

Comparative Genomics A New Arena of Biological Research: A Review

Anjali Tripathi*

Department of Biotechnology, Shri Ramswaroop Memorial University, Lucknow, India

*Corresponding author: Tripathi Anjali, Department of Biotechnology, Shri Ramswaroop Memorial University, Lucknow, India, Tel: +91-7985981955; E-mail: tripathianjali046@gmail.com

Received date: Oct 24, 2018; Accepted date: Jan 22, 2019; Published date: Jan 31, 2019

Copyright: © 2019 Anjali T. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Citation: Anjali T (2019) Comparative genomics a new arena of biological research: A review. Genet Mol Biol Res Vol No: 3 Iss No: 1:10

Abstract

Genomics is one of the profligate developing disciplines of science; where the innovation was the first complete genome sequencing of *Haemophilus influenzae* in 1995. The explosion of sequenced genomes has allowed the appraisal of the character of natural selection at that level of organization. A simple evaluation of the common features of genomes such as genome extent, total of genes, and chromosome quantity presents an entry point into comparative genomic analysis. This comprises of defining orthologous sections of DNA that descend from the same area in the common predecessor of the species compared, and paralogous sections that ascended by replication events preceding to the separation of the species compared. Matrix Assisted Laser Desorption Ionization Time Of Flight mass spectrometry (MALDI-TOF) is conventional in mass spectrometry in general as well as protein breakdown in particular. The CRISPR (Clustered, Regularly Interspaced, Short Palindromic Repeats)/CRISPR-associated protein 9 (Cas9) systems offers a speedy as well as proficient technology for targeted genome editing. In association with a precisely intended guide RNA (gRNA), Cas9 can accomplish site-specific DNA recognition in addition to cleavage

which used precisely, for discovery of contamination, identification of strain or intra precise alternative, determination of drug resistance and quantification of parasitic weight/DNA [5].

Comparative genomics is an arena of genetic study in which the genome arrangements of dissimilar species-mouse, human, and an extensive diversity of other organisms from bacteria to chimps are compared. A simple evaluation of the common features of genomes such as genome extent, total of genes, and chromosome quantity presents an entry point into comparative genomic analysis [6,7]. At the present time, it is likely to detect precise genetic factor accountable of biological variation. It is a wonderful chance to redefine the natural science, diversity and host-specificity. Fortuitously, there is an enormous quantity of genomic data waiting to be discovered to expose new genetic material accountable of adaptive and varying process, as well as reductive growth [8]. Genome biology goals at spending the whole genome arrangements to modernize all metabolic as well as signaling pathways that could activate in the goal organisms and recognize the likely governing centres and prospective drug targets. Such enquiry needs inclusive useful footnote of all proteins programmed in each sequenced genome [9]. Comparative genomics was necessary in aligning arrangements from divergent species for advance enquiry. Expanses of those composites that possibly interrelate near the DNA binding domain were segmented into "interrelating or expected co-operating areas" and "non-interacting areas". These areas were then imperilled to our fingerprinting procedure, which figures a linear "information signature" trusting on fragment geometric as well as physicochemical properties to examine a large pregenerated complex library shown in **Table 1** (20 million compounds) [10].

Introduction

In former era most of the biological study orbited around genomics and molecular biology interrelated studies. In genomic studies a foremost area of learning was engrossed on comparative genomics as well as genome sequencing [1]. Genomics is one of the profligate developing disciplines of science, where the innovation was the first complete genome sequencing of *Haemophilus influenzae* in 1995 [2].

Genome-wide studies are progressively becoming a must, specifically for intricate diseases such as cancer where numerous genes and varied molecular mechanisms are recognized to be involved in genes' function modification [3]. The explosion of sequenced genomes has allowed the appraisal of the character of natural selection at that level of organization [4]. The first methodology in the analysis of the parasitic infections even in tiny amounts of target template is the PCR as well as sequencing

Table 1 Comparison of the sizes of eight complete eukaryotae genome sequences and examples of complete bacteria and archaea genomes and eukaryotae draft genomes [11].

Domain	Organism	Genome size (kbp)
Archaea	<i>Thermoplasma acidophilum</i>	1565
	<i>Archaeoglobus fulgidus</i>	2178
	<i>Sulfolobus solfataricus</i>	2992

	<i>Methanosarcina acetivorans str. C2A</i>	5751
Bacteria	<i>Salmonella typhi</i>	180
	<i>Helicobacter pylori</i> 26695	1668
	<i>Haemophilus influenzae</i> Rd	1830
	<i>Escherichia coli</i> K12	4639
Eucaryot a	<i>Guillardia theta nucleomorph</i>	551
	<i>Encephalitozoon cuniculi</i>	2500
	<i>Saccharomyces cerevisiae</i> S288C	12,069
	<i>Caenorhabditis elegans</i>	97,000
	<i>Arabidopsis thaliana</i>	1,15,400
	<i>Drosophila melanogaster</i>	1,37,000
	<i>Oryza sativa</i> L. ssp. indica (draft)	4,20,000
	<i>Homo sapiens</i> (draft)	30,00,000

Tools and softwares for comparative genomics

MetNet: A great device being developed, which highlights propelled representation and measurable investigation devices for the examination of post-genomic informational indexes acquired with *A. thaliana*, is the MetNet bundle (<http://www.public.iastate.edu/~mash/MetNet/exchange.html>). It coordinates factual and grouping bundles and will in the long run incorporate abilities to show metabolic and administrative systems. MetNet has a JAVA-based interface to a database (MetNetDB) that contains data on known connections in metabolic and administrative systems [12].

To confirm GeneChip® exhibit articulation information, qPCR (Quantitative real-time PCR) was performed on *T. caerulescens* and *T. arvense* qualities for which coding successions were accessible in GenBank. As an outcome of the accessibility of seed resources, distinctive *Thlaspi* populaces were utilized for the qPCR affirmation. Seeds of *T. caerulescens* ('Ganges' population, France) and *T. arvense* (gathered from Wharf Ground field, Wellesbourne, Warwickshire, UK) were sterilized, soaked up and sown in agar as portrayed previously, using a 10% basal salt plan to decrease the ambient external Zn focus ([Zn]ext [13].

Non-coding RNA (ncRNA) qualities deliver an utilitarian RNA item rather than an interpreted protein. These items are parts of the absolute most critical cell machines, for example, the ribosome (ribosomal RNAs), the spliceosome (U1, U2, U4, U5 and U6 RNAs) and the (telomerase RNA). The known collection of ncRNA cell capacities is growing quickly [14]. An expansive division of eukaryotic genomes comprises of DNA that isn't converted into protein arrangement, and little is thought about its useful essentialness. [15].

Adjusting of Noncoding Regions to assess the level of protection of noncoding areas amongst mouse and human qualities, we chose to build up another arrangement calculation that would discover saved collinear obstructs in two DNA groupings. Examinations of practically moderated locales of DNA

grouping, for example, potential transcriptional administrative areas, posture specific issues for standard calculations [16].

Bioinformatics programming can be utilized to assess grouping likeness among various nucleotides and amino acids, arrangement transformation, cancellations and inclusions, succession recombination, and hereditary advancement of infections, microorganisms, and different species [17]. Succession likeness, grouping transformation, cancellation, or addition, and arrangement recombination among reference Boca viruses were acquired utilizing DNA Star programming, and its hereditary advancement was resolved utilizing MEGA 5.1 programming [18]. Grouping comparability profiles between two genomes are intricate and hard to picture. By gathering contiguous districts of likeness into bigger syntenic obstructs, the information can be refined into a visual shape that is both rational and interpretable [19].

Current genome perception and information investigation strategies are attempting to keep up as it turns into a normal prerequisite for researcher to contrast another genome with scores, if not hundreds, of different genomes on the double [20]. BRIG is fit for producing round examination pictures for prokaryote genomes, demonstrating numerous genome correlations in a solitary picture, and showing likeness between a reference genome in the inside against other inquiry groupings as an arrangement of concentric rings hued as indicated by BLAST (Basic Local Alignment Search Tool) character [21].

In the current version, data has been curated from various primary databases such as Orphanet [2], OMIM (Online Mendelian Inheritance in Man) [3], Ensembl [4], Drugbank [5], GHR (Growth Hormone Receptor) [6], dbSNP (Single Nucleotide Polymorphism Database) [7] and Rare diseases India [8]. Also, Orthologous genes and GO terms were analyzed for different rare diseases, which are included in Rare DDB. This study has made an attempt to gather the information from literatures, various databases which are shown in **Table 2** [22].

VISTA first VISTA server at <http://www-gsd.lbl.gov/vista/> was propelled in the mid-year of 2000 and was intended to adjust long genomic arrangements and envision these arrangements with related practical comments. Presently the VISTA site incorporates numerous comparative genomics instruments and furnishes clients with rich capacities to peruse pre-figured entire genome arrangements of huge vertebrate genomes and different gatherings of living beings with VISTA Browsers [23].

ENSEMBL Ensembl maps ESTs to the genome utilizing a mix of Exonerate, BLAST and EST2Genome. These are then handled by consolidating the repetitive ESTs and setting graft locales to the most widely recognized finishes. This technique finds the right inside graft destinations bunches 5' and 3' ESTs into UTRs and joins the sections into longer transcripts structures. The subsequent transcripts are prepared by Genome wise, which finds the longest ORF over every one [24].

KEGG (Kyoto Encyclopedia of Genes and Genomes) is a learning base for methodical investigation of quality capacities, connecting genomic data with higher request utilitarian data. The genomic data is put away in the GENES database, which is an accumulation of quality lists for all the totally sequenced

genomes and some incomplete genomes with exceptional explanation of quality capacities. The KEGG databases are day by day refreshed and made uninhibitedly accessible (<http://www.genome.ad.jp/kegg/>) [25].

Table 2 Internet resources for whole-genome comparative analysis and associated tools Resource URL [26].

S.No.	Tools	URL
1	KEGG	https://www.genome.jp/kegg/
2	MapView	http://www.ncbi.nlm.nih.gov/mapview/
3	VISTA Genome Browser	http://pipeline.lbl.gov/
4	PipMaker and MultiPipMaker	http://www.bx.psu.edu/
5	zPicture server	http://zpicture.dcode.org/
6	UCSC Genome4 Bioinformatics	http://genome.ucsc.edu/
7	K-BROWSER	http://hanuman.math.berkeley.edu/cgi-bin/kbrowser2
8	GALA	http://www.bx.psu.edu/
9	VISTA server	http://www-gsd.lbl.gov/vista/
10	Comparative Regulatory Genomics	http://corg.molgen.mpg.de/
11	EnsMart	http://www.ensembl.org/
12	EnsMart/ ETOPE	http://www.bx.psu.edu/
13	MAVID server	http://baboon.math.berkeley.edu/mavid/
14	Ensembl	http://www.ensembl.org/
15	rVISTA server	http://rvista.dcode.org

Comparative genomics in drug discovery

Genomics and the related downstream advancements are producing huge informational collections that give new chances to comprehension and battling both irresistible and hereditary infections in people [27].

The predominant worldview in medicate disclosure is the idea of outlining maximally particular ligands to follow up on singular medication targets, the two single most vital explanations behind weakening in clinical improvement are (i) Absence of adequacy and (ii) Clinical wellbeing or toxicology, which each record for 30% of disappointments [28].

The use of structure based medicate configuration has turned out to be moderately regular in lead streamlining. Structural learning of proteins and their ligands has helped in enhancing drug strength and selectivity. This approach has brought about quicker meaning of medication restricting properties and has made it less demanding to recognize 'hit' mixes through screening programs. Both the utilization of X-beam crystallography and NMR (Nuclear Magnetic Resonance) has permitted high throughput approaches for structure-based lead revelation [29].

Comparative or homology demonstrating is a procedure to anticipate protein structure in light of the general perception that proteins with comparable arrangements have comparable structures. The procedure of homology or comparative modeling of proteins comprises of the accompanying steps identification of known 3D structure(s) of a related protein that can fill in as format; (1) arrangement of target and template

proteins; (2) display working for the objective in view of the 3D structure of the layout and the arrangement; (3) refining/ approval/palatable model is fabricated [30]. Comparative genomics also provides a powerful tool for studying evolution. By taking advantage and analyzing the evolutionary relationships between species and the corresponding differences in their DNA, scientists can better understand how the appearance, behavior and biology of living things have changed over time.

Approaches used in comparative genomics

Molecular biology and pharmacology are gradually participating high-throughput facts, with the objective of a more modified form of medication. Genome correspondence the technique of defining the precise correspondence of chromosomal sections and efficient essentials across the species related is the principal stage in comparative genomics. This comprises of defining orthologous (genes diverged consequently a speciation instance) sections of DNA that descend from the same area in the common predecessor of the species compared, and paralogous (genes deviated after a doubling event) sections that ascended by replication events preceding to the separation of the species compared [31].

Whole genome sequences of numerous strains of *P. aeruginosa* have been sequenced in an effort to classify the genetic basis of the alterations in virulence perceived amongst strains [32].

To recognize the basic pathophysiology of disease and to assist the detection of innovative therapeutic replicas of human

illness have long been used practice to investigate different genomes or different gene families in these genomes [33]. Among different species, the phylogenetic study can be achieved for gene families and the synteny study can be completed to categorize the orthologous as well as non-orthologous inheritable factor of a gene family amongst dissimilar species [34].

It is unfavourably desired to recognize the coordination among various functional classes of genes within the growing number of genome sequencing records. A genome wide study of gene organization as well as their promoter functions can provide a substantial perception into the operational functionally associated inheritable factor as gene networks [35].

In accumulation to interspecific studies, intraspecific sequence evaluation yields insights into the evolutionary forces that have represented on a species in the past. In cooperation, both intra- and interspecific sequence evaluations are based on a multiplicity of computational techniques, including alignment, coalescent theory, and phylogenetic reconstruction [36].

There is a constant necessity for similarity search and prognostic tools for the segregation of protein function, non-coding region, coding region, genes, orthologs group and phylogenetic relation in contemporary biological study. Particularly, gene prediction tool GeneMark HMM (Hidden Markov model) Lukashin and Borodovsky and similarity search tool like BLAST Altschul et al. are very active as well as trust commendable computational tools for finding domains and gene prediction individually [37].

For the identification of species and subspecies phylogenetic study using 16S ribosomal RNA (16S rRNA) sequencing is a title standard. However, the procedure cannot discriminate between narrowly related strains such as those found in the genus *Neisseria*; this dispute can be fixed using WGS (Whole-genome sequencing). An excessive number of species and subspecies are now signified in the WGS databank. Matrix Assisted Laser Desorption Ionization Time of Flight mass spectrometry (MALDI-TOF) is conventional in mass spectrometry in general as well as protein breakdown in particular [38].

With the introduction of reasonable techniques for whole genome sequencing, study on extracted DNA may ultimately become archaic. The practice of whole genome sequencing of the DNA of grown-ups is previously the issue of some virtuous conversation [39].

Adversative drug responses are inclined by numerous features, including health, environmental influences as well as genetic features. Pharmacogenetics is an examination field still in expansion and treatment individualization keep on a challenge for the imminent. It is significant to escalate that many genes may affect the reaction to drugs, and the genetic polymorphisms present traditional distinction, which confounds the identification of genetic dissimilarities which are maximum applicable [40].

In cases of gene repetition, both facsimiles of the gene might gather mutations that for instance diminish the practical efficacy of the programmed proteins without hindering this function

overall. In such a circumstance, the molecular purpose (e.g. protein/enzyme activity) would still be accessible to the cell at least to the degree that was obtainable before replication [41].

In the post genomic era, protein arrangement and structural information provide important assistance in reviewing its activity, signalling networks, biochemical pathways, human disease and drug design. One of the vigorous and preeminent methods used to produce protein 3D arrangement is homology modelling or comparative protein modelling. MODELLER is a computer software database used in homology modelling also it practices acceptable 3-D restraint to form a model of a target protein based on homologues protein template [42].

Homology modelling is based on the practical hypothesis of two homologous proteins which shares very uniform like arrangements. The term homology modelling expresses accurately what this technique is about; demonstrating an arrangement using homologous model as template (which is generally a precise X-ray or NMR-determined construction). In homology modelling it is significant that modeller finds a template assembly with the maximum conceivable sequence-identity [43].

DNA transactions, such as duplication, restoration, and recombination comprise DNA blend and successively necessitate the accomplishment of DNA constructing enzymes called DNA polymerases (Pol). A eukaryotic cell comprises at least six dissimilar Pols, entitled alpha, beta, gamma, delta, epsilon, and zeta. Amongst them, Pol, delta inhabits significant characters in DNA duplication, base excision repair, nucleotide excision repair also VDJ recombination [44].

The CRISPR (Clustered, Regularly Interspaced, Short Palindromic Repeats)/CRISPR-Associated Protein 9 (Cas9) systems offers a speedy as well as proficient technology for targeted genome editing. In association with a precisely intended guide RNA (gRNA), Cas9 can accomplish site-specific DNA recognition in addition to cleavage. The site-specific DNA double-strand breaks (DSB) induced by Cas9 activates a Non-homologous End-joining (NHEJ) progression of DNA repair, which leads to small insertions or deletion in the nucleotidic sequence. This practice is therefore subjugated to produce loss-of-function of protein coding genes, via variation of the open reading frame [45].

Horizontal Gene Transfer (HGT) which is also entitled as lateral gene transfer is a non-sexual moment of genetic material between the two organisms. The impact of HGT was unequal across the eukaryotes. Many microbial eukaryotes and some of the plant mitochondria are good examples for HGT in eukaryotes. Some other extractions are performing to be immune for attaining the new genes. Another imperative arrangement in eukaryotes is in HGT eukaryotes implicates the genes from the bacteria. Genes that are learnt from the bacterial sources are diverse in dissimilar species. The number may be zero to hundred [46].

Plasmid mediated HGT of β -lactamase genes convening confrontation to third generation prolonged spectrum beta lactams in addition to fourth generation carbapenems transpired to an azide-resistant recipient *E. coli* if contributor as

well as receiver cells were mixed together on contemporary stainless steel surfaces and in suspension but not on copper alloy surfaces [47]. For developing parts of genomics, proteomics, and transcriptomics in the revelation as well as authorization of human colorectal tumour biomarkers from DNA/RNA sequencing information under synchrotron radiation, it has exposed numerous solicitations in natural systems. Lion's share of restorative and pharmaceutical nano synergist hydrogenations is still completed utilizing heterogeneous nano impetuses because of the procedure focal points, for example, capacity, simple division, and extensive variety of relevant response conditions [48]. Microbial communities have been studied using several different analysis approaches, including the analysis of 16S ribosomal variable regions and direct sequencing of the "metagenome". However, metatranscriptomics is a more powerful tool for understanding microbial processes within communities and consortia as it not only reveals the organisms present, but also provides information on their function and how this is influenced in different environments. Transcriptomics or proteomics these complex atomic strategies in blend with bioinformatics are utilized for top-down ways to deal with create new speculations, for instance on cell motioning in immature microorganisms that are regularly confirmed by ensuing research facility experiments [49].

One of remarkable achievements in transcriptomics is to identify noncoding RNA in the human genome. Coding regions of the human genome occupy less than 5% of the genome, while rest of the genome comprises noncoding regions that generate noncoding RNAs [50].

Tree portrayal of the family history of set of arrangements that offer a typical progenitor is known as a Phylogenetic Tree. A phylogeny tree demonstrates the association among different creatures and weight of the branches in the tree shows time between advancements of various organisms [51]. To appraise the quantity of substitutions that really happened in transformative history, a model of arrangement development is expected to foresee the impact of developmental separation on phenotypic separation. Probabilistic techniques in view of progress rate frameworks have been produced to catch the impact of the irregularity of the mutational procedure and of here and now determination on long haul advancement [52].

Future of comparative genomics

Looking at in excess of two genomic successions gives much all the more settling power. The adequacy of numerous arrangements for useful forecast is demonstrated drastically by the examinations of 13 genomic groupings from species going from fish to people. Other methodologies utilizing different successions from all the more firmly related species significantly enhance the settling intensity of near genomics. The Human Genome Project perceives the intensity of this wide similar examination [53].

NHGRI spearheaded the improvement of DNA sequencing strategies and advancements - including informatics - and has subsidized research to contemplate the genomes of a wide scope of animal groups. The National Institutes of Health (NIH) Intramural Sequencing Center has been instrumental in the

sequencing of numerous life forms. In mod ENCODE, analysts found shared examples of quality action and direction among fly, worm and human genomes. The mouse ENCODE Consortium exhibited that, by and large, the frameworks that are utilized to control quality movement have numerous likenesses in mice and people [54].

- Aligning entire genomes with part arrangements
- Looking at all the more indirectly related genomes (managing non-colinearity)
- Finding shared trait among obviously inconsequential or inaccessible genomes (information digging for shared systems) Our present apparatuses require a rodent Massive scaling of alignments
- Comparison of different draft genomes
- Better managing inaccurate matches
- Use of huge new PCs [55]

The database and web-tool STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) is a meta-resource that aggregates most of the available information on protein-protein associations, scores and weights it, and augments it with predicted interactions, as well as with the results of automatic literature-mining searches [56].

Conclusion

Genomic tools area unit remarkably complementary to alternative views, together with morphology, useful biology, development and biological science, and area unit serving to unify antecedently freelance analysis programs in these areas. Currently that order information area unit less costly to gather than information on several alternative organism attributes, genomes are going to be progressively helpful as a primary check up on organism biology that helps guide alternative forms of observations. Most methodologies created to anticipate harmful transformations were prepared utilizing human information and as a rule, must be utilized for human proteins.

References

1. Sharma RK (2016) A Review on New Horizons of Bioinformatics in Next Generation Sequencing, Viral and Cancer Genomics. *Int J Biomed Data Min* 5:1000122.
2. Ali A, Soares SC, Barbosa E, Santos AR, Barh D, et al. (2013) Microbial Comparative Genomics: An Overview of Tools and Insights Into The Genus *Corynebacterium*. *J Bacteriol Parasitol* 4:167.
3. Brim H, Ashktorab H (2016) Genomics of Colorectal Cancer in African Americans. *Next Generat Sequenc Applic* 3:133.
4. Mukherjee D, Diehl WJ (2017) Speciation Genomics of Protein-Coding Genes Common to Mycoplasmales. *J Phylogenetics Evol Biol* 5:175.
5. Sudhakara Reddy B, Sivajothi S (2017) Proteomics and Genomics in Veterinary Parasitology -A Diagnostic Tool. *J Mol Biomark Diagn* 8:367.
6. Touchman J (2010) Comparative Genomics. *Nature Education Knowledge* 3:13.

7. Sivashankari S, Shanmughavel P (2007) Comparative genomics - A perspective. *Bioinformation* 1:376-378.
8. Nava GM, Hernandez YE (2016) Comparative Genomics of Salmonella Could Reveal Key Features of Adaptation. *J Data Mining Genomics Proteomics* 7:e121.
9. Galperin MY (2013) Comparative genomics: From genome sequences to genome biology. *J Data Mining Genomics Proteomics* 4:5.
10. Reinhold WC (2013) The current state of comparative genomics and pharmacogenomics, and the application of the NCI-60 resources and cell miner tools to these problems. *J Data Mining Genomics Proteomics* 4:5.
11. Wei L, Liu Y, Dubchak I, Shon J, Park J (2002) Comparative genomics approaches to study organism similarities and differences. *J Biomed Inform* 35:142-150.
12. Brocchieri L (2013) Phenotypic and Evolutionary Distances in Phylogenetic Tree Reconstruction. *J Phylogen Evolution Biol* 1:e106.
13. Lange BM, Ghassemian M (2005) Comprehensive post-genomic data analysis approaches integrating biochemical pathway maps. *Phytochemistry* 66:413-451.
14. Hammond JP, Bowen HC, White PJ, Mills V, Pyke KA, et al. (2006) A comparison of the *Thlaspi caerulescens* and *Thlaspi arvense* shoot transcriptomes. *New Phytol* 170:239-60.
15. Jones SG, Moxon S, Marshall M, Khanna A, Eddy SR, et al. (2005) Rfam: Annotating non-coding RNAs in complete genomes. *Nucleic Acids Res* 33:D121-D124.
16. Andolfatto P (2005) Adaptive evolution of non-coding DNA in *Drosophila*. *Nature* 437:1149-1152.
17. Jareborg N, Birney E, Durbin R (1999) Comparative Analysis of Noncoding Regions of 77 Orthologous Mouse and Human Gene Pairs. *Genome Res* 9:815-824.
18. Zhai SL, Li XP, Chen QL, Chen SN, Wei WK, et al. (2015) Sequence Comparison and Phylogenetic Analysis of Several Emerging Porcine Bocaviruses and a Proposal Regarding Nomenclature. *Virology* 4:142.
19. Hukowska-Szematowicz B, Tokarz-Deptuła, Deptuła W (2014) Phylogenetic Analysis of Selected RHDV Strains on the Basis of a Fragment of the Gene Encoding C-Terminal End of VP60 Capsid Structural Protein. *J Phylogen Evolution Biol* 2:130
20. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, et al. (2009) Circos: an Information Aesthetic for Comparative Genomics. *Genome Res* 19:1639-1645.
21. Alikhan NF, Petty NK, Zakour NLB, Beatson SA (2011) BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12:402.
22. Gupta H, Badapanda C, Ghosh A, Chikara S (2016) RareDDB: An Integrated Catalog of Rare Disease Database. *Clin Med Biochemistry* 2:2.
23. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I (2004) VISTA: computational tools for comparative genomics. *Nucleic Acids Res* 32:273-279.
24. Clamp M, Andrews D, Barker D, Bevan P, Cameron G, et al. (2003) Ensembl 2002: accommodating comparative genomics. *Nucleic Acids Res* 31:38-42.
25. Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes *Nucleic Acids Res* 28:27-30.
26. Brown JB, Celniker SE (2004) Lessons from mod ENCODE. *Annu Rev Genom Human Genet* 5:15-56.
27. Cole ST (2002) Comparative mycobacterial genomics as a tool for drug target and antigen discovery. *Eur Respir J* 20:78-86.
28. Lundstrom K (2007) Structural genomics and drug discovery. *J Cell Mol Med* 11:224-238.
29. Cavasotto CN, Phatak SS (2009) Homology modeling in drug discovery: current trends and applications. *Drug Discov Today* 4:676-683.
30. Kellis M, Patterson N, Birren B, Berger B, Lander ES (2004). Methods in comparative genomics: genome correspondence, gene identification and regulatory motif discovery. *J Comput Biol* 11:319-55.
31. L Brocchieri (2014) The coding potential of *Pseudomonas aeruginosa*: Gene discovery across strains by computational comparative genomics and ribosome profiling. *J Mol Genet Med* 8:3.
32. Stoll M, Kwitek-Black AE, Cowley AW, Harris EL, Harrap SB, et al. (2000) New Target Regions for Human Hypertension via Comparative Genomics. *Genome Res* 10:473-482.
33. Rawal HC, Singh NK, Sharma TR (2012) CGPSA- Comparative Genomics Methodology for Phylogenetic and Synteny Analysis of gene families in different species. *J Biotechnol Biomater* 2:6.
34. Nalla VK, Datta A (2012) Comparative genomics and genome wide analysis of bidirectional promoters in *Candida albicans*. *J Biotechnol Biomater* 2:6.
35. Haubold B, Wiehe T (2004) Comparative genomics: methods and applications. *Naturwissenschaften* 91:405-421.
36. Saravanan V (2010) Mass Blaster V1.0 - A Perl Gui Tool for Mass Sequence Blast and Gene Prediction. *J Proteomics Bioinform* 3: 302- 304.
37. Jeongsook Y (2017) Recent Technology of Clinical Microbiology; Whole Genome Sequencing (WGS) and Matrix-Assisted, Laser Desorption Ionization-Time of Flight Mass Spectrometry (MALDI-TOF MS). *J Microb Biochem Technol* 9:138-142.
38. Hens K (2011) Whole Genome Sequencing of Children's DNA for Research: Points to Consider. *J Clin Res Bioeth* 2:106e.
39. Bandu R, Kim KP (2016) Editorial: Pharmacogenetics, Rozana, J Drug Metab Toxicol 7:3.
40. Mukherjee P, Roy P (2013) Cloning, Sequencing and Expression of Novel Trichloroethylene Degradation Genes from *Stenotrophomonas maltophilia* PM102: A Case of Gene Duplication. *J Bioremediat Biodegrad* 4:177.
41. Kumar V, Meyer D, Lall N (2017) Comparative Investigation of Techniques Used for Mtr Homology Model Refinement. *J Proteomics Bioinform* 10: 284-297.
42. Mishra S, Gomase VS (2016) Computational Comparative Homology Based 3D-structure Modelling of the HSp70 Protein from GWD. *J Health Med Informat* 7:233.
43. Ikram Manzoor (2017) Homology modeling, binding site identification and molecular docking of DNA polymerase delta in *Plasmodium falciparum*. 3rd World Congress on Human Genetics and Genetic Disorders in Pakistan. *J Mol Genet Med*.
44. Tongelen AV, Lorient A, De Backer O, De Smet C (2015) Application of CRISPR/cas9-Directed Homologous Recombination to the Generation of Human Tumor Cells with Conditional Knockout of an X-Linked MicroRNA Locus. *Gene Technol* 4:124.

45. Rajarapu G (2014) Horizontal Gene Transfer. *J Phylogen Evolution Biol* 2:136.
46. Keevil CW (2016) Prevention of horizontal gene transfer of antibiotic resistance genes on touch surfaces. *Clin Microbiol* 5:5.
47. Heidari A (2017) Integrative Approach to Biological Networks for Emerging Roles of Proteomics, Genomics and Transcriptomics in the Discovery and Validation of Human Colorectal Cancer Biomarkers from DNA/RNA Sequencing Data under Synchrotron Radiation. *Transcriptomics* 5:e117.
48. Gorni C (2014) Metatranscriptomics in the NGS Era. *Transcriptomics* 2:e107
49. Gosau M, Morsczech C (2012) Bioinformatics, Transcriptomics, Proteomics and Dental Follicle Precursor Cells. *J Biochip Tissue chip* 2:e111.
50. Kurokawa R (2013) Transcriptomics Inspires Long Noncoding RNA. *Transcriptomics* 1:1.
51. Kudipudi S, Allam AR, SridharGR, Srinubabu G (2008) Methodology for Phylogenetic Tree Construction. *J Proteomics Bioinform* S1:S005-S011.
52. Hardison RC (2003) Comparative Genomics. *PLoS Biol* 1:e58.
53. <https://www.genome.gov/11509542/comparative-genomics-fact-sheet/>
54. Chain P, Kurtz S, Ohlebusch E, Tom Slezak (2003) An applications-focused review of comparative genomics tools: Capabilities, limitations and future challenges. Henry Stewart Publications 4:105–123
55. Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, et al. (2008) STRING 8- A global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res* 37:412-416.
56. Nobrega MA, Pennacchio LA (2004) Comparative genomic analysis as a tool for biological discovery. *J Physiol* 1:31-39.