THE HUMAN VARIOME PROJECT

sharing data • reducing disease

an NGO official partner of UNESCO

Coordinating Office Position Paper

AP07: Standards for disease and phenotype descriptions in gene variant databases (LSDBs)
ICO Comments

PP-ICO-11-2013

2013-11-12

Human Variome Project International Ltd.
Level 5, 234 Queensberry Street
University of Melbourne VIC 3010
AUSTRALIA

Prepared by:
Timothy D. Smith

Contact:
Timothy D. Smith
Level 5, 234 Queensberry Street
University of Melbourne VIC 3010
AUSTRALIA
+61 (0)3 8344 2096
tim@variome.org

Authorisation:

This Position Paper has been prepared by Timothy D. Smith and represents the official position of the Human Variome Project Coordinating Office only. It does not represent an official position of the Human Variome Project, its Consortium, Advisory Councils or International Scientific Advisory Committee.

-s-

Heather Howard

I. Context

The International Coordinating Office (ICO) received the attached Activity Proposal (AP) on the 3rd of September, 2013, from Human Variome Project Consortium member Peter Taschner. Pursuant to the Project's Standards Development Process (PD06-2011) it was forwarded it to the International Scientific Advisory Committee (ISAC) for referral to a Sponsoring Council (SC). Further pursuant to the Standards Development Process we have prepared the following comments on the AP for the SC.

II. Comments

Background

The Human Variome Project exists to aid the continual improvement of genetic variation knowledge resources and to facilitate their use in and integration into routine clinical practice. A core component of this work must be a focus on improving the utility and interoperability of these resources and the ways in which they describe clinical features, phenotypes, diseases and disorders. This proposed activity seeks to provide a set of minimum content requirements for:

1. the description of clinical features and observed phenotypes that prompted genetic testing in an individual patient; and
2. the description of concluded diseases and/or phenotypes following genetic analysis.

Such descriptions of diseases and phenotypes are vital to ensuring that genetic variation knowledge resources are useful in the clinical context. It is also vital that these descriptions be standardised across resources to allow data exchange, distributed querying and improved understanding across clinical specialties.

The proposer notes that the curators of genetic variation knowledge resources would like to state that they follow the HVP recommendations regarding content and display. This is a noble ambition and the development of such standards by the HVP, as the desire implies, is an activity that the Human Variome Project ICO strongly supports.

The proposed activity strongly complements existing work that the Project Consortium is undertaking.

Harmonisation of Ontologies

The Human Variome Project ICO recently supported a meeting of several groups involved in the development of ontologies and terminologies for human phenotypes. Groups included:

- PhenoDB/OMIM
- Human Phenotype Ontology
- Elements of Morphology
- Orphanet
- Possum
- NCBI

At this meeting, the groups decided on a set of ~2,000 top-level terms for describing human phenotypes that will be used across the various ontologies and nomenclatures represented. The ICO strongly believes that the outcomes of this activity should be integrated into the proposed Human Variome Project activity.

WG03: Minimal content for gene variant databases (LSDBs)

The Human Variome Project has a Working Group (Chair: Peter Taschner, Sponsoring Council: ISAC) that is developing standards for minimal content for genetic variation databases, including standards for the description of "associated phenotype."

As describing phenotype is a considerable activity that involves a degree of specialised expertise, it probably does merit the spinning out into an activity in its own right. If the Sponsoring Council for this proposed activity decides to move this activity forward, some thought will need to be given as to how to best coordinate the work between the two Working Groups.

Scope of Work

The ICO believes that the scope of work outlined in the Activity Proposal is necessary and sufficient to achieve the stated objectives of the proposed activity.

The minimum scope should cover:

- data fields to be included describing the disease(s) and/or phenotype(s) detected leading to the genetic analysis of the individuals analysed
- data fields to be included describing the concluded disease(s) and/or phenotype(s) following to the genetic analysis of the individuals analysed
- standards to follow for data entry & storage

The Sponsoring Council should decide if the following work is in or out of scope:

- the method(s) and/or equipment used to identify a specific phenotype or trait
- minimal query options that should be offered
- additional recommended data fields

Possible Working Group Members

The ICO strongly believes that the individuals involved in the existing work of harmonising ontologies should be included in any Working Group chartered to undertake the proposed activity. Further, the ICO will, as always, remind the Sponsoring Council that the Human Variome Project is an international organisation and to ensure that any Working Group be as representative of the Consortium as possible.

Available Assistance from the ICO

The ICO is able to provide the following assistance to the WG and the ensuing Interest Group:

- secretarial assistance;
- face-to-face meeting organisation and venues;

- teleconferencing facilities;
- an email list/discussion group.

The ICO is currently investigating a number of options for creating online collaborative workspaces for Working Groups. We anticipate being able to offer such functionality early in 2014.

III. Summary of Recommendations

We recommend that:

- The activity being proposed be taken up by the Project
- The GDSDBAC should decide how best to proceed with this activity:
    - If the work being proposed necessitates a new Working Group
        - The GDSDBAC should charter a new Working Group to undertake the scope of work outlined in the Activity Proposal;
    - If the work would be better conducted by the existing Working Group (WG03: Minimal content for gene variant databases)
        - The GDSDBAC should advise the Sponsoring Council of WG03 of this decision
- The Sponsoring Council should decide if the additional work outlined in the Activity Proposal is in scope, or should be the subject of further work by another Working Group
- The resultant work be published as an HVP Standard

ATTACHMENT A
Activity Proposal

Standards for disease and phenotype descriptions in gene variant databases (LSDBs)

Need
Standardization of disease and phenotype descriptions in gene variant databases (LSDBs) is highly desired to improve database interoperability and database quality. To indicate their quality, curators of gene variant databases (LSDBs) would like to state that they follow the HVP/HGVS recommendations regarding content and display of disease and phenotype descriptions. However, such recommendations do not yet exist.
Existing recommendations would also help gene variant database software developers, ensuring that their design includes all minimal requirements for adequate storage and display of disease and phenotype details.

Scope

The "***Standards for disease and phenotype descriptions in gene variant databases (LSDBs)***" should cover at least;
 data fields to be included describing the disease(s) and/or phenotype(s) detected leading to the genetic analysis of the individuals analysed
 data fields to be included describing the concluded disease(s) and/or phenotype(s) following to the genetic analysis of the individuals analysed
 standards to follow for data entry & storage

It should be discussed whether the recommendations should be extended to include;
 the method(s) and/or equipment used to identify a specific phenotype or trait
 minimal query options that should be offered
 additional recommended data fields

Since disease and phenotype descriptions are likely to become part of the minimal requirements, it seems appropriate to come to a set of "minimal requirements" and a set of "highly desired requirements", allowing existing databases to cope with the new demands. Databases can then choose to follow either the minimal or extended requirements.

Plan of action
Assign a group of experts (3-5) that write a draft of the document suggested under "Expected deliverables", give HGVS/HVP members a chance to comment on the proposal and finally complete the document and submit it for publication.

Resources required
Time, administrative support from HVP, a room to meet & discuss during an upcoming meeting (first within the group of experts, next with HGVS/HVP members).

Expected deliverables
A document describing the "***Standards for disease and phenotype descriptions in gene variant databases (LSDBs)***". Preferably the document will be published in a peer-reviewed scientific journal.

Recommendation
The "***Standards for disease and phenotype descriptions in gene variant databases (LSDBs)***" should become the (minimal) standard for gene variant databases and allow them to state that they follow the HVP recommendations.