

## REVIEW ARTICLE

# Evolution of dental informatics as a major research tool in oral pathology

Sasidhar Singaraju, Prasad H, Medhini Singaraju

Department of Oral and Maxillofacial Pathology, Modern Dental College and Research Center, Airport Road, Gandhi Nagar, Indore, India

**Address for correspondence:**

Dr. Sasidhar Singaraju,  
Department of Oral and Maxillofacial Pathology,  
Modern Dental College and Research Center,  
Airport Road, Gandhi Nagar, Indore, India.  
E-mail: [ssingaraju\\_64@yahoo.com](mailto:ssingaraju_64@yahoo.com)

**ABSTRACT**

Bioinformatics is a relatively new field of science that incorporates the principles of biology and computer science. It is mainly concerned with collecting, storing, and retrieving data from large databases. Ever since the successful completion of the Human Genome Project, there has been an exponential growth in the volumes of biological data that is being generated worldwide. The evolution of bioinformatics has made it possible to access these databases and apply the information for better research. One discipline that has been benefitted from the advent of bioinformatics is oral pathology. Oral pathology is a branch of dentistry which deals with the diseases of the head and neck region. Remarkable progress has been made in the diagnosis and treatment of diseases with the aid of bioinformatics. This article attempts to review the evolution and progress of dental informatics as a specialty and its applications in oral pathology.

**Key words:** Bioinformatics, databases, dental informatics, oral pathology

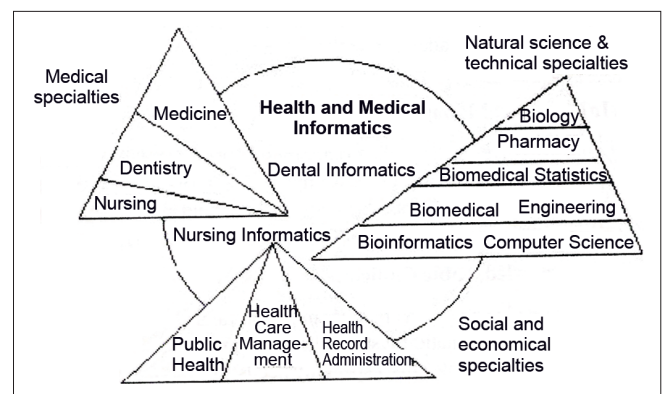
## INTRODUCTION

Bioinformatics is a combination of biology and information technology. The discipline encompasses many computational tools and methods used to manage, analyze, and manipulate large sets of biological data.<sup>[1]</sup> According to Hwa A. Lim, bioinformatics is the new subject that is concerned with the collection and analysis of genetic data and its dissemination to the research community.<sup>[2]</sup> The ultimate goal of this field is to facilitate new biological insights as well as to create a global perspective from which the unifying principles in biology can be discerned. Over the past few decades, huge volumes of information have been generated through advances in biomedical research. Bioinformatics is important because it enables us to manage these data more easily, to sift through the data for useful encoded gene segments, and to use this information for clinical decision making and developing new drugs and therapies.<sup>[3]</sup> Biomedical informatics incorporates the knowledge of health sciences (medicine, dentistry, pharmacy, nursing, and allied health sciences) with computer science, management and decision science, biostatistics,

engineering, and information technologies. This involves the integrated use of several approaches and techniques from these sciences to solve problems relevant to healthcare delivery, health sciences, pharmaceutical research, education, and clinical/medical decision making [Figure 1].<sup>[4]</sup>

The primary goal of bioinformatics is to increase the understanding of biological processes. What sets bioinformatics apart from other approaches, however, is its focus on developing and applying computationally intensive techniques (e.g., pattern recognition, data mining, machine learning algorithms, and visualization) to achieve this goal. Major research efforts in the field include sequence alignment, gene finding, genome assembly, drug design, drug discovery, protein structure

Access this article online	
<b>Quick Response Code:</b> 	<b>Website:</b> <a href="http://www.jomfp.in">www.jomfp.in</a>
	<b>DOI:</b> 10.4103/0973-029X.92979



**Figure 1:** Flowchart showing interrelationship between health and medical informatics and other closely related specialties

alignment, protein structure prediction, prediction of gene expression and protein–protein interactions, genome-wide association studies, and the modeling of evolution.<sup>[5]</sup>

## EVOLUTION OF DENTAL INFORMATICS

Dental informatics refers to the application of computer and information science to improve dental practice, research, education, and management.<sup>[6]</sup> Dental informatics can be considered a specialty of medical informatics. A number of models, methods, and applications can be shared and/or transferred between dental practice and computer sciences. For instance, the National Library of Medicine's (NLM), MEDLINE is the world's largest biomedical literature database and stores data from all health care disciplines. The same applies to GenBank, a database of gene and protein sequence information, to which researchers from several domains contribute. Methods for building expert and decision-support systems, such as neural and Bayesian networks, can be applied across disciplines.<sup>[7]</sup> The main goal of dental informatics is to improve patient outcomes. Thus, the discipline must support and improve diagnosis, treatment and prevention of disease and traumatic injury; relieve pain; and preserve and improve oral health. A secondary goal is to make the delivery of dental care more efficient, for example, by maintaining or improving cost–benefit ratios. Dental informatics also must support research and education; improvements in these areas should, and often do, translate into improved patient care. Dentists must be as familiar with the problem-solving process as they are with the problem domain itself. Computers can help practitioners maintain their continuing competency, and many dentists are already using computers to keep abreast of new developments.<sup>[8,9]</sup>

Over a period of several years dentistry has evolved into a major specialty in itself, with many branches (e.g., orthodontics, periodontics, etc.) taking care of different aspects of oral diseases. One such discipline is oral pathology, which represents the confluence of the basic sciences and clinical dentistry. Knowledge in this field is acquired through the adaptation of methods and disciplines of those sciences basic to dental practice, such as gross and microscopic anatomy, chemistry, microbiology, and physiology, and through information obtained by clinical histories and observation of patients. Through the science of oral pathology, an attempt is made to correlate human biology to the signs and symptoms of human disease so that it can be properly diagnosed and adequately treated.<sup>[10]</sup> The subject of oral pathology mainly deals with pathologies associated with the head and neck region, which comprises oral cancer, odontogenic cysts and tumors, salivary gland pathology, dental caries and associated pathology, and other microbiological diseases or lesions affecting the head and neck region. The beginning of the 21<sup>st</sup> century saw an enormous amount of information being generated on head and neck pathology, accessing which

necessitated comprehensive integrated systems in the field of dental informatics.

## APPLICATIONS OF DENTAL INFORMATICS IN ORAL PATHOLOGY

### Early detection of oral cancer

Oral cancer is one of the most common causes of cancer-related deaths worldwide. Detection of oral cancer in the early asymptomatic stage dramatically improves cure rates and patients' quality of life by minimizing extensive debilitating treatments.<sup>[11]</sup> Unfortunately, more than 50% of patients with oral cancer display evidence of spread to regional lymph nodes and metastases at the time of diagnosis, and approximately two-thirds of patients have apparent symptoms at presentation, a negative prognostic indicator.<sup>[12]</sup> Although screening has been emphasized as a method of reducing the morbidity and mortality associated with oral cancers, the visual detection of oral cancer at an early stage is significantly hindered by the difficulty in clinically differentiating premalignant and malignant lesions from similar-looking benign lesions.<sup>[13,14]</sup> Precancers and early-stage oral cancers cannot be adequately identified by visual inspection alone and may easily be overlooked and neglected, even by highly-trained professionals with broad experience. Thus, a method of detection at early, curable stages is crucial and may lead to a reduction in the currently unacceptably high oral cancer morbidity and mortality rates.<sup>[13]</sup> Evolution of the OralCDx<sup>®</sup> system has significantly contributed to negating this drawback. A critical component of OralCDx<sup>®</sup> is the use of computer-assisted image analysis of the oral brush biopsy sample. Although automated cytology had been proposed in the late 1950s as a method of reducing false negative findings, early attempts that relied on analysis by algorithmic computers were not successful. This limitation was finally overcome by the application of new, non-algorithmic, neural network computers that were developed in the late 1980s for missile defense. In recent years, neural networks have been successfully applied to several medical diagnostic procedures, including cervical smear screening and interpretation of digital radiologic images such as chest radiographs and mammograms.<sup>[15–17]</sup>

The OralCDx<sup>®</sup> neural network assists in the search for potentially abnormal cells in oral brush biopsy samples, which then are interpreted by the pathologist. The identification of these abnormal cells is labor intensive, fatiguing, and time consuming; more importantly, these abnormalities are also easily overlooked.

The OralCDx<sup>®</sup> images of the neural network–selected cells are presented to the pathologist for review. With this method—optimizing human and computer capabilities—it becomes possible to identify cellular abnormalities that might otherwise have been missed with manual microscopic screening.<sup>[18]</sup>

## Genomics and proteomics in oral pathology

Scientific progress drives innovation in health care. From digital radiographs to smart dental materials that stimulate tooth repair, new technologies have significantly improved diagnostic quality, patient comfort, and efficiency in dental care. Today's scientific breakthroughs are leading to increasingly powerful and sophisticated tools that promise to improve the lives of millions across the globe. The mapping of the human genome in 2003 triggered the development of powerful research tools such as comparative genomic hybridization, serial analysis of gene expression, and DNA arrays. The combination of these analytic tools with dizzyingly fast computers is allowing us to understand complex biological systems from a molecular perspective. Scientists now can extract and integrate vital genomic and proteomic information to diagnose diseases more precisely at their earliest stages, to direct personalized therapy, and to predict and evaluate disease outcomes more accurately.<sup>[19]</sup>

Genomics is a discipline in genetics that is concerned with the study of the genomes of organisms. The field includes intensive efforts to determine the entire DNA sequence of organisms and fine-scale genetic mapping efforts. Thus, genomics is the study of all the genes of a cell or tissue at the DNA (genotype), mRNA (transcriptome), or protein (proteome) levels.<sup>[20]</sup> Proteomics involves the identification and compilation of the complete catalog of proteins produced by cells in normal and diseased tissues. Metabolomics is a large-scale approach to monitoring and providing a complete score card of all small molecules, such as lipids, sugars, and amino acids involved in daily cellular function.

Many oral diseases and disorders, including dental caries, cleft lip/palate, and a host of craniofacial syndromes, are complex conditions that arise from the actions of multiple genes and their interactions with one another, the environment, and other factors. It will take more than adding up the molecular parts to understand and address such multifactorial disorders. An interdisciplinary approach is needed to integrate the complex web of molecular information with clinical information, particularly for diseases where the diagnosis is based primarily on clinical findings.<sup>[19]</sup>

Management of head and neck cancer is still based primarily on the evaluation of macroscopic tumor characteristics and extent of disease. But if genomic information could be factored into treatment decisions, it might be possible to predict which patients require aggressive treatment and which do not. Surgeons might be able to distinguish healthy tissue from cancerous tissue more precisely, thus minimizing disfiguration and loss of function.<sup>[19]</sup>

The word 'proteome' originated in 1994 when Marc Wilkins coined the term to describe the set of all PROTEINS expressed by a genome (Wilkins *et al.* 1996). While genomics and

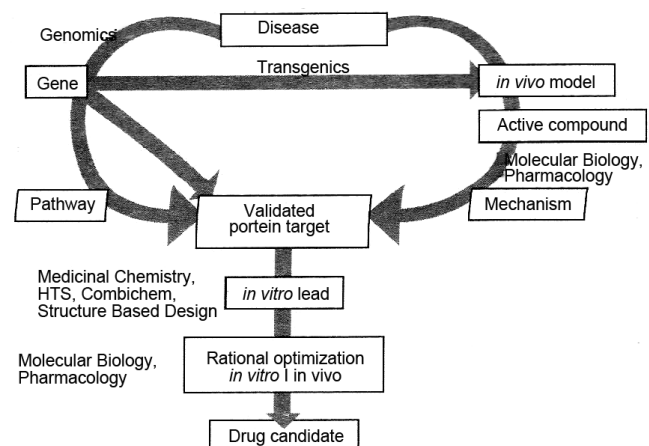
transcriptomics provide basic information on DNA sequences, regulatory elements, and gene expression, proteomics provides quantitative information on the total protein profile of a cell, tissue, or organism. Proteomics allows the level of protein expression to be evaluated and can be used to determine the presence of protein isoforms and post-translational modifications or to examine protein-protein interactions.<sup>[21]</sup>

Comparative proteomics allows the identification of several diseases/causative pathogens of the head and neck region. Some proteomic studies have already been conducted in the field of oral pathology and have led to the identification of risk factors and therapeutic targets at a molecular level.<sup>[22,23]</sup> Further advances in the application of proteomics to dental research and oral pathology are being considered, such as the identification of molecules and molecular pathways involved in oral diseases, the limitation of the inflammatory response after dental procedures, and the possible programming of cellular and tissue response upon the placement of restorative materials or implants [Figure 2].

The National Institute of Dental and Craniofacial Research (NIDCR) in the US has undertaken a project to catalog all the proteins contained in human oral fluids. This Salivary Proteome Project and other efforts to map small-molecule metabolites are a key step toward developing saliva-based diagnostic systems capable of rapid, noninvasive, and safe identification of molecules and substances that can otherwise be detected only through blood tests. The progress being made toward salivary diagnostics is the result of collaboration between engineers, biologists, dental researchers, and dental clinicians.<sup>[19]</sup>

## Microarrays

Microarray analysis is a new technology in which all of the genes of an organism are represented by oligonucleotide sequences spread out in an 80 × 80 array on microscope



**Figure 2:** Flowchart illustrating the role of genomics and proteomics in the identification of risk factors and therapeutic targets at a molecular level

slides, and can also be synthesized directly on the slides at densities of up to one million per square centimeter. The oligonucleotides are collectively hybridized to a labeled cDNA library prepared by reverse transcribing mRNA from cells. The amount of label binding to each oligonucleotide spot reflects the amount of mRNA present in the cell. From this analysis, a set of genes that respond in an identical manner may be identified. Microarray analysis is designed to detect global changes in transcription in the genome but does not provide information about the levels of protein products of the genes, which may also be subject to translational regulation.<sup>[24]</sup> Protein arrays can be used in such instances.

### *Applications of microarrays in oral pathology*

In the oral cavity, several conditions including oral cancer and precancer have been found to have some genetic basis. As of now, several precancerous conditions and lesions have been identified, though only a small fraction of these turn into malignancy. However, there is no single method by which the behavior of these lesions can be predicted. It is understood that genetic alterations take place much before morphologic changes in the oral cavity occur. Microarrays may be able to identify those genetic alterations that are more likely to determine the progression of a premalignant lesion to frank malignancy.<sup>[25,26]</sup>

Microarrays can also be used in several other fields. Information about differences in gene expression between tissue types can help us to uncover the mysteries about growth and development.<sup>[27,28]</sup> One can also examine targets for drug discovery upon identifying the gene that plays a significant role in a particular disease. For example, if we compare normal cells to cancer cells and discover a gene that is turned on in that particular type of cancer, we may be able to target that gene in cancer therapy. Today, therapies that directly target malfunctioning genes are already in use and are showing exceptional results. Microarrays are used commonly to detect viruses and other pathogens from blood samples. More recently they have been used to identify inheritable markers, and therefore have been used as a genotyping tool.<sup>[29]</sup>

### **DATABASES USED IN DENTAL INFORMATICS**

The rapid development of genome technologies, especially automatic sequencing techniques, has produced a huge amount of data consisting essentially of nucleotide and protein sequences. DNA sequence databases were first assembled at Los Alamos National Laboratory (LANL), New Mexico, by Walter Goad and colleagues in the GenBank database and at the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany. GenBank is now under the auspices of the National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov>). In 1984, the DNA DataBank of Japan (DDBJ) (<http://www.ddbj.nig.ac.jp>), Mishima, Japan, came into existence. GenBank, EMBL, and DDBJ have now formed the International Nucleotide Sequence Database

Collaboration (<http://www.ncbi.nlm.nih.gov/collab>), which acts to facilitate exchange of data on a daily basis.

A number of databases have similarly been established for protein information; these include Protein Information Resource (PIR) (<http://pir.georgetown.edu/>), Kyoto Encyclopedia of Genes and Genomes (KEGG) (<http://www.genome.jp/kegg/>), Database of Interacting Proteins (DIP) (<http://dip.doe-mbi.ucla.edu>), and Biomolecular Interaction Network Database (BIND) (<http://www.bind.ca>). The structural databases store data on proteins and nucleic acid structure. Such databases include Protein Data Bank (PDB) (<http://www.pdb.org/>), Nucleic Acid Data Base (NADB) (<http://ndbserver.rutgers.edu/>), and Molecular Modeling Data Base (MMDB) (<http://www.ncbi.nlm.nih.gov/Structure/MMDB/mmdb.shtml>).

### **TOOLS FOR DATABASE SEARCHING FOR SIMILAR SEQUENCES**

Searching a sequence database for sequences that are similar to the query sequence is the most common type of database similarity search. The first rapid search method was FASTA, which found short common patterns in the query and database sequences and joined these into an alignment. The Basic Local Alignment Search Tool (BLAST) is similar to FASTA but is gaining popularity because of its ability to search rarer and more significant sequences. PAM250 Scoring Matrix is a scoring matrix based on an evolutionary model that predicts the type of amino acid changes over a long period of time. PSI-BLAST and PHI-BLAST are iterative search methods that improve on the detection rate of BLAST and FASTA. Pfam and Simple Modular Architecture Research Tool (SMART) are used for protein domain family analysis. Phylogenetic analysis using parsimony (PAUP) and phylogenetic inference package (PHYIP) are software packages available for the phylogenetic analysis of macromolecular sequences. The program RasMol is the best-known viewer for macromolecular structures.<sup>[24]</sup>

### **CONCLUSION**

Genomics and proteomics hold the promise to change the practice of dentistry and oral pathology; their potential to identify risk factors and therapeutic targets at a molecular level is established. However, mass-scale molecular genomics and proteomics suffer some major drawbacks: Gene and protein expression is not significant per se, but become useful only when applied to gene–gene, gene–protein, and protein–protein interactions. Genomics and proteomics together, i.e., epigenetics, will play a vital role in providing better understanding of disease processes at a molecular level. Theoretical disciplines like bioinformatics and data mining will aid us greatly in understanding this complex picture. Current state of the art in both genomics and proteomics and the theoretical disciplines suggests that a proper combination of experimental and theoretical results, obtained with different methods, will soon become the gold standard for the study of oral diseases.

## REFERENCES

1. Westhead DR, Parish JH, Twyman RM. *Bioinformatics*. 1<sup>st</sup> ed. UK: BIOS Scientific Publishers Limited; 2003.
2. Hwa A Lim. *Genetically yours: Bioinforming, biopharming, and biofarming*. New Jersey : Word Scientific Publishing Co; 2002. p. 273-95.
3. Srivastava. *Introduction to bioinformatics*. 1st ed. India: Shree Publications; 2007..
4. Bansal M. *Medical informatics: A primer*. 1st ed. India: Tata Mcgraw-Hill; 2003.
5. Rao VS, Das SK, Rao VJ, Srinubabu G. Recent developments in life sciences research: Role of bioinformatics. *Afr J Biotechnol* 2008;7:495-503.
6. Eisner J. The future of dental informatics. *Eur J Dent Educ* 1999;3 Suppl 1:61-9.
7. White SC. Decision-support systems in dentistry. *J Dent Educ* 1996;60:47-63.
8. Kalkwarf KL. How the licensure process will evolve. *J Am Dent Assoc* 1999;130:1737-42.
9. Schleyer TK, Forrest JL, Kenney R, Dodell DS, Dovgy NA. Is the Internet useful for clinical practice? *J Am Dent Assoc* 1999;130:1501-11.
10. Rajendran R, Sivapathasundaram B. *Shafer's Textbook of Oral Pathology*. 6<sup>th</sup> ed. India; Elsevier; 2009.
11. National Cancer Institute. *Cancer statistics review, 1973-1990*. Bethesda, Md.: U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health; DHHS publication (NIH) 93-2789; 1993.
12. Silverman S. *American Cancer Society. Oral cancer*. Hamilton, Ontario, Canada: B.C. Decker; 1998: xvi, 174.
13. Silverman S Jr. Early diagnosis of oral cancer. *Cancer* 1988;62(8 Suppl):1796-9.
14. Shugars DC, Patton LL. Detecting, diagnosing, and preventing oral cancer. *Nurse Pract* 1997;22:105,109-10,113-5.
15. Ishida T, Katsuragawa S, Ashizawa K, MacMahon H, Doi K. Application of artificial neural networks for quantitative analysis of image data in chest radiographs for detection of interstitial lung disease. *J Digit Imaging* 1998;11:182-92.
16. Mango LJ. Computer-assisted cervical cancer screening using neural networks. *Cancer Lett* 1994;77:155-62.
17. Wu Y, Giger ML, Doi K, Vyborny CJ, Schmidt RA, Metz CE. Artificial neural networks in mammography: Application to decision making in the diagnosis of breast cancer. *Radiology* 1993;187:81-7.
18. Sciubba JJ. Improving detection of precancerous and cancerous oral lesions. *J Am Dent Assoc* 1999;130:1445-57.
19. Garcia I, Tabak LA. Beyond the 'omics': Translating science into improved health. *J Am Dent Assoc* 2008;139:392-5.
20. EPA Interim policy on genomics. Available from: <http://epa.gov/osa/spc/pdfs/genomics.pdf>. [Last accessed on 2011 Feb 8].
21. Macarthur DJ, Jacques NA. Proteome analysis of oral pathogens. *J Dent Res* 2003;82:870-6.
22. Wu Y, Shu R, Luo LJ, Ge LH, Xie YF. Initial comparison of proteomic profiles of whole unstimulated saliva obtained from generalized aggressive periodontitis patients and healthy control subjects. *J Periodontol Res* 2009;44:636-44.
23. Oluwadara O, Giacomelli L, Christensen R, Kossan G, Avezova R, Chiappelli F. LCK, survivin and PI-3K in the molecular biomarker profiling of oral lichen planus and oral squamous cell carcinoma. *Bioinformation* 2009;4:249-57.
24. Mount DW. *Bioinformatics. Sequence and genome analysis*. 1<sup>st</sup> ed. New York: Cold Spring Harbor Laboratory Press; 2001.
25. Bowden JR, Brennan PA. DNA microarray technology: Insights for oral and maxillofacial surgeons. *Br J Oral Maxillofac Surg* 2004;42:542-5.
26. Kuo WP, Whipple ME, Epstein JB, Jenssen TK, Santos GS, Ohno-Machado L, *et al.* Deciphering gene expression profiles generated from DNA microarrays and their applications in oral medicine. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2004;97:584-91.
27. Available from: <http://www.gene-chips.com/>. [Last accessed on 2011 Aug 8].
28. Available from: <http://www.microarrayworld.com/>. [Last accessed on 2011 Aug 8].
29. Kuo WP, Whipple ME, Jenssen TK, Todd R, Epstein JB, Ohno-Machado L, *et al.* Microarrays and clinical dentistry. *J Am Dent Assoc* 2003;134:456-62.

**How to cite this article:** Singaraju S, Prasad H, Singaraju M. Evolution of dental informatics as a major research tool in oral pathology. *J Oral Maxillofac Pathol* 2012;16:83-7.

**Source of Support:** Nil. **Conflict of Interest:** None declared.