

Initiating a Human Variome Project Country Node

Jumana AlAama,¹ Timothy D. Smith,^{2,3,4*} Alan Lo,⁵ Heather Howard,² Alexandria A. Kline,² Matthew Lange,⁶ Jim Kaput,⁶ and Richard G.H. Cotton^{2,3}

¹Department of Genetic Medicine, Faculty of Medicine and the Princess Al Jawhara Center of Excellence in the Research of Hereditary Disorders, King Abdulaziz University, Jeddah, Saudi Arabia; ²Genomic Disorders Research Centre, Florey Neurosciences Institute, Victoria, Australia; ³Department of Medicine (St. Vincent's), University of Melbourne, Victoria, Australia; ⁴Department of Information Systems, University of Melbourne, Victoria, Australia; ⁵Victorian Partnership for Advance Computing, Carlton South, Victoria, Australia; ⁶Division of Personalized Nutrition and Medicine, National Center for Toxicological Research. U.S. FDA, Jefferson, Arkansas

For the HVP Bioinformatics Special Issue

Received 10 January 2011; accepted revised manuscript 25 January 2011.

Published online 29 March 2011 in Wiley Online Library (www.wiley.com/humanmutation). DOI 10.1002/humu.21463

ABSTRACT: Genetic diseases are a pressing global health problem that requires comprehensive access to basic clinical and genetic data to counter. The creation of regional and international databases that can be easily accessed by clinicians and diagnostic labs will greatly improve our ability to accurately diagnose and treat patients with genetic disorders. The Human Variome Project is currently working in conjunction with human genetics societies to achieve this by establishing systems to collect every mutation reported by a diagnostic laboratory, clinic, or research laboratory in a country and store these within a national repository, or *HVP Country Node*. Nodes have already been initiated in Australia, Belgium, China, Egypt, Malaysia, and Kuwait. Each is examining how to systematically collect and share genetic, clinical, and biochemical information in a country-specific manner that is sensitive to local ethical and cultural issues. This article gathers cases of genetic data collection within countries and takes recommendations from the global community to develop a procedure for countries wishing to establish their own collection system as part of the Human Variome Project. We hope this may lead to standard practices to facilitate global collection of data and allow efficient use in clinical practice, research and therapy. *Hum Mutat* 32:501–506, 2011. © 2011 Wiley-Liss, Inc.

KEY WORDS: variation; standard; data collection; human variome project

Background

The Human Variome Project is an international initiative committed to reducing the burden of genetic disease on the world's population by collecting and sharing data on all instances of genetic variation effecting human disease. The vision of the

*Correspondence to: Timothy D. Smith, Genomic Disorders Research Centre, University of Melbourne, Level 2, 161 Barry Street, Carlton South, Victoria 3053, Australia. E-mail: tim@variome.org

Contract grant sponsors: The Australian National Data Service (ANDS) through The Education Investment Fund (EIF) Super Science Initiative; The Australian Research Collaboration Service (ARCS) through The National Collaborative Research Infrastructure Strategy Program.

Human Variome Project is to be a catalyst for reduction in human disease in the 21st century by facilitating the establishment and maintenance of standards, systems, and infrastructure for the worldwide collection and sharing of genetic information.

Collecting information on all instances of human genetic variation as they are discovered is an ambitious prospect and one that has generated much discussion within the Human Variome Project Consortium. In addition to the obvious technical and organizational challenges of a data collection program of this scale, the nature of genetic variation information introduces a number of ethical, legal, and sociocultural challenges. To address these issues, the Human Variome Project Roadmap (available at <http://www.humanvariomeproject.org/>) proposes a two tiered method of collection via an integrated network of gene and disease specific databases and country based repositories or *HVP Country Nodes*.

HVP Country Nodes are repositories of genetic variation data that have been generated within each country hosting a Node. These repositories are built, managed, and resourced by individual countries to meet the ethical, legal, and information requirements of those individual countries. The information within each node is collected directly from laboratories and clinics within each country and stored safely and securely in a national repository. This model of collection allows diagnostic lab workers and clinicians access to a core set of information about each instance of genetic variation reported in their own country, while keeping access and storage managed at the local level. Access to such a repository will:

- provide faster and more accurate diagnosis of genetically based illnesses within the country's populations, reducing the cost and suffering of patients;
- help clinicians make more accurate prognoses and develop better treatment plans;
- improve the quality of genetic counseling for families and genetic testing; and
- improve national healthcare planning leading to reduced costs within national healthcare systems.

Ultimately, these repositories will become key sources of data for international gene and disease/specific databases by automatically passing on their data to the appropriate databases.

The HVP Country Node model provides a level of flexibility and modularization that would not be available in a centrally mandated system. As each individual country is responsible for the funding, collection, and storage of the data generated by their own country, they can ensure that it is handled according to their

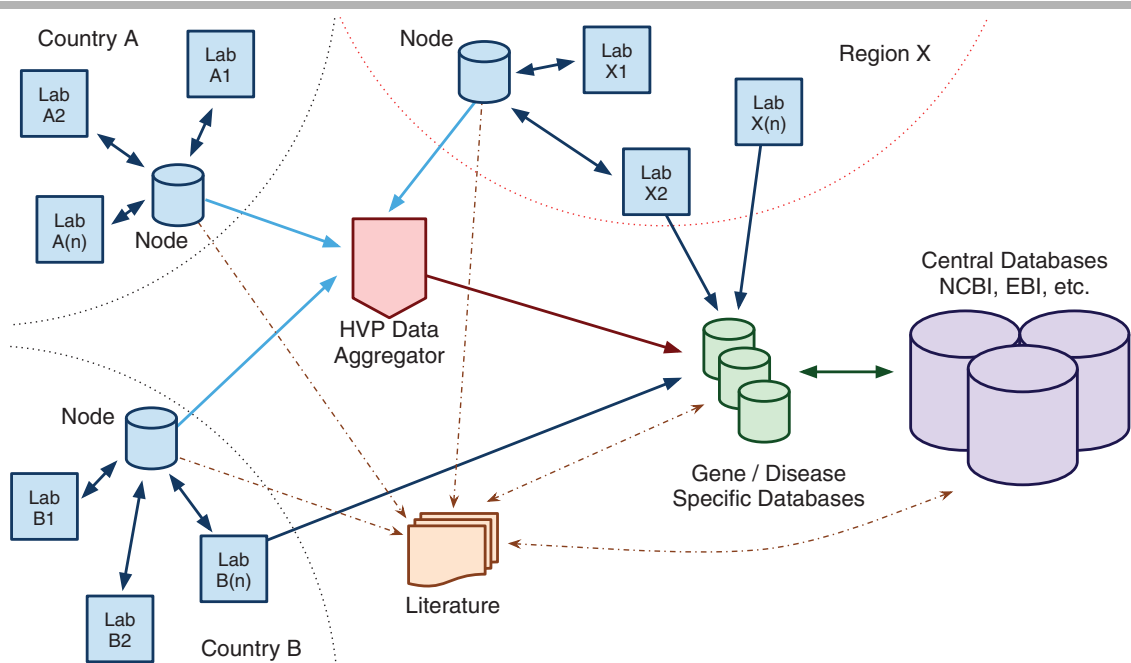


Figure 1. The proposed data collection architecture of the Human Variome Project. Institutions within specific countries generate genetic data and store these within national repositories or HVP Country Nodes (some countries in a geographically related area may wish to create regional nodes). These data are passed on to Gene or Disease Specific Databases through the HVP Data Aggregator. In these databases, the data are curated to a high standard before being passed on to the existing central databases for storage and reuse. Importantly, countries without a national repository are still able to contribute to the global collection effort with individual labs being able to submit data to the Gene or Disease Specific Databases directly. [Color figures can be viewed in the online issue, which is available at www.wiley.com/humanmutation.]

own nation's laws and in an ethical and culturally sensitive manner. Countries that cannot afford a node of their own may choose to partner with other countries in the creation of regional nodes that can be used by multiple country affiliates.

Figure 1 shows how HVP Country Nodes fit into the global collection architecture proposed by the Human Variome Project. In this example, Region X is a partnership of multiple countries, none of whom would be able to host an HVP Country node by themselves. Some laboratories within a country share data with only country/regional nodes: A(1)-A(n), B(1), B(2), X(1); some laboratories share data with both country/regional nodes and gene/disease specific databases B(n), X(2); and one laboratory shares only with disease specific databases: X(n). In addition to acting as a local repository for country specific data, HVP Country Nodes will share core data elements with global gene or disease specific databases for use around the world. This sharing will be mediated by the proposed HVP Data Aggregator.

This article is divided into two sections. The first section introduces a number of existing and planned systems from various countries. Although not all of these systems can be considered HVP Country Nodes, they serve to illustrate how such repositories can function. The second part of the article details the process a country can use to develop a new HVP Country Node and illustrates that process with examples and experiences from the Australian Node of the Human Variome Project.

Existing Systems

A number of countries have now established nationwide efforts in collecting genetic variants [see Cotton et al., 2009]. There has also been a review of the earlier activities [Patrinos, 2006]. Each of these accomplishments is a success in demonstrating the feasibility of the Human Variome Project model of collection. Although

infrastructure is an obvious common challenge, experiences reported from these countries also include topics such as the need for communities, education, government involvement, and sustainability. These experiences will assist other countries to develop their own nationwide databases.

The Korean Mutation Database (KMD) was developed with an emphasis for clinicians and researchers. KMD features four major components: advisory committees, mutation data collection, Web-enabled database, and the upgrades that allow connectivity with other databases. Governed by the Korean National Institute of Health (KNIH), KMD also has an advisory committee consisting of both diagnostic and genetics researchers. Data are collected from clinics, four major university hospitals, and also individual researchers. KMD's scope includes all diseases and collects minimal clinical information (disease and gender). The Web-enabled interface allows both the searching and submission of data. The main proponents for success in KMD are (1) obtaining up to date information, (2) working in conjunction with genetic societies, (3) funding from its national government, (4) creation of active advisory committees, and (5) frequent system upgrades. KMD can be found at <http://kmd.cdc.go.kr/>.

In China, the Center for Genetic and Genomic medicine has begun data collection using the Leiden Open Variation Database (LOVD; <http://www.lovd.nl/>) [Fokkema et al., 2005]. The Chinese Node has connected the data collection effort with healthcare, basic research, world experts, and industry partners. An initial consortium of 11 institutions in six provinces in China has already created substantial data for research. This effort will compliment the new investments in genetic sequencing hardware by that nation's healthcare research system. Student volunteers helped make this achievement possible in a short amount of time: an option other countries may consider.

Orphan Net Japan (<http://onj.jp/>) is a nationwide effort designed to create sustainable support for testing rare genetic

diseases. Much like other countries, Orphan Net Japan has an established network infrastructure with a mutation database for the collection of data. To ensure the sustainability of the project Orphan Net was established as a not-for-profit organization, providing a fee for service approach for hospitals in testing of rare diseases. Education is an important component of the project, spanning areas from genetic testing through to patient consultation. Orphan Net Japan also uses this platform to create quality control and standardization for genetic testing in Japan.

In the United Kingdom, the Diagnostic mutation Database, DMuDb (<http://www.ngrl.org.uk/Manchester/dmudb.html>), was created specifically to collect data from diagnostic labs rather than just from the literature. These data should be of high quality for use in guiding clinical decisions and are collected with proprietary software.

In the Arab world, the Consortium of Arab Genetics Societies (CAGS) has established a listing of diseases found in their populations. To date, mutations are not yet being systematically collected (unpublished). Dr. Al Sayali has recently documented the mutations in the United Arab Emirates [Al-Gazali and Ali, 2010] but has not yet established a database. Other local and regional efforts to create clinically relevant genetic databases may be occurring, but these have not yet been linked to the HVP.

These successful outcomes demonstrate the possibility of creating nationwide genetic variation collections despite the local challenges involved. These models may inform other countries interested in constructing HVP Country Nodes.

Proposed Systems

Malaysia, Kuwait, Saudi Arabia, South Africa, Brazil, and Lithuania are beginning the organization of data collection methodologies, building momentum for the global initiative.

For example, a consortium of academicians, clinicians, researchers and other professionals in Malaysia have been established. This consortium is creating and applying genomic technology and informatics in life sciences, forensics, ethno-archaeology, as well as improving the social well-being of Malaysians in the process. The Malaysian Node was officially launched on October 9, 2010.

The Kuwait Medical Genetics Centre (KMGC) has in a period of 30 years worked with over 32,000 families diagnosed with genetic disorders. With this vast amount of data, Kuwait faces unique challenges in developing an HVP Country Node and participating in the international effort.

The Saudi Arabian Consortium has accomplished a considerable amount of ground work for creating an HVP Country Node. Consortium members have networked with academic and clinical centres to define existing services and created plans for future capacity development. The Saudi initiative placed an emphasis on creating a DNA bank linked to clinical phenotypes. During this process, a number of obstacles were identified such as the cost for diagnostic laboratories to participate and the difficulties of communication. Because the laboratory capabilities and infrastructure are not yet in place through the Kingdom, samples may have to be sent abroad for testing for certain mutations. Country-specific laws and customs may have to be addressed, a challenge other countries may have in common with Saudi Arabia.

The development of the Human Variome Project will help create incentives for addressing these challenges and obstacles. Databases and interfaces that are easy to use for data collection with some type of reward and recognition for contributors may increase participation of clinical laboratories; funding, and

infrastructure will also support these processes. Organization, cooperation, and trust are also required to build a country node: this can be achieved by initiating a national committee of representatives and subcommittees of diseases, medical centers, and/or specialties. Education is also an important factor: the Saudi Arabian effort has led the way on this, with material presented via their website, local newspaper, and television interviews.

The Australian Node of the Human Variome Project

The first official HVP Country Node was established in Australia in 2009 via funding provided by the Australian federal government. The project is currently in the process of installing and testing data collection systems in diagnostic laboratories that will facilitate the automated collection of genetic data from those laboratories and the secure transmission of data to a central national repository. The initial phase of the project is projected to be completed in mid-2011. The Australian Human Variome Database will be made accessible to researchers, clinicians, and diagnostic labs throughout the country, and technologies developed by the project will be exportable to other countries.

These examples of country-specific efforts have arisen in the last 5 years and may provide models for countries with similar infrastructures, cultures, and conditions. Systematic efforts within countries that are linked to the worldwide initiative are needed to efficiently collect and deliver data to those caring for families with inherited disease.

The following represents a distillation of suggestions, guidelines, and recommendations developed by the Human Variome for establishing genetic data collection within a country.

Initiating an HVP Country Node

What is Needed to Establish a HVP Country Node?

Establishing an HVP Country Node requires the creation of a central national database into which genetic variations reported by diagnostic labs within a country can be submitted and linked to deidentified information concerning each variant's consequences and predicted pathogenicity. Ideally, systems to allow the automatic deposition of these data by diagnostic labs would also be developed and introduced into diagnostic labs and clinics throughout each country through standard and secure Web technologies. Creating a country node requires support from diverse stakeholders: from patient advocacy groups through local and national government agencies.

The first crucial step in setting up an HVP Country Node is the creation of a Country Node Consortium to assist in garnering political support and obtaining funding (see Box 1).

Consortium Formation

The Human Variome Project was enlarged by local champions, geneticists who recognized the need for creating country-specific nodes. In almost all cases, the process was initiated through a regional or national meeting of clinical geneticists and representatives from diagnostic laboratories. Perhaps most importantly, an HVP Country Node should be formed in association with the respective local human genetics society. In most cases, local genetic societies are listed at the International Federation of Human Genetics Societies Website (<http://www.ifhgs.org/>).

Developing funding is crucial, and key stakeholders to identify are representatives of local governmental and funding bodies who should be engaged at an early stage in the process. The Human Variome Project Coordinating Office can assist in the process by providing support letters and other related services.

Box 1. Strategy for the Formation of an HVP Country Node

Administrative/Funding

Network with clinical and diagnostic geneticists, bioinformaticians, and other stakeholders within the country to create a consortium of interested people to support the project

Create a working group from the members of the consortium

Employ a project manager to oversee the running of the project

Enlist the support of key government representatives

Develop a data access policy consistent with the ethical and legal requirements of your country

Elect or create a permanent committee or board to ensure that the database is maintained to the highest standards and operated in a fair, legal, ethical, and transparent manner.

Technical

Evaluate the technology and systems that could best be employed to collect and store genetic and clinical data

Develop/implement software systems to ensure smooth collection and linking of data from diagnostic labs and clinics. These systems should be capable of evolving and being continuously updated with new content. The Australian Node is developing a model with the specific intent of making it available for use by other country nodes.

Decide on a data model

Practical Considerations

What will be the process for collecting this information? (i.e., what model of data flow will be used from lab/clinic to the node?)

Identify the organizations that may play a collaborative role to leverage the resources, systems and capabilities already in existence

Automate as many data submission (or extraction) steps as possible

Ensure interoperability of systems locally and globally.

Develop methods for integrating genetic variant data from laboratories or clinics with phenotypic data

Legal/ Ethical Considerations

Identify and work with agencies or organizations that provide nationally recognized ethical approval for the collection and sharing of genetic data

Will ethical oversight and/or agreements be needed for each point of collection?

What security measures will be in place to protect patient privacy?

Is it necessary to obtain patient consent?

Individual countries should decide on the governance structure best suited to their requirements and local conditions. For example, Australia created one committee with representatives from each of their six states. Another approach would be a central committee with subcommittees representing individual cities, states or provinces, with each sending a representative to the central body. In some instances, individual countries may not have the resources for an HVP node. In these cases, a geographically located hub may be established for several countries within a region. Sharing costs and resources regionally may allow development of country-specific nodes in the future. The Human Variome Project International Confederation of Countries Advisory Council was formed to assist with these efforts.

Join the International Community

An HVP Country Node is best formed with advice and guidance from the Human Variome Project Coordinating Office. Once established, each node can then apply for membership of the International Confederation of Countries Advisory Council using a simple application process. The Human Variome Project has established procedures to assess and approve membership. Each member node is issued a Warrant of Partnership. This warrant allows the country to officially use the Human Variome Project

logo and to call their repository a node of the Human Variome Project. A policy document outlining the application process can be found on the Human Variome Project Website at <http://www.humanvariomeproject.org/>.

Current interim members of the Human Variome Project Confederation of Countries Advisory Council are: China, Kuwait, Malaysia, Australia, Egypt, and Belgium.

Decide on Data to Collect

Deciding on data fields to collect is not a simple matter. A new node should draw input from local representatives of genetic clinicians and diagnostic laboratories. The data model must reflect data that is feasible to collect in a country based on factors such as geographical limitations, local practices and procedures, and the capacity of contributing organizations. Each node must ensure that the data will be collected by appropriate parties to safeguard the quality of the information. Different countries also have different levels of privacy laws and ethical obligations that must be considered in the data collection process. The Human Variome Project suggests that informed consent be collected from each patient before their data is stored, although this may not always be practicable. It should be noted however, that some researchers or clinicians in other countries may not be able to use data obtained without informed consent. Audit trails for data acquisition and access, and for tracking and validating consent, should be instituted before patient information is collected.

Keeping all these factors in mind, the following baseline fields should be collected in most situations to ensure the utmost utility of the resulting repository:

- Gene Name—described in the form of both the HUGO Nomenclature Committee approved gene name and a sequence accession number and version number.
- Variant Name—written as HGVS nomenclature (<http://www.hgvs.org/mutnomen/>).
- Pathogenicity—classified as five levels of pathogenicity [see Plon et al., 2008].
- Test date—the date that the results were produced.
- Patient ID—a deidentified code which is unique to a patient.
- Patient Age—the age of the patient when tested.
- Patient Gender.
- Submission date.
- Disease associated with the mutation—if diagnosed.
- Lab Operator ID—a code that identifies the operator who uploaded the data.
- Laboratory Name/ID.
- Country/Region Name/ID—if a regional repository is used.
- Level of consent obtained.
- Can the patient be recontacted for other studies?
- Can clinical and/or molecular data be used for statistical analyses (with options for local laboratory, country, and/or international)?

Details of the test such as the method used, sample type, test range, sensitivity, and accuracy, are beneficial but are optional. Information about the laboratory and operators should be submitted to the country node in order to ensure data provenance and the ability to trace back information to the source of the submission. When the data is submitted from the country/region node to the HVP data aggregator, metadata about the country node should be submitted as well.

Clinical data, family history data, segregation analysis, and other disease-specific information are also important but may be

limited by legal and ethical considerations. Other more detailed descriptors may be more difficult to obtain, catalog, or store because of practical considerations of who, when, how, and cost of effort of the data collected. In most countries, it is anticipated that access to detailed clinical data may not be possible due to reidentification concerns. In such circumstances, records in the node repository should be linked to clinical data maintained in that country's mandated clinical data repository (e.g., an electronic health record, disease registry, etc.) to ensure that clinicians and researchers using the node still have access to this data. When data from such a node is shared with an international gene/disease specific databases, summary information such as SISA (in press), should then be included.

Systems Building

Broadly speaking, a country node consists of three components: the physical data collection systems, the Web-enabled tools for uploading data, and the underlying database.

The data collection systems need to be simple and easy to use. Where possible, they should interface with existing laboratory information management systems to automate the collection of data, reducing the effort required for submission. Each laboratory should maintain full control over the collection process in terms of what fields will be sent on to a node, subject to necessary minimal reporting standards. This gives the system flexibility to negotiate with organizations on their level of participation. The system will therefore need to:

- Integrate data sources from other systems (sequence processing software, laboratory information management systems) in a laboratory setting to automatically or semiautomatically identify and collect the data of interest.
- Provide a simple user interface for the medical scientist who collects any remaining fields that could not be collected automatically (if any) for each record; for example, prompting the scientist for an assessment of the pathogenicity of a variant.
- Be configurable for the data custodian at a laboratory to control precisely what fields will be submitted to the Node versus fields that will remain within the laboratory. For example, permitting the variant name to be submitted but omitting a patient's name for privacy reasons.

In Australia, the MAWSON project (<http://www.mawsonproject.com/>) is being utilized for these data collection functions.

Transportation of data over the Internet between a laboratory and a central repository may be compromised and exposed to malicious online activities. The transport mechanism, which allows data collected from laboratories to be deposited to the Country Node, must consider the issue of IT security. Encryption protocols will be mandatory to transfer the data. For example:

- Using Hypertext Transfer Protocol Secure (HTTPS) communication.
- Using Virtual Private Network (VPN) technology.
- Using Internet Protocol Security (IPSec) technology.

This will help ensure that data submitted to the Node is protected from theft or tampering.

For the protection of patient privacy, a patient's identity will need to be hidden within the Node repository. On the Country Node database, a patient should be codified and unidentifiable by users. This ethical concern will generally be supported by legal

implications in most countries. Techniques for ensuring adequate deidentification of patient data include:

- Using nonreversible cryptographic hashing algorithms that codify a patient's identity using the patient details. This allows information from two sources about the same patient to be linked without revealing the patient identity.
- Using a simple unique incremental number—this method does not allow information from multiple sources to be reliably linked.
- The Node itself needs to be a data storage facility where data is submitted utilizing the data model identified as appropriate for the country. It should include:
- Database Management System—hosts the country node data model and provides facilities for management of the database instance such as backup and revival.
- User registration—to allow users within the country to access the data provided they qualify and accept the countries terms and conditions of usage.
- Basic data access facilities—to allow users to browse the data.
- Geographic collocation and disaster recovery plans in the event that the physical structure of the country node data center is compromised.
- Limited and auditable access to the physical hardware and network infrastructure housing the country node.

Roll Out and Uptake

Once a system of collection has been identified and the necessary systems and organizational structures to support them have been established, then the process of rolling the system out to the laboratories and clinics within the country doing genetic testing can begin. The smooth integration into the existing systems and workflows of these sites will require all the stakeholders within the organizations to be both informed and accepting of the underlying principles and purpose of the Country Node. Detailed management and education planning should be done well in advance of the initiation of roll out. It is strongly suggested that all stakeholders be involved as much as possible at the beginning of the process as possible so that all parties have a strong understanding of what is required.

Link to International Repositories

It will be necessary for HVP Country Nodes to establish links with the Gene/Disease Specific Databases to share their data with the international community. The Human Variome Project's International Confederation of Countries Advisory Council will work together with the Project's Gene/Disease Specific Database Advisory Council to provide guidelines and standards for the secure transfer of data. The volume of data generated by the HVP Country Nodes will become, over time, quite large, and looking toward the future, an increasing amount of redundant data could potentially be collected, especially relating to variations that are already well defined (e.g., CFTR:p.Phe508del). In these cases, the curators of the relevant gene/disease specific databases will need to determine in what circumstances such data should be submitted to their databases. But such data should still be collected in the HVP Country Nodes in order to provide an accurate picture of the genetic component of each population's disease burden.

The overall Human Variome Project data collection architecture (Fig. 1) provides for the transfer of data from node to gene/disease specific database to central databases (and other interested

repositories). This is to ensure universal access to any interested researchers and to provide safe and stable repositories into the future.

Summary

The concept of HVP Country Nodes evolved from discussions within the Human Variome Project Consortium as the best way to collect sufficient molecular and clinical information on instances of genetic variation in a manner that is compliant with the diverse legal and ethical landscape of member countries. By engaging local consortiums to manage the collection, management, storage, and transfer of genetic data, the control and ownership of that data is maintained at the local level and individual Nodes can decide what information is collected and the manner in which the collection is conducted.

The HVP Country Node concept will facilitate the worldwide sharing of genetic variation information. As well as becoming a vital source of data for the highly curated gene and disease specific databases, these uncurated repositories of diagnostic and clinical test reports will provide both a service for the diagnostic labs within host countries as well as a valuable platform for medical research and healthcare planning and delivery.

Although this article has outlined a way forward for those wishing to initiate an HVP Country Node, it is by no means meant to be a prescriptive set of methods; rather, this article seeks to explain the rationale behind the HVP Country Node concept and provide a single method of how that concept can be practically achieved. The Human Variome Project Consortium is

keen to encourage participation in this program of work from any country, and the authors invite those wishing to contribute to contact them.

Acknowledgments

This work is part of a National eResearch Architecture Taskforce (NeAT) project, supported by the Australian National Data Service (ANDS) through the Education Investment Fund (EIF) Super Science Initiative, and the Australian Research Collaboration Service (ARCS) through the National Collaborative Research Infrastructure Strategy Program. Note: The views in this article do not necessarily represent those of the U.S. FDA.

References

- Al-Gazali L, Ali BR. 2010. Mutations of a country: a mutation review of single gene disorders in the United Arab Emirates (UAE). *Hum Mutat* 31(5):505–520.
- Cotton RGH, Al Aqeel AI, Al-Mulla F, Carrera P, Claustres M, Ekong R, Hyland VJ, Macrae FA, Marafie MJ, Paalman MH, Patrinos GP, Qi M, Ramesar RS, Scott RJ, Sijmons RH, Sobrido M, Vihinen M. 2009. Capturing all disease-causing mutations for clinical and research use: Toward an effortless system for the Human Variome Project. *Genet Med* 11(12):843–849.
- Fokkema IF, den Dunnen JT, Taschner PE. 2005. LOVD: Easy creation of a locus-specific sequence variation database using an “LSDB-in-a-box” approach. *Hum Mutat* 26(2):63–68.
- Patrinos GP. 2006. National and ethnic mutation databases: recording populations’ genography. *Hum Mutat* 27(9):879–887.
- Plon SE, Eccles DM, Easton D, Foulkes WD, Genuardi M, Greenblatt MS, Hogervorst FB, Hoogerbrugge N, Spurdle AB, Tavtigian SV, for the IARC Unclassified Genetic Variants Working Group. 2008. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum Mutat* 29(11):1282–1291.