ORIGINAL RESEARCH

DataGenno: building a new tool to bridge molecular and clinical genetics

Fabricio F Costa[1,2]
Luciano S Foly[1]
Marcelo P Coutinho[1]

[1]DataGenno Interactive Research Ltd., Itaperuna, Rio de Janeiro, Brazil; [2]Cancer Biology and Epigenomics Program, Children's Memorial Research Center, Northwestern University's Feinberg School of Medicine, Chicago, IL, USA

Abstract: Clinical genetics is one of the most challenging fields in medicine, with thousands of children born every year with congenital defects that have no satisfactory diagnosis. There are more than 6,000 known single-gene disorders that can cause birth defects or diseases in approximately 1 in every 200 births. Clinical and molecular information on genetic diseases and syndromes are widespread in the literature, and there are few databases combining this information. Therefore, it is very challenging for health care professionals and researchers to translate the latest advances in science and medicine into effective clinical interventions and new treatments. In order to overcome this obstacle and promote networking, we are building DataGenno, an online medical and scientific portal. DataGenno has been developed to be a source of information on genetic diseases and syndromes for the needs of all heath care professionals and researchers. Our database will be able to integrate both clinical and molecular aspects of genetic diseases in a fully interactive environment. DataGenno's system already contains clinical and molecular information for 300 diseases, with approximately 6,000 signs and symptoms of these diseases in a database combined with a search engine. Our main goal is to cover all genetic diseases described to date, providing not only clinical information such as morphological and anatomical features but also the most comprehensive molecular genetics/genomics features and available testing information. We are also developing ways to connect DataGenno's portal with Electronic Health Records in order to improve the efficiency of patient care. Additionally, DataGenno's system and search engine will be able to provide tools that will facilitate the discovery and description of new genetic syndromes. In conclusion, we believe that DataGenno's portal will be a helpful and innovative tool for health care professionals, scientists, genetic counselors, and other professionals in the clinical genetics field.

Keywords: genetic diseases, signs and symptoms, molecular genetics, genomics, search engine, database

Introduction

There are more than 6,000 known single-gene disorders, which occur in about 1 in every 200 births;[1] examples include cystic fibrosis, sickle cell anemia, Marfan syndrome, Huntington's disease, and hereditary hemochromatosis. Clinical and molecular information for genetic diseases and syndromes, especially monogenic diseases, are extensive in the scientific literature and all over the Internet, in several public and private websites. Thus, there is an increasing need for better tools to support and connect professionals working with diseases that have a genetic cause, in both medical and scientific communities. DataGenno is a portal (www.datagenno.com) with a search engine and a database, being developed to operate in the online health information field, providing medical and scientific information on genetic diseases and syndromes to health

Correspondence: Fabricio F Costa
Children's Memorial Research Center, Northwestern University's Feinberg School of Medicine 2430 N. Halsted, Box 220, Chicago, IL 60614-3394, USA
Tel +1 773 880 4000 ext 57312
Fax +1 773 755 6551
Email fcosta@childrensmemorial.org

45

care professionals, genetic counselors, research scientists and other interested professionals. The main advantage of our database is the ability to link clinical information from genetic diseases with the most recent molecular discoveries being published in the literature and in other databases. DataGenno will interconnect clinical information with molecular genetics using the Human Genome Browser,[2] the National Center for Biotechnology Information (NCBI),[3] and other portals as a source for genes, genetic testing information, new molecular discoveries, and genomic loci involved in each disease.

The main clinical features that have been described in the database so far are common signs and symptoms for genetic diseases and specific morphological features that are already established as standard in clinics. Molecular information will include the gene and network of genes associated with that specific disease, genomic loci of the gene(s) in the human genome, and other important molecular information. The description of the molecular features in this database will be of importance especially now that the cost of sequencing human genomes is decreasing, and several genomes will be sequenced in the next decade.[4,5] In that regard, it will be convenient to be able to combine genetic information with the phenotype that is observed in genetic diseases and syndromes.[6] DataGenno will also allow interaction and information exchange between health care professionals and scientists, helping to improve the diagnosis and discovery of new genetic diseases. Since DataGenno's main focus will be to provide information for health care professionals and scientists, laboratory tests and treatments for genetic diseases and syndromes will be also described in this database.

With emerging technologies and multiplying information sources, health care professionals and scientists need trustworthy, interactive, and contextual access to specialized information. DataGenno has the means to be the medical and health care professionals' first choice for information, tools and solutions for genetic diseases and syndromes. Thus, we believe that DataGenno will become an important source of information in biomedical sciences for professionals in health care, medicine and science.

Methods for construction and content
Database scheme

DataGenno's database has been developed after an extensive review of the scientific and technical literature including books on genetic diseases, genetic anomalies and birth defects, genetic syndromes as well as genetic tests that are already available. The database contains a list of genetic syndromes with clinical descriptions such as the presentation of the disease, natural history, etiology, genes and genetics information, signs and symptoms and a glossary with descriptions for technical terms. The system was developed using the PHP (a recursive acronym for Hypertext Preprocessor, originally Personal Home Page) tool,[7] and all data is stored in MySQL format[8] to manage the database content. The system follows the methodology of object orientation. Each object is defined by a class and represents a specific query. DataGenno's database uses the MVC (Model – View – Controller) model, which is widely used in web-based systems, because it facilitates communication between the browser (client) and the data stored in the database (server) (Figure 1).
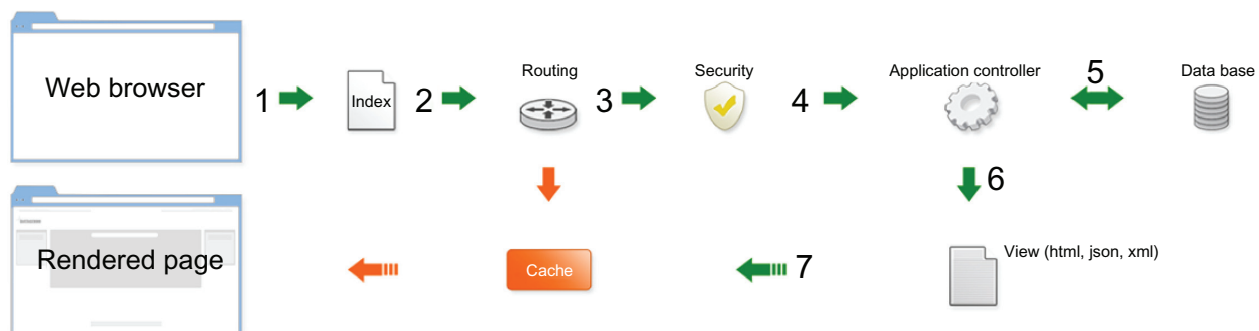

Figure 1 Schematic representation of DataGenno's computational interface between the website and the database.
Notes: 1) The index serves as the main driver, initializing the resources needed to run the system; 2) The router (Routing) examines the HTTP (Hypertext Transfer Protocol Overview) request to determine the action; 3) If there is something in the cache, it goes directly to the browser, bypassing the normal flow in the system; 4) The Application Controller, the HTTP request and any data submitted by the user will be filtered for security reasons; 5) The Application Controller loads the necessary resources required for the requisition, such as libraries to access the database; 6) After obtaining the data, the Application Controller generates an output according to the data format of the application; 7) The visualization of the data is finalized and sent to be generated by the browser. If the cache is enabled, the information is cached first, in order to be saved for subsequent applications.

Search engine

The database search engine uses the javascript library JQuery,[9] which is responsible for building the display panels and requesting data to the server using the AJAX technology.[10] JQuery is a fast and concise JavaScript Library that simplifies HTML (Hyper Text Markup Language) document handling and conversion.[9] Using the search engine, the user will be able to select one or more signs and/or symptoms, according to filters such as "genitourinary" and/or "low ear implantation" (or any of the approximately 6,000 signs and symptoms currently included in the database). The system will then select the syndrome(s) that have the sign(s) and/or symptom(s) that were chosen, allowing the user to have access to natural history, other signs and symptoms, genes that are involved in that specific disease, images and illustrations of patients, etc (Figures 2 and 3). The users will also be able to share different types of media such as pictures, videos and animations to make the information exchange a fully interactive experience.

Interactive platform

The database has an interactive platform in which the users (health care professionals, research scientists, and others) each have a personal web page describing their professional interests, field(s) of work, published articles and projects in progress, clinical studies being conducted by the professional, contact information, etc (Figure 4). The user will be able to register and sign up, sharing information about the patients and also the clinical studies that they are participating in with other users. Users will also be able to link with collaborators on a project, form discussion groups for specific diseases and share all types of information. All patient information in the database will have informed consent signed by the patient and provided by the health care professional and/or researcher according to the regulations in current use for each specific country. This interactive platform will enable the exchange of information and assistance in diagnosis of known genetic diseases and those that are still unknown or not yet described in the literature.

Figure 2 DataGenno's website and Down Syndrome.
Notes: The Control Panel in the Main Page of the website is indicated on the left. The Control Panel has the e-mail for each user and links to a glossary of medical and scientific terms, a list of signs and symptoms and a list of genetic diseases already described in the database in alphabetical order. Each genetic syndrome contained in the database has information about signs and symptoms, presentation of the morphological features, natural history of the disease, genes and genetics that include the genes and/or chromosomal regions that are implicated in each specific disease, and a picture with the morphological characteristics presented by the patients. The magnification on the right depicts a drawing for some specific features that patients with Down Syndrome have, such as phenotypic defects in the face, feet and hands. It is important to note that all diseases will have a different degree of phenotypic manifestation; however the pictures displayed in the database will contain the main characteristics for that specific genetic and/or chromosomal disease.
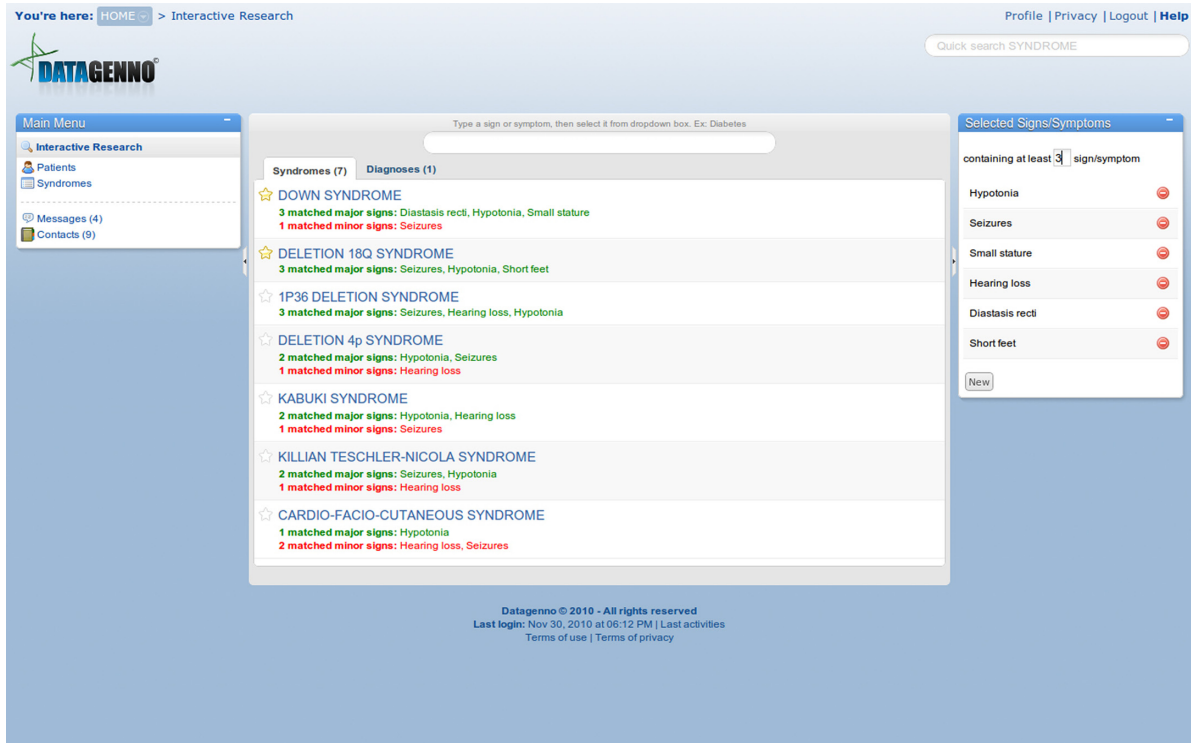
Figure 3 Interactive Research.
Notes: This panel represents the part of the database that contains the search engine for genetic syndromes. Using this search engine, the user will be able to search for a specific genetic syndrome in the database based on signs and symptoms.
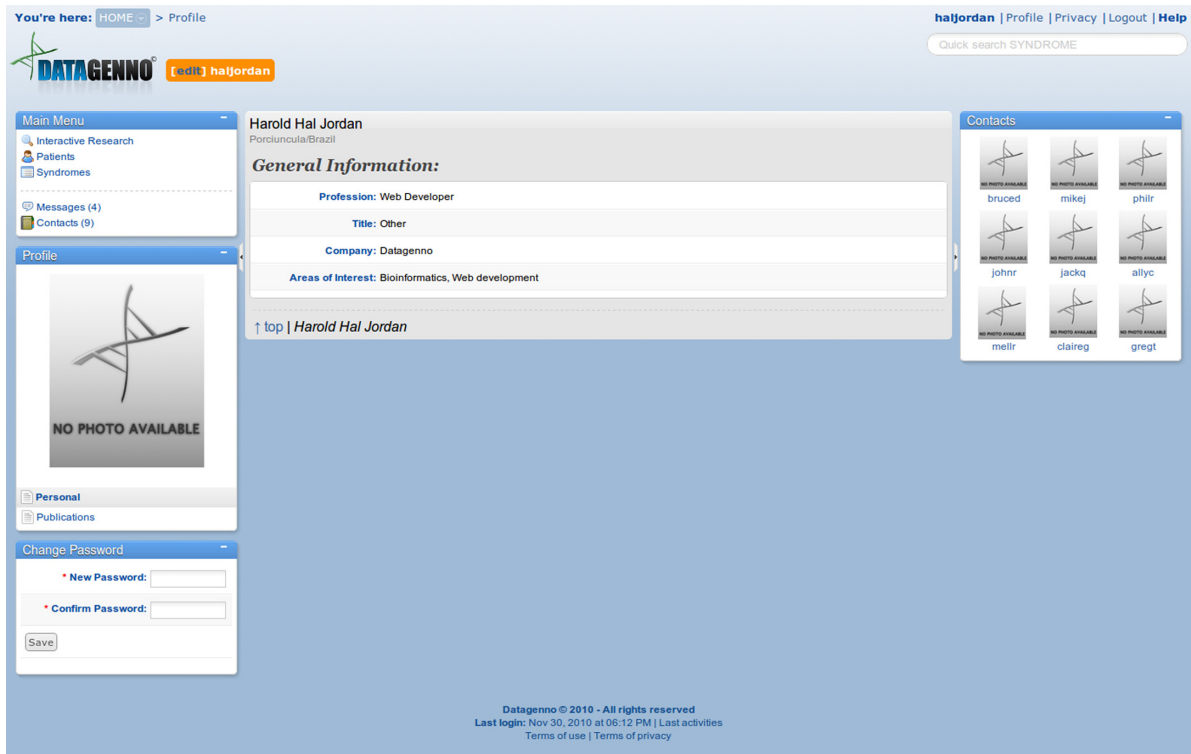

Figure 4 User profile web page.
Notes: This panel represents the area of the website where the user will be able to upload his or her profile picture, include personal information such as profession, title, company or institution and areas of interest. The user will be also able to include a list of publications and contacts to interact with other users signed up in the system.

Utility and discussion
Applications of the database and search engine

DataGenno's database and system will provide services for the medical, health care and scientific communities through its website (www.datagenno.com), where users will be able to: 1) have access to a database containing comprehensive clinical and molecular information about different genetic diseases and syndromes (see the Down Syndrome example in Figure 2); 2) have a professional profile, including job and research descriptions, in a personal web page, thereby facilitating interaction with other colleagues via multimedia tools. This feature will enable the downloading and/or uploading of clinical and research data as well as the exchange of data and information with other users signed up in the system (Figure 4); 3) create a list of signs and symptoms of patients with unknown genetic disorders which will allow professionals to perform computer searches between their patients' symptoms and those of others in order to find similarities and improve the ability to identify and describe new genetic diseases; and 4) have access to interactive 3D tools for educational and teaching purposes.

We believe that the features present in DataGenno's portal will contribute to the advancement of medical and scientific knowledge of genetic diseases and syndromes. One example of how a database such as DataGenno could be helpful in the biomedical field is the story of Dr Rienhoff, whose daughter had an unidentified genetic disorder.[11] Dr Rienhoff has several years of training in clinical genetics and has worked as a researcher in clinical and molecular genetics for decades. However, after many tests and dozens of medical interviews, there was no clear conclusion about his daughter's genetic condition until very recently.[11] With the advent of new DNA sequencing technologies and extensive collaboration with his peers, Dr Rienhoff has recently found evidence implicating a mutation in the gene for copine-1 (CPNE1) as the cause of his daughter's condition.[11] Like his daughter, about one-third of the patients that have morphological abnormalities and other symptoms leave the doctor's office or research/diagnostic laboratories with an uncertain diagnosis or no diagnosis at all. Unfortunately, the medical community often abandons cases of rare conditions since it is difficult to treat a patient without a diagnosis. This picture is changing since the National Institutes of Health (NIH) introduced a service named "The Office for Undiagnosed Diseases".[12] One successful case that was unraveled using this program was recently published with a description of a new genetic disease, which causes arterial calcifications and has severe consequences in patients' muscles, joints and heart.[13] This study was able to link clinical symptoms of individuals from different families with similar disease phenotypes to mutations in the gene NT5E.[13] A group of professionals have used genetic, genomic and molecular biology tools in association with clinical and radiographic analyses to reach a conclusive diagnosis in this case.[13] While this is an example of success, the NIH Program for rare conditions is very static and the professionals have no portal to exchange information; they are restricted to a website that relies on e-mail and phone calls to make connections. We believe that DataGenno will offer a better level of connectivity that will improve the diagnosis of orphan rare genetic diseases such as the one described above.

In addition to NIH, there are other public and private databases that provide information on genetic diseases, such as The National Organization for Rare Disorders (NORD),[14] WrongDiagnosis,[15] Revolution Health[16] etc. We believe the most important features for a database include interactivity, ease of implementation, user friendliness and content tailored for health care professionals and their professional needs. There is no database or system currently in existence on the Internet with these qualities, making our portal unique. It is clear that DataGenno will differ from other databases that are already available in four major ways: 1) DataGenno is targeted mainly at health care professionals, genetic counselors and research scientists; 2) it will display updated clinical and genetic information about genetic diseases (mostly monogenetic diseases and chromosomal abnormalities; in the future, complex diseases such as familial cases of cancer will be included); 3) while most of the available databases are focused only on clinical information, DataGenno provides an association of both clinical and genetic/genomic information for its users; 4) the combination of a database, a search engine and a website as a portal will provide a greater level of interactivity for each user and better support for clinical diagnosis and scientific research on genetic diseases.

Information exchange in the biomedical scientific community: developing a better tool

The Internet has emerged as a major communications media and has already fundamentally changed many sectors of the economy, including marketing and sales of financial services, travel and entertainment, among others. The Internet is also changing the health care industry and has transformed how consumers and physicians find and utilize health care information.[17] It has also become a primary source of information for physicians and researchers seeking to improve clinical practice and scientific studies.[18]

Thus, the Internet is changing the way professionals obtain information compared to traditional information sources, such as conferences, meetings, and printed journals. In addition, in the Web 2.0 era, open access scientific journals that have their contents published online and free of charge are becoming a reality.

Physician's offices, hospitals, and research institutions are becoming overwhelmed by the tasks and time involved in addressing the needs of patients. The gap between the patients' desire for information and the physicians'/scientists' ability to provide it in a timely manner is likely to increase. This gap will be filled by two related but distinct emerging options. One is a stand-alone Personal Health Record (PHR), such as the Internet-based tools for patients developed by Dossia[19] (a nonprofit consortium of major employers), Google Health,[20] Microsoft HealthVault,[21] WebMD,[22] health insurance plans and other Web services that are seeking expanded roles in the United States health care system.[23] Online repositories will allow patients to store, retrieve, manage, and share their health data – including lists of medical problems, medical history, medications, allergies, immunizations, test results, insurance information, and doctor's visits – over the Internet. The other option, the "integrated PHR",[24] an extension of the physicians' Electronic Health Records (EHRs), will go further and facilitate the type of physician–patient relationship that will improve health and health care at a lower cost. Integrated PHRs are already used by millions of patients, and their adoption is reaching a tipping point in some regions of the United States.[25]

DataGenno's website users will certainly benefit from the mass migration of the physician's EHR to the Internet. Our database will also benefit from the necessary standardization process that is currently in progress in the PHR field, which will make it much easier to exchange medical information. The use of EHRs and PHRs will finally drive the medical community's attention to the Internet, connecting health care providers, pharmaceutical and biotechnology companies, and millions of new consumers to the Online Medical Information community.

Considerable sums of money are being spent on a myriad of public and private databases intended to connect research and clinical data in order to improve our understanding of health and disease so that medical care can evolve. These databases include the London Dysmorphology Database,[26] the POSSUM-web database,[27] OrphaNet,[28] UpToDate,[29] and Online Mendelian Inheritance in Man (OMIM),[30] among others. The London Dysmorphology Database currently contains information on over 4450 dysmorphic, multiple congenital anomalies and mental retardation syndromes. It includes single-gene disorders, sporadic conditions, and those caused by environmental agents. The POSSUM-web Database is a computer-based system intended to help clinicians to diagnose syndromes in their patients. It contains information on more than 3000 syndromes, including multiple malformation chromosomal, metabolic conditions and skeletal dysplasias. OrphaNet is a database of information on rare diseases and orphan drugs for the public in general. Its purpose is to contribute to the improvement of the diagnosis, the care and treatment of patients with rare diseases. UpToDate is a website that provides information for health care professionals and their patients. The company's editorial board has more than 3,800 authors, editors, and peer reviewers that contribute by writing and continuously updating more than 7,400 topics addressing questions that arise in clinical practice in several diseases. Finally, OMIM, part of the National Center for Biotechnology Information (NCBI), is a comprehensive, authoritative, and timely compendium of human genes and genetic phenotypes. The full-text referenced overviews in OMIM contain information on all known Mendelian disorders and over 12,000 protein-coding genes. OMIM focuses on the relationship between phenotype and genotype. Even though the OMIM database is regularly updated, it has limited querying capacity. All of these databases and portals have similarities to DataGenno; however our database and system will provide better levels of interactivity by the development of personal profiles for the professionals that have an account in the website. Additionally, most of these portals and databases already available have incomplete or no genetic information about diseases. In this regard, we will also develop ways to interconnect the data from personal genomics and DNA sequencing information in our system to associate the clinical and molecular features of genetic diseases. As pointed out above, we believe that some characteristics present in DataGenno's portal differentiate us from the current tools available on the Internet and make DataGenno a valuable tool to study and analyze genetic diseases.

Although imprecise knowledge and uncertain data are an essential part of research, the clinical world requires more straightforward and reproducible information for better decision-making. These important issues are well-known to those who work at the interface of bioinformatics and medical informatics. Thus, close attention to how this

information is presented is essential, as researchers and clinicians typically have very different expectations when searching for information. Perhaps one of the real unmet challenges for existing databases is extracting data from basic research, securing conclusions and predictions to aid health care professionals in diagnosing a patient. Therefore, we believe that the next generation of databases and services providing reliable clinical and genetic information should remain highly focused on the needs of the medical and scientific community.[6]

Patient confidentiality is an important concern for existing clinical databases and also in the development of the EHRs. DataGenno will require informed consent from the physicians and other professionals dealing with their patients and/or samples (body fluids, tissue, cells, etc) before any information is included in our database. It is imperative that any patient and/or sample information that is included in DataGenno's website has approval from the patients and the institutions that are involved, in the case of research and clinical studies. We believe that this ethical aspect is important to facilitate data sharing and communication between professionals about their patients and samples in specific studies. Forms will be available in the personal profile of the professional according to the regulations and standards that have been used in the clinics. These forms will need to be completed in order to avoid using and depositing patient information without consent in DataGenno's database. The user will be responsible for any patient and sample information with consent forms that are uploaded in the portal. Key efforts designed to overcome these private information issues include the development of EHRs by different private companies such as General Electric Healthcare (GE Healthcare), and Microsoft HealthVault,[21] among others, together with the work of the Health Level Seven (HL7) organization.[31] DataGenno is prepared to follow this trend, offering a fully interactive database to health care professionals that will benefit from EHRs and standardization processes that are currently ongoing in the medical information community.

Applications for personal genetics and genomics

As discussed previously, EHRs have the potential to allow clinical integration of genetic and genomic medicine and improve the delivery of personalized health care, but more structured and standardized data elements and functions are needed.[32] While the advancement of genomic technologies after the human genome sequencing more than a decade ago is facilitating our understanding of human genetics,[33,34] for most clinicians, and some research scientists, the genomics era has not yet arrived. Thus, for genomics to have a profound impact in clinical practice, as it is having in scientific research, genomic literacy of health care providers and professionals is required.[32] We believe that DataGenno will be able to improve and facilitate genomics literacy for professionals who have little knowledge in this field.

The genomics field is entering in a new era. The cost of sequencing an entire genome is quickly decreasing and in a few years it will be possible to sequence entire human genomes at lower costs.[35] Several companies have been developing new DNA sequencing technologies with better resolution and accuracy.[36,37] However, big challenges lie ahead to understand the information behind the DNA sequences. These new advances will represent a breakthrough for medical science, and health care professionals will have the ability to scan the genome of a patient to detect susceptibility for diseases, drug metabolism defects, etc. There are already companies offering whole genome sequencing, analyses and data interpretation; examples include Knome[38] and Genomera.[39] There are also specific companies offering to scan and analyze parts of the human genome already known for susceptibility to diseases based on Genome Wide Association Studies (GWAS). These companies include 23andMe,[40] Navigenics,[41] Decode Genetics[42] and others. We will develop platforms and integrated graphical interfaces in DataGenno's system to include tools that will facilitate the visualization of entire human genomes and also to incorporate data generated from personal genome analyses. These features will include a search engine for specific genomic regions and genes in order to identify mutations, sequence variations, and sequence changes such as deletions or insertions, among other features. This part of DataGenno is still under development but we believe it will become the most important section of our portal.

Whole-genome sequencing is becoming a reality and we believe it will revolutionize medical diagnostics by enabling the identification of regions of the genome and/or genes that are responsible for diseases.[43] As new technologies continue to lower the cost of sequencing a genome, researchers and health professionals have begun to use large-scale genomic sequence data as a tool for novel discovery.[44] Genome sequencing is now accessible to researchers across a broad

range of disciplines and it will no longer be the sole domain of large research institutions. The relationship between disease phenotypes and their corresponding genetic changes is complicated, especially for diseases that have variations in clinical phenotype manifestation. A comprehensive diagnostic tool for genetic syndromes using whole-genome DNA sequencing will enable the identification of all possible genetic and/or genomic changes in each individual and their relatives, thus determining specific genetic defects that are responsible for underlying disorders. With the advent of new DNA sequencing technologies, whole-genome sequencing will become part of the clinical information relevant to disease diagnosis in the next five to ten years. In fact, recent studies utilized new technologies to sequence the DNA of individuals with rare genetic syndromes to gain insight of the genomic regions that could be responsible for the patient's phenotype. A patient with Charcot-Marie-Tooth Neuropathy had his entire genome sequenced and several genomic regions were identified as linked to the clinical phenotype, with most of them having an impact on the clinical manifestation of the disease.[45] Moreover, individuals with a Mendelian disorder with unknown causes – Miller Syndrome – had their genomes sequenced using different strategies to correlate genomic information to the observed phenotype.[46,47] The patients' relatives were also sequenced to determine which gene defects are present in their families.[46,47] Another group was also successful in applying whole-exome sequencing in the analyses of the Mendelian disorder Fowler Syndrome.[48]

These studies are just the beginning for clinical genomics and whole-genome sequencing for individuals with genetic disorders.[49] The molecular genetic cause of over 3,000 monogenic disorders is currently unknown,[50] and DataGenno's database will be extremely useful when more genetic and genomic data is available for many of these diseases. We believe that the data generated in these studies and from private companies can be used in our platform and interface to connect groups of professionals studying the same genetic diseases. Using DataGenno's tools, they will be able to exchange genomic data from different patients with unidentified diseases in order to uncover mutations and other genetic abnormalities that are shared by these patients. Thus, we believe that DataGenno will improve the knowledge in the clinical genetics and genomics fields by providing innovative tools designed to help establish a correlation between genotype and phenotype for genetic syndromes in a fully interactive web-based environment.

Conclusion

We believe that DataGenno's portal will help health care professionals, genetic counselors, scientists and other professionals interested in clinical genetics improve their knowledge and practice of medicine and science. The tools under development in DataGenno's portal will help researchers and scientists in the study of genetic diseases and syndromes that are still unknown or have no established clinical diagnosis. Finally, our database can revolutionize the way information about genetic diseases is organized and exchanged, especially now that whole-genomes can be sequenced faster and at a lower cost, adding a new dimension to the field of clinical genetics.

Acknowledgements

Disclosure

The authors declare that they have financial interests; FFC and MPC are research scientists and business partners at DataGenno Interactive Research Ltd. DataGenno has been sustained and funded by the partners' private resources and there is no government or academic funding to support it through to the present date. FFC is also the founder and CEO of Genomic Enterprise (www.genomicenterprise.com). A patent (Application Number 12/951,060) is pending in the USPTO for DataGenno's database and search engine.

Author contributions

FFC was responsible for drafting this manuscript, collecting and reviewing the molecular genetics and genomics information that has been inserted in DataGenno's database; LSF was responsible for data mining and revision of all information being incorporated in the database and MPC was responsible for collecting and revising the literature for clinical data on genetic diseases and syndromes.

Availability and requirements

The DataGenno website will be available free of charge at www.datagenno.com. There are no restrictions for use by

academic and private sectors. A group of institutions that are already using DataGenno as a tool include: 1) the Federal University of Rio de Janeiro (UFRJ); 2) the Medical Genetics Service of IPPMG (Pediatric Hospital); 3) the Neonatal Intensive Care Nicola Albano; 4) the Medical Genetics ER from Hospital Marcilio Dias (Navy Hospital); 5) the Genetics ER from Unifeso; 6) the Neonatal Intensive Care from Hospital Ferreira Machado of Campos dos Goytacazes; 7) the Medical Genetics Service of Campos dos Goytacazes; and 8) the Reference Center for Infants and Teenagers at Polo (CRTCA II) in Rio de Janeiro, Brazil. A group of international institutions that will test DataGenno's portal services in the future include: 1) The Department of Pediatrics at Children's Memorial Hospital and Northwestern University in Chicago, IL; 2) The Johns Hopkins University in Baltimore, MD and 3) The Department of Human Genetics at The University of Chicago in Chicago, IL. Institutions that want to be test sites for DataGenno can send an e-mail with inquiries to fcosta@datagenno.com or mcoutinho@datagenno.com.

References

1. Genetic Disease Information Available at: http://www.ornl.gov/sci/techresources/Human_Genome/medicine/assist.shtml. Accessed January 1, 2011.
2. Human Genome Browser Available at: http://genome.ucsc.edu/. Accessed January 1, 2011.
3. National Center for Biotechnology Information Available at: http://www.ncbi.nlm.nih.gov/. Accessed January 1, 2011.
4. Collins F. Has the revolution arrived? *Nature*. 2010;464:674–675.
5. Venter JC. Multiple personal genomes await. *Nature*. 2010;464:676–677.
6. Thorisson GA, Muilu J, Brookes AJ. Genotype-phenotype databases: challenges and solutions for the post-genomic era. *Nat Rev Genet*. 2009;10:9–18.
7. PHP Available at: http://php.net/manual/en/intro-whatis.php. Accessed January 1, 2011.
8. MySQL Available at: http://www.mysql.com/. Accessed January 1, 2011.
9. Jquery Available at: http://jquery.com/. Accessed January 1, 2011.
10. AJAX Available at: http://www.dmoz.org/Computers/Programming/Languages/JavaScript/AJAX/. Accessed January 1, 2011.
11. Maher B. Personal genomics: His daughter's DNA. *Nature*. 2007;449:773–776.
12. Undiagnosed Diseases Program Available at: http://rarediseases.info.nih.gov/Resources.aspx?PageID=31. Accessed January 1, 2011.
13. St. Hilaire C, Ziegler SG, Markello TC, et al. NT5E Mutations and Arterial Calcifications. *N Engl J Med*. 2011;364:432–442.
14. NORD – National Organization for rare Diseases Available at: http://www.rarediseases.org/search/rdbsearch.html. Accessed January 1, 2011.
15. Wrong Diagnosis Available at: http://www.wrongdiagnosis.com/. Accessed January 1, 2011.
16. Revolution Health Available at: http://www.revolutionhealth.com/. Accessed January 1, 2011.
17. Preliminary Estimates of Electronic Medical Record Use by Office-based Physicians: United States, 2008. Available at: http://www.cdc.gov/nchs/data/hestat/physicians08/physicians08.htm. Accessed January 1, 2011.
18. DesRoches CM, Campbell EG, Rao SR, et al. Electronic health records in ambulatory care – a national survey of physicians. *N Engl J Med*. 2008;359:50–60.
19. Dossia Available at: http://www.dossia.org/. Accessed January 1, 2011.
20. Google Health Available at: http://www.google.com/intl/en-US/health/about/. Accessed January 1, 2011.
21. Microsoft Health Vault Available at: http://www.healthvault.com/personal/index.aspx. Accessed January 1, 2011.
22. WebMD Available at: http://www.webmd.com/. Accessed January 1, 2011.
23. Steinbrook R. Personally controlled online health data – the next big thing in medical care? *N Engl J Med*. 2008;358:1653–1656.
24. Tang PC, Lee TH. Your doctor's office or the Internet? Two paths to personal health records. *N Engl J Med*. 2009;360:1276–1278.
25. The National Alliance for Health Information Technology. Report to the Office of the National Coordinator for Health Information Technology on Defining Key Health Information Technology Terms, 2008.
26. London Dysmorphology Database Available at: http://www.lmdata-bases.com/. Accessed January 1, 2011.
27. POSSUM Database Available at: http://www.possum.net.au/index.html. Accessed January 1, 2011.
28. OrphaNet Available at: http://www.orpha.net/consor/cgi-bin/Disease_Search_List.php?lng=EN. Accessed January 1, 2011.
29. UpToDate Available at: http://www.uptodate.com/home/index.html. Accessed January 1, 2011
30. Online Mendelian Inheritance in Man (OMIM) Available at: http://www.ncbi.nlm.nih.gov/omim. Accessed January 1, 2011.
31. HL7 Available at: http://www.service-architecture.com/xml/articles/health_level_seven_hl7.html. Accessed January 1, 2011.
32. Scheuner MT, de Vries H, Kim B, et al. Are electronic health records ready for genomic medicine? *Genet Med*. 2009;11:510–517.
33. Mardis ER. The impact of next generation sequencing technology on genetics. *Trends Genet*. 2008;24:133–141.
34. Kidd JM, Cooper GM, Donahue WF, et al. Mapping and sequencing of structural variation from eight human genomes. *Nature*. 2008;453:56–64.
35. Dolgin E. Talkin' 'bout my (third) generation. *Nat Med*. 2010;16:1053.
36. Branton D, Deamer DW, Marziali A, et al. The potential and challenges of nanopore sequencing. *Nat Biotechnol*. 2008;26:1146–1153.
37. Eid J, Fehr A, Gray J, et al. Real-time DNA sequencing from single polymerase molecules. *Science*. 2009;323:133–138.
38. Knome Available at: http://www.knome.com/. Accessed January 1, 2011.
39. Genomera Available at: http://genomera.com/. Accessed January 1, 2011.
40. 23andme Available at: https://www.23andme.com/. Accessed January 1, 2011.
41. Navigenics Available at: http://www.navigenics.com/. Accessed January 1, 2011.
42. Decode Genetics Available at: http://www.decode.com/. Accessed January 1, 2011.
43. Guttmacher AE, Porteous ME, McInerney JD. Educating health-care professionals about genetics and genomics. *Nat Rev Genet*. 2007;8:151–157.
44. Zaghloul NA, Katsanis N. Functional modules, mutational load and human genetic disease. *Trends Genet*. 2010;26:168–176.
45. Lupski JR, Reid JG, Gonzaga-Jauregui C, et al. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N Engl J Med*. 2010;362:1181–1191.
46. Ng SB, Buckingham KJ, Lee C, et al. Exome sequencing identifies the cause of a Mendelian disorder. *Nat Genet*. 2010;42:30–35.
47. Roach JC, Glusman G, Smit AF, et al. Analysis of Genetic Inheritance in a Family Quartet by Whole-Genome Sequencing. *Science*. 2010;328:636–639.

48. Lalonde E, Albrecht S, Ha KC, et al. Unexpected allelic heterogeneity and spectrum of mutations in Fowler syndrome revealed by next-generation exome sequencing. *Hum Mutat*. 2010;31:918–923.

49. Giallourakis C, Henson C, Reich M, et al. Disease gene discovery through integrative genomics. *Annu Rev Genomics Hum Genet*. 2005;6: 381–406.

50. Kuhlenbäumer G, Hullmann J, Appenzellerm S. Novel genomic techniques open new avenues in the analysis of monogenic disorders. *Hum Mutat*. 2010;32:144–151.