patient with a family history of vascular disease and very early death but no critical clinical record, and determined elevated post-test probability risks of coronary artery disease and myocardial infarction (Pushkarev et al., 2009). However, whole genome sequence may not always be sufficient to determine the health of the patients, because other environmental factors also trigger diseases (Su et al., 2012). In a study Roberts et al. (2012) determined the risk of about 24 diseases after whole genome sequencing in the monozygotic twins. In addition to the whole genome sequencing approach transcriptome, epigenome, metabolome and proteome of the human body and omics approach such as microRNA and gut microbiome and various immune receptor may also be significant factor for the monitoring of health and personalized medicine, either in combination with other iPOP omics or alone. MicroRNA is one of the important markers for the development of personalized medicine and it controls various tumour initiation drivers (Sumazin et al., 2011). Omics approach also used to predict highly complex diseases such as, asthma, immune response and inflammation (Peretz et al., 2012). The WES was used to determine complex disorder, Mendelian phenotypes and Miller Syndrome and this method is useful for various clinical applications because this area covers only functional genome to analyze the variations in the exon region and to determine the disease-causing mutations (Weigelt et al., 2018).

19. Conclusions and future perspectives

In this review article, I analyzed the biological basis of ABO blood group, biology of important organs including endoplasmic

reticulum, and nucleus and associated diseases. Whole genome sequencing data has several advantages and variation of expressed genes have biomedical and therapeutic potential. Recent findings revealed reversible changes of chromosome structure and unprecedented changes taken place in the interphase-mitotic cell division. In human genome more than 50% of a genome was arranged in repetitive manner, 2% of genome was coded and the remaining 48% were unique, originated through mobile elements. However, whole genome sequence may not always be sufficient to determine the health of the patients, because other environmental factors also trigger diseases. In humans, alternations in gene expression were determined without any changes in DNA sequences.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Amundadottir, L., Kraft, P., Stolzenberg-Solomon, R.Z., Fuchs, C.S., Petersen, G.M., Arslan, A.A., Bueno-de-Mesquita, H.B., Gross, M., Helzlsouer, K., Jacobs, E.J., LaCroix, A., Zheng, W., Albanes, D., Bamlet, W., Berg, C.D., Berrino, F., Bingham, S., Buring, J.E., Bracci, P.M., Canzian, F., Clavel-Chapelon, F., Clipp, S., Cotterchio, M., de Andrade, M., Duell, E.J., Fox Jr, J.W., Gallinger, S., Gaziano, J.M., Giovannucci, E.L., Goggins, M., González, C.A., Hallmans, G., Hankinson, S.E., Hassan, M., Holly, E.A., Hunter, D.J., Hutchinson, A., Jackson, R., Jacobs, K.B., Jenab, M., Kaaks, R., Klein, A.P., Kooperberg, C., Kurtz, R.C., Li, D., Lynch, S.M., Mandelson, M., McWilliams, R.R., Mendelsohn, J.B., Michaud, D.S., Olson, S.H., Overvad, K., Patel, A.V., Peeters, P.H.M., Rajkovic, A., Riboli, E., Risch, H.A., Shu, X.-O., Thomas, G., Tobias, G.S., Trichopoulos, D., Van Den Eeden, S.K., Virtamo, J., Wactawski-Wende, J., Wolpin, B.M., Yu, H., Yu, K., Zeleniuch-Jacquotte, A., Chanock, S.J., Hartge, P., Hoover, R.N., 2009. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. Nat. Genet. 41 (9), 986–990.

André, T., Kotelevets, L., Vaillant, J.C., Coudray, A.M., Weber, L., Prévot, S., Parc, R., Gespach, C., Chastre, E., 2000. Vegf, Vegf-B, Vegf-C and their receptors KDR, FLT-1 and FLT-4 during the neoplastic progression of human colonic mucosa. Int. J. Cancer. 86 (2), 174–181.

Baranzini, S.E., Mudge, J., Van Velkinburgh, J.C., Khankhanian, P., Khrebtukova, I., Miller, N.A., Zhang, L., Farmer, A.D., Bell, C.J., Kim, R.W., May, G.D., 2010. Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. Nature 464 (7293), 1351–1356.

Bernardi, G., 2000. Isochores and the evolutionary genomics of vertebrates. Gene 241 (1), 3–17.

Bohacek, J., Mansuy, I.M., 2015. Molecular insights into transgenerational non-genetic inheritance of acquired behavior. Nat. Rev. Genet. 16, 641–652.

Braakman, I., Hebert, D.N., 2013. Protein folding in the endoplasmic reticulum. Cold Spring Harb. Perspect. Biol. 5(5), a013201.

Brosius, U., Gärtner, J., 2002. Human Genome and Diseases: Cellular and molecular aspects of Zellweger syndrome and other peroxisome biogenesis disorders. Cell. Mol. Life Sci. 59 (6), 1058–1069.

Brown, A., Geiger, H., 2018. Chromosome integrity checkpoints in stem and progenitor cells: transitions upon differentiation, pathogenesis, and aging. Cell. Mol. Life Sci. 75 (20), 3771–3779.

Cheetham, S.W., Faulkner, G.J., Dinger, M.E., 2020. Overcoming challenges and dogmas to understand the functions of pseudogenes. Nat. Rew. Genetics 21 (3), 191–201.

Clausen, H., Bennett, E.P., Grunnet, N., 1994. Molecular genetics of ABO histo-blood groups. Transfus. Clin. Biol. 1 (2), 79–89.

Coffey, A.J., Brooksbank, R.A., Brandau, O., Oohashi, T., Howell, G.R., Bye, J.M., Cahn, A.P., Durham, J., Heath, P., Wray, P., Pavitt, R., Wilkinson, J., Leversha, M., Huckle, E., Shaw-Smith, C.J., Dunham, A., Rhodes, S., Schuster, V., Porta, G., Yin, L., Serafini, P., Sylla, B., Zollo, M., Franco, B., Bolino, A., Seri, M., Lanyi, A., Davis, J.R., Webster, D., Harris, A., Lenoir, G., de St Basile, G., Jones, A., Behloradsky, B.H., Achatz, H., Murken, J., Fassler, R., Sumegi, J., Romeo, G., Vaudin, M., Ross, M.T., Meindl, A., Bentley, D.R., 1998. Host response to EBV infection in X-linked lymphoproliferative disease results from mutations in an SH2-domain encoding gene. Nat. Genet. 20 (2), 129–135.

Dabir, P.D., Bruggeling, C.E., van der Post, R.S., Dutilh, B.E., Hoogerbrugge, N., Ligtenberg, M.J., Boleij, A., Nagtegaal, I.D., 2020. Microsatellite instability screening in colorectal adenomas to detect Lynch syndrome patients? A systematic review and meta-analysis. Eur. J. Human. Genetics 28 (3), 277–286.

Denomme, G.A., 2011. Molecular basis of blood group expression. Transfus. Aphere. Sci. 44 (1), 53–63.

Epstein, C.J., Martin, G.M., Schultz, A.L., Motulsky, A.G., 1985. Werner's syndrome: A review of its symptomatology, natural history, pathologic features, genetics and relationship to the natural aging process. Werner's Syn. Human Aging., 57–120

Everett, L.A., Glaser, B., Beck, J.C., Idol, J.R., Buchs, A., Heyman, M., Adawi, F., Hazani, E., Nassir, E., Baxevanis, A.D., Sheffield, V.C., Green, E.D., 1997. Pendred syndrome is caused by mutations in a putative sulphate transporter gene (PDS). Nat. Genet. 17 (4), 411–422.

Ewald, D.R., Sumner, S.C., 2016. Blood type biochemistry and human disease. Wiley Interdisciplin. Rew. Sys. Biol. Med. 8 (6), 517–535.

Faa, G., Manchia, M., Pintus, R., Gerosa, C., Marcialis, M.A., Fanos, V., 2016. Fetal programming of neuropsychiatric disorders. Birth. Defect. Res. C Embryo Today 108 (3), 207–223.

Filocamo, M., Morrone, A., 2011. Lysosomal storage disorders: molecular basis and laboratory testing. Human Genom. 5 (3), 1–14.

Fokkema, I.F., Taschner, P.E., Schaafsma, G.C., Celli, J., Laros, J.F., den Dunnen, J.T., 2011. LOVD v. 2.0: the next generation in gene variant databases. Hum. Mutat. 32 (5), 557–563.

Forsberg, K., Andersen, P.M., Marklund, S.L., Brännström, T., 2011. Glial nuclear aggregates of superoxide dismutase-1 are regularly present in patients with amyotrophic lateral sclerosis. Acta Neuropathol. 121 (5), 623–634.

Franchini, M., Bonfanti, C., 2015. Evolutionary aspects of ABO blood group in humans. Clin. Chim. Acta 444, 66–71.

Gabriel, S.E., Clarke, L.L., Boucher, R.C., Stutts, M.J., 1993. CFTR and outward rectifying chloride channels are distinct proteins with a regulatory relationship. Nature 363 (6426), 263–266.

Garg, A., Ahmed, S., Sinha, A., Singh, H.P., 2015. Tumor markers–Its advantages and limitations in diagnosis of oral cancer. Univ. J. Dent. Sci. 1, 42–45.

Gedeon, Á.K., Colley, A., Jamieson, R., Thompson, E.M., Rogers, J., Sillence, D., Tiller, G.E., Mulley, J.C., Gécz, J., 1999. Identification of the gene (SEDL) causing X-linked spondyloepiphyseal dysplasia tarda. Nature Genet. 22 (4), 400–404.

Gibbs, R.A., 2020. The human genome project changed everything. Nat. Rev. Genet. 21 (10), 575–576.

Goto, M., Miller, R.W., Ishikawa, Y., Sugano, H., 1996. Excess of rare cancers in Werner syndrome (adult progeria). Cancer Epidemiol. Biomarkers Prev. 5, 239–246.

Han, X., Zhou, Z., Fei, L., Sun, H., Wang, R., Chen, Y., Chen, H., Wang, J., Tang, H., Ge, W., Zhou, Y., Ye, F., Jiang, M., Wu, J., Xiao, Y., Jia, X., Zhang, T., Ma, X., Zhang, Q.i., Bai, X., Lai, S., Yu, C., Zhu, L., Lin, R., Gao, Y., Wang, M., Wu, Y., Zhang, J., Zhan, R., Zhu, S., Hu, H., Wang, C., Chen, M., Huang, H.e., Liang, T., Chen, J., Wang, W., Zhang, D., Guo, G., 2020. Construction of a human cell landscape at single-cell level. Nature 581 (7808), 303–309.

Harris, L., Fritsche, H., Mennel, R., Norton, L., Ravdin, P., Taube, S., Somerfield, M.R., Hayes, D.F., Bast, R.C., 2007. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. J. Clin. Oncol. 25 (33), 5287–5312.

Hashimoto, I., Kodama, J., Seki, N., Hongo, A., Yoshinouchi, M., Okuda, H., Kudo, T., 2001. Vascular endothelial growth factor-C expression and its relationship to pelvic lymph node status in invasive cervical cancer. Br. J. Cancer. 85 (1), 93–97.

Hentschel, C.C., Birnstiel, M.L., 1981. The organization and expression of histone gene families. Cell 25 (2), 301–313.

Hirai, M., Nakagawara, A., Oosaki, T., Hayashi, Y., Hirono, M., Yoshihara, T., 2001. Expression of vascular endothelial growth factors (VEGF-A/VEGF-1 and VEGF-C/VEGF-2) in postmenopausal uterine endometrial carcinoma. Gynecol. Oncol. 80 (2), 181–188.

Hollins, S.L., Cairns, M.J., 2016. MicroRNA: small RNA mediators of the brains genomic response to environmental stress. Prog. Neurobiol. 143, 61–81.

Hurvitz, J.R., Suwairi, W.M., Van Hul, W., El-Shanti, H., Superti-Furga, A., Roudier, J., Holderbaum, D., Pauli, R.M., Herd, J.K., Van Hul, E., Rezai-Delui, H., 1999. Mutations in the CCN gene family member WISP3 cause progressive pseudorheumatoid dysplasia. Nat. Genet. 23 (1), 94–98.

Jan, C.H., Williams, C.C., Weissman, J.S., 2014. Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling. Science 346 (6210), 1257521.

Johnson, P.J., 2001. A framework for the molecular classification of circulating tumor markers. Ann. NY. Aca. Sci. 945 (1), 8–21.

Johnson, T.D., 2010. "Developmental system theory," in Oxford Handbook of Developmental Behavioral Neuroscience, ed. M. S. Blumberg, J. H. Freeman and S. R. Robinson (New York, NY: Oxford University Press), 12–29.

Jung, M., Pfeifer, G.P., 2015. Aging and DNA methylation. BMC Biol. 13, 7.

Karl Landsteiner-Biographical. NobelPrize.org. Nobel Media AB, 2020. Available online: https://www.nobelprize.org/prizes/medicine/1930/landsteiner/biographical/ (accessed on 8 March 2021).

Kazazi-Hyseni, F., Beijnen, J.H., Schellens, J.H., Bevacizumab, 2010. Oncologist 15 (8), 819–825.

Kerem, B.S., Rommens, J.M., Buchanan, J.A., Markiewicz, D., Cox, T.K., Chakravarti, A., Buchwald, M., Tsui, L.C., 1989. Identification of the cystic fibrosis gene: genetic analysis. Science 245 (4922), 1073–1080.

Kumaresan, V., Pasupuleti, M., Arasu, M.V., Al-Dhabi, N.A., Arshad, A., Nurul Amin, S. M., Md Yusoff, F., Arockiaraj, J., 2019. A comparative transcriptome approach for identification of molecular changes in Aphanomyces invadans infected Channa striatus. Mol. Biol. Rep. 45, 2511–2523.

Lekkerkerker, H.N.W., Tuinier, R., 2011. Colloids and the Depletion Interaction. Lecture Notes in Physics 233, 833.

Lucaccioni, L., Righi, B., Cingolani, G.M., Lugli, L., Della Casa, E., Torcetta, F., Iughetti, L., Berardi, A., 2020. Overwhelming sepsis in a neonate affected by Zellweger syndrome due to a compound heterozygosis in PEX 6 gene: a case report. BMC Med. Genet. 21 (1), 1–4.
Maccari, S., Polese, D., Reynaert, M., Amici, T., Morley-Fletcher, S., Fagioli, F., 2017. Early-life experiences and the development of adult diseases with a focus on mental illness: the human Birth theory. Neuroscience 342, 232–251.
Makałowski, W., Boguski, M.S., 1998. Evolutionary parameters of the transcribed mammalian genome: an analysis of 2,820 orthologous rodent and human sequences. Proc. Natl. Acad. Sci. U.S.A. 95 (9407), 9412.
Mäkivuokko, H., Lahtinen, S.J., Wacklin, P., Tuovinen, E., Tenkanen, H., Nikkilä, J., Björklund, M., Aranko, K., Ouwehand, A.C., Mättö, J., 2012. Association between the ABO blood group and the human intestinal microbiota composition. BMC Microbiol. 12 (1), 1–12.
Mathers, J.C., Strathdee, G., Relton, C.L., 2010. Induction of epigenetic alterations by dietary and other environmental factors. Adv. Genet. 71, 3–39.
Miki, T., Nakura, J., Ye, L., Mitsuda, N., Morishima, A., Sato, N., Kamino, K., Ogihara, T., 1997. Molecular and epidemiological studies of Werner syndrome in the Japanese population. Mech. Ageing Dev. 98, 255–265.
Nagamine, K., Peterson, P., Scott, H.S., Kudoh, J., Minoshima, S., Heino, M., Krohn, K.J., Lalioti, M.D., Mullis, P.E., Antonarakis, S.E., Kawasaki, K., 1997. Positional cloning of the APECED gene. Nat. Genet. 17 (4), 393–398.
Nagpal, M., Singh, S., Singh, P., Chauhan, P., Zaidi, M.A., 2016. Tumor markers: A diagnostic tool. Nat. J. Maxillofacial Sur. 7 (1), 17–20.
Nicholson, J.M., Cimini, D., 2013. Cancer karyotypes: survival of the fittest. Front. Oncol. 3, 148.
O-charoenrat, P., Rhys-Evans, P., Eccles, S.A., 2001. Expression of vascular endothelial growth factor family members in head and neck squamous cell carcinoma correlates with lymph node metastasis. Cancer: Interdisciplin. Int. J. Am. Cancer Soc. 92 (3), 556–568.
Pang, T.Y., Short, A.K., Bredy, T.W., Hannan, A.J., 2017. Transgenerational paternal transmission of acquired traits: stress-induced modification of the sperm regulatory transcriptome and offspring phenotypes. Curr. Opin. Behav. Sci. 14, 140–147.
Peretz, Y., Cameron, C., Sékaly, R.P., 2012. Dissecting the HIV-specific immune response: a systems biology approach. Curr. Opin. HIV AIDS. 7 (1), 17–23.
Peschansky, V.J., Wahlestedt, C., 2014. Non-coding RNAs as direct and indirect modulators of epigenetic regulation. Epigenetics 9 (1), 3–12.
Politei, J.M., Cabello, J.F., Villalobos, J., Valadez, G., Loaeza, A., Linares, A., Martins, A. M., 2009. Enfermedad de Fabry: Nuevos conceptos en su historia natural, evolución y tratamiento, en relación a los hallazgos del Registro Fabry. Revista de nefrología, diálisis y transplante 29 (4), 145–152.
Portsteffen, H., Beyer, A., Becker, E., Epplen, C., Pawlak, A., Kunau, W.H., Dodt, G., 1997. Human PEX1 is mutated in complementation group 1 of the peroxisome biogenesis disorders. Nat. Genet. 17 (4), 449–452.
Pushkarev, D., Neff, N.F., Quake, S.R., 2009. Single-molecule sequencing of an individual human genome. Nat. Biotechnol. 27 (9), 847–850.
Ravichandran, G., Kumaresan, V., Bhatt, P., Arasu, M.V., Al-Dhabi, N.A., Arockiaraj, J., 2017. A cumulative strategy to predict and characterize antimicrobial peptides (AMPs) from protein database. Int. J. Pept. Res. Therapeut. 23, 281–290.
Ravichandran, G., Kumaresan, V., Mahesh, A., Dhayalan, A., Arshad, A., Arasu, M.V., Al-Dhabi, N.A., Pasupuleti, M., Arockiaraj, J., 2018. Bactericidal and fungistatic activity of peptide derived from GH18 domain of prawn chitinase 3 and its immunological functions during biological stress. Int. J. Biol. Macromol. 106, 1014–1022.
Reuber, B.E., Germain-Lee, E., Collins, C.S., Morrell, J.C., Ameritunga, R., Moser, H.W., Valle, D., Gould, S.J., 1997. Mutations in PEX1 are the most common cause of peroxisome biogenesis disorders. Nat. Genet. 17 (4), 445–448.
Roberts, N.J., Vogelstein, J.T., Parmigiani, G., Kinzler, K.W., Vogelstein, B., Velculescu, V.E., 2012. The predictive capacity of personal genome sequencing. Sci. Translat. Med. 4(133), 133ra58-133ra58.
Rossi, M.L., Ghosh, A.K., Bohr, V.A., 2010. Roles of Werner syndrome protein in protection of genome integrity. DNA Repair 9 (3), 331–344.
Sakuntabhai, A., Ruiz-Perez, V., Carter, S., Jacobsen, N., Burge, S., Monk, S., Smith, M., Munro, C.S., O'Donovan, M., Craddock, N., Kucherlapati, R., 1999. Mutations in ATP2A2, encoding a Ca 2+ pump, cause Darier disease. Nat. Genet. 21 (3), 271–277.
Samtleben, S., Jaepel, J., Fecher, C., Andreska, T., Rehberg, M., Blum, R., 2013. Direct imaging of ER calcium with targeted-esterase induced dye loading (TED). J. Vis. Exp. 75, e50317.
Sathyamoorthi, A., Bhatt, P., Ravichandran, G., Kumaresan, V., Arasu, M.V., Al-Dhabi, N.A., Arockiaraj, J., 2017. Gene expression and in silico analysis of snakehead murrel interleukin 8 and antimicrobial activity of C-terminal derived peptide WS12. Veterin. Immunol. Immunopathol. 190, 1–9.
Sathyamoorthi, A., Kumaresan, V., Palanisamy, R., Pasupuleti, M., Arasu, M.V., Al-Dhabi, N.A., Marimuthu, K., Nurul Amin, S.M., Arshad, A., Md Yusoff, F., Arockiaraj, J., 2019. Therapeutic Cationic Antimicrobial Peptide (CAP) Derived from Fish Aspartic Proteinase Cathepsin D and its Antimicrobial Mechanism. Int. J. Pept. Res. Therapeut. 25, 93–105.
Sathyamoorthi, A., Palanisamy, R., Arasu, M.V., Al-Dhabi, N.A., Pasupuleti, M., Arockiaraj, J., 2018. Fish Heat Shock Cognate 70 Derived AMPs CsHSC70 A1 and CsHSC70 A2. Int. J. Peptide Res. Therapeut. 24, 143–155.
Schunkert, H., König, I.R., Kathiresan, S., Reilly, M.P., Assimes, T.L., Holm, H., Preuss, M., Stewart, A.F., Barbalic, M., Gieger, C., Absher, D., 2011. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat. Gene. 43 (4), 333–338.
Schvartzman, J.M., Sotillo, R., Benezra, R., 2010. Mitotic chromosomal instability and cancer: mouse modelling of the human disease. Nat. Rev. Cancer. 10 (2), 102–115.
Schwartzentruber, J., Korshunov, A., Liu, X.Y., Jones, D.T., Pfaff, E., Jacob, K., Sturm, D., Fontebasso, A.M., Quang, D.A.K., Tönjes, M., Hovestadt, V., 2012. Driver mutations in histone H3. 3 and chromatin remodelling genes in paediatric glioblastoma. Nature 482 (7384), 226–231.
Shen, J.C., Loeb, L.A., 2000. The Werner syndrome gene: the molecular basis of RecQ helicase-deficiency diseases. Trends Genet. 16, 213–220.
Shoshani, T., Augarten, A., Gazit, E., Bashan, N., Yahav, Y., Rivlin, Y., Tal, A., Seret, H., Yaar, L., Kerem, E., Kerem, B.S., 1992. Association of a nonsense mutation (W1282X), the most common mutation in the Ashkenazi Jewish cystic fibrosis patients in Israel, with presentation of severe disease. Am. J. Hum. Genet. 50 (1), 222.
Sinclair, P., Bian, Q., Plutz, M., Heard, E., Belmont, A.S., 2010. Dynamic plasticity of large-scale chromatin structure revealed by self-assembly of engineered chromosome regions. J. Cell Biol. 190 (5), 761–776.
Smart, C., Strathdee, G., Watson, S., Murgatroyd, C., McAllisterWilliams, R.H., 2015. Early life trauma, depression and the glucocorticoid receptor gene—an epigenetic perspective. Psychol. Med. 45, 3393–3410.
Smith, L.M., Sanders, J.Z., Kaiser, R.J., Hughes, P., Dodd, C., Connell, C.R., Heiner, C., Kent, S.B., Hood, L.E., 1986. Fluorescence detection in automated DNA sequence analysis. Nature 321 (6071), 674–679.
Snyder, M., Du, J., Gerstein, M., 2010. Personal genome sequencing: current approaches and challenges. Genes Develop. 24 (5), 423–431.
Strunz, T., Kellner, M., Kiel, C., Weber, B.H., 2021. Assigning Co-Regulated Human Genes and Regulatory Gene Clusters. Cells 10 (9), 2395.
Su, M.W., Tung, K.Y., Liang, P.H., Tsai, C.H., Kuo, N.W., Lee, Y.L., 2012. Gene-gene and gene-environmental interactions of childhood asthma: a multifactor dimension reduction approach. PloS One 7, (2) e30694.
Sumazin, P., Yang, X., Chiu, H.S., Chung, W.J., Iyer, A., Llobet-Navas, D., Rajbhandari, P., Bansal, M., Guarnieri, P., Silva, J., Califano, A., 2011. An extensive microRNA-mediated network of RNA-RNA interactions regulates established oncogenic pathways in glioblastoma. Cell 147 (2), 370–381.
Swanson, J.M., Entringer, S., Buss, C., Wadhwa, P.D., 2009. Developmental origins of health and disease: environmental exposures. Semin. Reprod. Med. 27, 391–402.
Thompson, S.L., Bakhoum, S.F., Compton, D.A., 2010. Mechanisms of chromosomal instability. Curr. Biol. 20 (6), R285–R295.
Todrank, J., Heth, G., Restrepo, D., 2011. Effects of in utero odorant exposure on neuroanatomical development of the olfactory bulb and odour preferences. Proc. Biol. Sci. 278, 1949–1955.
Van Laer, L., Huizing, E.H., Verstreken, M., van Zuijlen, D., Wauters, J.G., Bossuyt, P.J., Van de Heyning, P., McGuirt, W.T., Smith, R.J., Willems, P.J., Legan, P.K., 1998. Nonsyndromic hearing impairment is associated with a mutation in DFNA5. Nat. Genet. 20 (2), 194–197.
van Otterdijk, S., Michels, B.K., 2016. Transgenerational epigenetic inheritance in mammals: how good is the evidence? FASEB J. 30, 2457–2465.
Velde, C.V., Miller, T.M., Cashman, N.R., Cleveland, D.W., 2008. Selective association of misfolded ALS-linked mutant SOD1 with the cytoplasmic face of mitochondria. Proc. Natl. Acad. Sci. U.S.A. 105 (10), 4022–4027.
Venter, J.C., Adams, M.D., Sutton, G.G., Kerlavage, A.R., Smith, H.O., Hunkapiller, M., 1998. Shotgun sequencing of the human genome. Science 280 (5369), 1540–1542.
Wang, J., Shiels, C., Sasieni, P., Wu, P.J., Islam, S.A., Freemont, P.S., Sheer, D., 2004. Promyelocytic leukemia nuclear bodies associate with transcriptionally active genomic regions. J. Cell Biol. 164, 515–526.
Weigelt, B., Bi, R., Kumar, R., Blecua, P., Mandelker, D.L., Geyer, F.C., et al., 2018. The landscape of somatic genetic alterations in breast cancers from ATM germline mutation carriers. J. Natl. Cancer Inst. 110, 1030–1034.
Westrate, L.M., Lee, J.E., Prinz, W.A., Voeltz, G.K., 2015. Form follows function: the importance of endoplasmic reticulum shape. Annu. Rev. Biochem. 84, 791–811.
Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C., Micklem, G., Barfoot, R., 1995. Identification of the breast cancer susceptibility gene BRCA2. Nature 378 (6559), 789–792.
Yamamoto, F., Cid, E., Yamamoto, M., Blancher, A., 2012. ABO research in the modern era of genomics. Transfus. Med. Rev. 26 (2), 103–118.
Zhang, M.Q., 1998. Statistical features of human exons and their flanking regions. Hum. Mol. Genet. 7 (5), 919–932.