Date: AugustFebruary 2009

Phenotype and Genotype Object Model (PAGE-OM)

Beta 32 - without changebars

OMG Document Number: dtc/2009-08-01dtc/2009-02-04
Standard document URL: http://www.omg.org/spec/PAGE-OM/1.0/Beta2
Associated File(s)*: http://www.omg.org/spec/PAGE-OM/20080101
http://www.omg.org/spec/PAGE-OM/20080102
http://www.omg.org/spec/PAGE-OM/20080103

* original files: lifesci/2008-01-03 (PIM-normative), lifesci/2008-01-04 (PSM-normative), lifesci/2008-01-02 (normative and non-normative)

This OMG document replaces the submission document (dtc/200908-0924-0943, Beta). It is an OMG Adopted Beta Specification and is currently in the finalization phase. Comments on the content of this document are welcome, and should be directed to issues@omg.org by October 20, 2008.

You may view the pending issues for this specification from the OMG revision issues web page http://www.omg.org/issues/.

The FTF Recommendation and Report for this specification will be published on AugustNovember 1709, 200908. If you are reading this after that date, please download the available specification from the OMG Specifications Catalog.

USE OF SPECIFICATION - TERMS, CONDITIONS & NOTICES

The material in this document details an Object Management Group specification in accordance with the terms, conditions and notices set forth below. This document does not represent a commitment to implement any portion of this specification in any company's products. The information contained in this document is subject to change without notice.

LICENSES

The companies listed above have granted to the Object Management Group, Inc. (OMG) a nonexclusive, royalty-free, paid up, worldwide license to copy and distribute this document and to modify this document and distribute copies of the modified version. Each of the copyright holders listed above has agreed that no person shall be deemed to have infringed the copyright in the included material of any such copyright holder by reason of having used the specification set forth herein or having conformed any computer software to the specification.

Subject to all of the terms and conditions below, the owners of the copyright in this specification hereby grant you a fully-paid up, non-exclusive, nontransferable, perpetual, worldwide license (without the right to sublicense), to use this specification to create and distribute software and special purpose specifications that are based upon this specification, and to use, copy, and distribute this specification as provided under the Copyright Act; provided that: (1) both the copyright notice identified above and this permission notice appear on any copies of this specification; (2) the use of the specifications is for informational purposes and will not be copied or posted on any network computer or broadcast in any media and will not be otherwise resold or transferred for commercial purposes; and (3) no modifications are made to this specification. This limited permission automatically terminates without notice if you breach any of these terms or conditions. Upon termination, you will destroy immediately any copies of the specifications in your possession or control.

PATENTS

The attention of adopters is directed to the possibility that compliance with or adoption of OMG specifications may require use of an invention covered by patent rights. OMG shall not be responsible for identifying patents for which a license may be required by any OMG specification, or for conducting legal inquiries into the legal validity or scope of those patents that are brought to its attention. OMG specifications are prospective and advisory only. Prospective users are responsible for protecting themselves against liability for infringement of patents.

GENERAL USE RESTRICTIONS

Any unauthorized use of this specification may violate copyright laws, trademark laws, and communications regulations and statutes. This document contains information which is protected by copyright. All Rights Reserved. No part of this work covered by copyright herein may be reproduced or used in any form or by any means--graphic, electronic, or mechanical, including photocopying, recording, taping, or information storage and retrieval systems--without permission of the copyright owner.

DISCLAIMER OF WARRANTY

WHILE THIS PUBLICATION IS BELIEVED TO BE ACCURATE, IT IS PROVIDED "AS IS" AND MAY CONTAIN ERRORS OR MISPRINTS. THE OBJECT MANAGEMENT GROUP AND THE COMPANIES LISTED ABOVE MAKE NO WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, WITH REGARD TO THIS PUBLICATION, INCLUDING BUT NOT LIMITED TO ANY WARRANTY OF TITLE OR OWNERSHIP, IMPLIED WARRANTY OF MERCHANTABILITY OR WARRANTY OF FITNESS FOR A PARTICULAR PURPOSE OR USE.

IN NO EVENT SHALL THE OBJECT MANAGEMENT GROUP OR ANY OF THE COMPANIES LISTED ABOVE BE LIABLE FOR ERRORS CONTAINED HEREIN OR FOR DIRECT, INDIRECT, INCIDENTAL, SPECIAL, CONSEQUENTIAL, RELIANCE OR COVER DAMAGES, INCLUDING LOSS OF PROFITS, REVENUE, DATA OR USE, INCURRED BY ANY USER OR ANY THIRD PARTY IN CONNECTION WITH THE FURNISHING, PERFORMANCE, OR USE OF THIS MATERIAL, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

The entire risk as to the quality and performance of software developed using this specification is borne by you. This disclaimer of warranty constitutes an essential part of the license granted to you to use this specification.

RESTRICTED RIGHTS LEGEND

Use, duplication or disclosure by the U.S. Government is subject to the restrictions set forth in subparagraph (c) (1) (ii) of The Rights in Technical Data and Computer Software Clause at DFARS 252.227-7013 or in subparagraph (c)(1) and (2) of the Commercial Computer Software - Restricted Rights clauses at 48 C.F.R. 52.227-19 or as specified in 48 C.F.R. 227-7202-2 of the DoD F.A.R. Supplement and its successors, or as specified in 48 C.F.R. 12.212 of the Federal Acquisition Regulations and its successors, as applicable. The specification copyright owners are as indicated above and may be contacted through the Object Management Group, 140 Kendrick Street, Needham, MA 02494, U.S.A.

TRADEMARKS

MDA®, Model Driven Architecture®, UML®, UML Cube logo®, OMG Logo®, CORBA® and XMI® are registered trademarks of the Object Management Group, Inc., and Object Management GroupTM, OMGTM , Unified Modeling LanguageTM, Model Driven Architecture LogoTM, Model Driven Architecture DiagramTM, CORBA logosTM, XMI LogoTM, CWMTM, CWM LogoTM, IIOPTM , MOFTM , OMG Interface Definition Language (IDL)TM , and OMG Systems Modeling Language (OMG SysML)TM are trademarks of the Object Management Group. All other products or company names mentioned are used for identification purposes only, and may be trademarks of their respective owners.

COMPLIANCE

The copyright holders listed above acknowledge that the Object Management Group (acting itself or through its designees) is and shall at all times be the sole entity that may authorize developers, suppliers and sellers of computer software to use certification marks, trademarks or other special designations to indicate compliance with these materials.

Software developed under the terms of this license may claim compliance or conformance with this specification if and only if the software compliance is of a nature fully matching the applicable compliance points as stated in the specification. Software developed only partially matching the applicable compliance points may claim only that the software was based on this specification, but may not claim compliance or conformance with this specification. In the event that testing suites are implemented or approved by Object Management Group, Inc., software developed using this specification may claim compliance or conformance with the specification only if the software satisfactorily completes the testing suites.

OMG's Issue Reporting Procedure

All OMG specifications are subject to continuous review and improvement. As part of this process we encourage readers to report any ambiguities, inconsistencies, or inaccuracies they may find by completing the Issue Reporting Form listed on the main web page *http://www.omg.org*, under Documents, Report a Bug/Issue (http://www.omg.org/technology/agreement.htm).

Table of Contents

Preface

About the Object Management Group

OMG

Founded in 1989, the Object Management Group, Inc. (OMG) is an open membership, not-for-profit computer industry standards consortium that produces and maintains computer industry specifications for interoperable, portable and reusable enterprise applications in distributed, heterogeneous environments. Membership includes Information Technology vendors, end users, government agencies and academia.

OMG member companies write, adopt, and maintain its specifications following a mature, open process. OMG's specifications implement the Model Driven Architecture® (MDA®), maximizing ROI through a full-lifecycle approach to enterprise integration that covers multiple operating systems, programming languages, middleware and networking infrastructures, and software development environments. OMG's specifications include: UML® (Unified Modeling LanguageTM); CORBA® (Common Object Request Broker Architecture); CWMTM (Common Warehouse Metamodel); and industry-specific standards for dozens of vertical markets.

More information on the OMG is available at *http://www.omg.org/.*

OMG Specifications

As noted, OMG specifications address middleware, modeling and vertical domain frameworks. A catalog of all OMG Specifications is available from the OMG website at:

http://www.omg.org/technology/documents/spec_catalog.htm

Specifications within the Catalog are organized by the following categories:

OMG Modeling Specifications

- UML
- MOF
- XMI
- CWM
- Profile specifications.

OMG Middleware Specifications

- CORBA/IIOP
- IDL/Language Mappings
- Specialized CORBA specifications
- CORBA Component Model (CCM).

Platform Specific Model and Interface Specifications

- CORBAservices

- CORBAfacilities
- OMG Domain specifications
- OMG Embedded Intelligence specifications
- OMG Security specifications.

Relationship to Existing Domain OMG Specifications

- Single Nucleotide Polymorphism (SNP)

  The described specification extends the "Single Nucleotide Polymorphism" (formal/2005-11-01, formal/2005-12-01). Because it requires some changes in the SNP specification, the original SNP entities are also mentioned here, in this PAGE-OM specification.

- Bibliographic Query Service (BQS)

  This specification uses classes and attributes defined in the BQS (formal/02-05-03). It does not directly include the original BQS entities because its model is expressed as a CORBA model. However, it uses all bibliographic entities, except the query part (the query is not in the scope of this specification). The above is also true for the SNP specification. Therefore, this specification uses BQS only indirectly, via SNP.

- Life Sciences Identifiers (LSID)

  It is used in the Identifiable class (again, as with BQS, only indirectly, via SNP specification).

All of OMG's formal specifications may be downloaded without charge from our website. (Products implementing OMG specifications are available from individual suppliers.) Copies of specifications, available in PostScript and PDF format, may be obtained from the Specifications Catalog cited above or by contacting the Object Management Group, Inc. (as of January 16, 2006) at:

OMG Headquarters
140 Kendrick Street
Building A, Suite 300
Needham, MA 02494
USA
Tel: +1-781-444-0404
Fax: +1-781-444-0320
Email: *pubs@omg.org*

Certain OMG specifications are also available as ISO standards. Please consult *http://www.iso.org*

Issues

The reader is encouraged to report any technical or editing issues/problems with this specification to *http://www.omg.org/technology/agreement.htm.*

1 Scope

Please see Chapter 6 - Introduction.

2 Conformance

The normative parts of this specification are:

- Platform independent model (Chapter 7) expressed in the attached XML file created according to XMI format rules, v2. 1, using program Enterprise Architect (http://www.sparxsystems.com/), version 7.5.847~~7.0.817.1~~, UML version 2.1.
- Platform specific model (Chapter 8), representing an XML-based data exchange format, defined by an XML Schema definition.

Any implementation using or producing data exchange format defined by the Platform specific model defined by this specification is considered complying with this specification.

Any platform specific model that is derived from the platform independent model defined by this specification is also considered compliant.

If there is any inconsistency, or discrepancy between generality and specificity, between the platform independent and platform specific model, the platform specific model has precedence.

The normative parts are expressed in the ~~accompanied~~accompanying files in a document whose number is given in Appendix A (or elsewhere in this document). Parts of these files may also appear in the explanatory text of this document. If they do and if there are some differences or discrepancies the contents of the normative ~~accompanied~~accompanying files has precedence.

Regarding the use of ontology, it is not normative to use the ones listed in the “Ontology” section of this document, but it is highly suggested. The reason why this specification is not stricter about it is the reflection of the fact that ontology is a moving target in the bio-community and insisting on using only specific ones, may harm the usability of this specification.

3 Normative References

There are no normative references associated with this specification.

4 Terms and Definitions

Please see Annex C – Glossary.

5 Supporting Organizations

The following organizations have been involved in the process of developing, prototyping, and/or reviewing this specification. The authors thank them for participating and giving their valuable input.

- Institute for Molecular Medicine Finland, University of Helsinki, Finland
- University of Bergen, Norway
- Cold Spring Harbor Laboratory (CSHL), USA
- National Center for Biotechnology Information (NCBI), USA
- Stanford University, USA
- Yale University, USA
- University College London,
- Centre for Health Informatics and Multiprofessional Education, UK
- University of Leicester, UK
- South African National Bioinformatics Institute, South Africa
- School of Medicine, Tsinghua University, China
- Shanghai Center for Bioinformation Technology, China
- Shanghai Information Center for Life Sciences, China
- Institute of Genomics & Integrative Biology (IGIB), India
- Japan Science and Technology Agency (JST), Japan
- Medical Institute of Bioregulation, Kyushu University, Japan
- Mie University Life Science Research Center, Japan
- National Institute of Genetics (NIG), Japan
- RIKEN, Japan
- Tokyo Medical and Dental University (TMDU), Japan
- Tokyo Metropolitan Institute of Gerontology, Japan
- Graduate School of Medicine, University of Tokyo, Japan
- UNISYS, USA

5.1 Acknowledgements

The publishers of this document wish to express their appreciation to those listed below (in non-significant order) for their contributions of ideas and experience. Ultimately, the ideas expressed in this document are those of the authors and do not necessarily reflect the views or ideas of these individuals, nor does the inclusion of their names imply an endorsement of the final product.

Kimitoshi Naito, Akihiko Konagaya, Albert V. Smith, Anthony J. Brookes, David Fredman, Debasis Dash, Haseena Rajeevan, Heikki Lehväslaiho, Hideaki Sugawara, Hiroshi Mizushima, Juha Muilu, Jun Nakaya, Katsushi Tokunaga, Kei-Hoi Cheung, Kenshi Hayashi, Mark Woon, Masako Kuroda, Masashi Tanaka, Martin Senger, Matthew Darlison, Takeshi Tomiki, Toshio Kojima, Yasumasa Shigemoto, Yoshiji Yamada.

6 Introduction

Modern biology is striving to understand what factors generate inter-individual differences in structure, development, or behaviour within a species. Genetic factors (‘genotypes’) are undoubtedly of key importance in this equation, as are environmental conditions and even stochastic events, and much research is being done to elucidate how these things impact a range of normal and disease ‘phenotypes’ (i.e., the characteristics that can be observed). Progress in this area will ultimately lead to improved and increasingly personalised medical care, more productive agricultural and farming systems, and better solutions for environmental monitoring and control.

Many organisms are being explored and utilised with regards to deciphering genotype-phenotype relationships. On the genetic level, it is now possible to determine DNA sequences at extremely high-throughput, thereby revealing normal and pathogenic variation in and around genes, across individuals. On the level of the phenotype, an almost unlimited number of things could be influenced by genetic variants, and increasingly precise and standardised assays are being devised to measure and assess such things.

Both genetic and phenotype datasets need to be managed and combined to elucidate genotype-phenotype relationships. The genetic datasets have so far mostly comprised assays of specific sites of variation (creating ‘genotypes’), which have recently been growing exponentially due to technical advances. A subsequent new era is emerging wherein large DNA stretches (and perhaps complete genetic complements) are being fully (re)sequenced in relevant individuals. Phenotype datasets, both in human and non-human model organisms, are also being scaled up via the investigation of thousands or tens of thousands of individuals. Indeed, study integration is a major feature of current research; for example, existing Genome-wide Association (GWA) data done on various different studies are being merged in the ENAGE (ref) program leading to a potential GWA resource of 80,000 samples. As a result of this rapid progress in the modes of both genetic and phenotype analysis, the amount of genotype-phenotype data being generated is increasing at an exponential rate. The potential for converting this raw 'data' into useful ‘knowledge’ is therefore immense, but achieving this effectively will require distinct studies to be inter-related, cross-validated, and compared. In short, there is now an urgent need to ensure that genotype-phenotype investigations are supported by technologies that make them interoperable. Key to achieving this is a robust reference model for these types of data, via which all relevant domain knowledge can be effectively integrated and disseminated. That is the precisely the motivation behind building the Phenotype And GEnotype Object Model (PAGE-OM).

Anticipated user communities for PAGE-OM include biomedical researchers, clinicians, people involved in teaching and training such groups, and information technologists working in support of these domains.

Given the modern data production systems now at their disposal, biomedical researchers are constantly facing challenges of genotype-phenotype data management (e.g., storage, retrieval, tracking, reformatting, merging) and this directly impacts their ability to effectively analyse, share, and report their various datasets. Clinicians - for whom genotype-phenotype relationships are only one of many things to be considered - require streamlined and standardised access to explanations about DNA sequence alternatives and how and in which situations they may impact their clinical work. All of these needs will be helped by the development of PAGE-OM, which specifies the rational and functional segmentation of genotype-phenotype information into sub-components that are operationally distinct (the high-level objects in the model) and lays out the natural inter-relationships between these components. Once biomedical researchers and clinicians come to be familiar with these logical components and their relationships, it will be far easier for both groups to process and make sense of genotype-phenotype information, both within their own fields and in the important task of communicating between their domains of work. For this reason, it is also highly desirable that teachers of their two professions become familiar with the object modeling concepts, in order that they may suitably educate the researchers and doctors of the future.

Information technologists working in the biomedical fields will, more than any other group, benefit from understanding PAGE-OM in depth. These individuals have the job of providing the universe of databases, data pipelines, analysis tools, search functionalities, and exchanges protocols in which genotype-phenotype information will exist and be processed for exploitation. In the absence of any data model for genotype-phenotype information, these workers would each have to devise their own data model. Not only would this constant re-invention of the wheel be highly wasteful of manpower, it is also likely that many of the solutions they come up with would be suboptimal - given the limited experience that any one IT individual or IT team would usually have. Instead, PAGE-OM has been designed and piloted by a consortium of expert participants from nearly 50 Institutions engaged in many diverse genotype-phenotype projects. The PAGE-OM therefore provides a first-version common language, well-documented entity list, and carefully considered array of entity-relationships that IT workers can take directly off the shelf, with confidence that it should need nothing more than small adjustments or additions to serve their own specific needs. Additionally, as more and more IT groups adopt PAGE-OM and base their systems upon it, those systems will naturally become more and more similar and able to inter-communicate. This is an absolute necessity if the totality of genotype-phenotype information is ever to be merged into a single virtual corpus, for holistic and optimal utilisation.

To enable PAGE-OM to achieve its goals, it is structured as several high level concepts that can be used independently on the conceptual level (but not necessarily when using a modeling tool)~~ each of which can be used as a standalone model with its own object classes and relationships thereof.~~ ~~These concepts are~~Examples of these concepts are: Marker, Assay, Sample, Genotype, Frequency, Phenotype, and Experiment. ~~By way of example uses~~For example:~~;~~ A~~a~~ company providing DNA analysis kits might only need to use the Marker and Assay parts of this model. A genome variation database might use solely the Marker component, or perhaps the Assay, Sample, Genotype, and Frequency entities as well listing variant findings in population groups. A project involved in collecting and examining clinical samples might use only Sample and Phenotype parts, and if that team subsequently undertook DNA analysis it could extend their data systems by incorporating some or all of the remaining concepts in order to fully describe whatever genotype-phenotype relationships they might discover. Such flexibility is a necessary and innate characteristic of PAGE-OM, and that extends into the way we have matched the design principles and various classes of the model to those of other data standards initiates that cover other domains of biomedical research.

In conclusion, we believe PAGE-OM provides a timely, robust, and useful data model, sufficiently developed and tested to justify formal registration and deployment to the many and various communities engaged in genotype-phenotype data handling.

7 Platform Independent Model

The platform independent model is expressed as a UML model, UML version 2.1. The normative is its XMI representation, attached in the "Accompanying files" as the file PIM/ PAGE-OM_uml_2. 1_xmi_2. 1 .xml. The XMI was generated by the tool Enterprise Architect (http://www.sparxsystems.com/), version 7.5.847~~7.0.817.1~~, ~~without~~ including the EA specific extensions.

The XMI representing the platform independent model includes, because of the technical reasons and because of lacking package names, also the classes of the SNP-PML specification (that are used by this specification).

The basic data types (e.g., string or dateTime) are those inspired by the data types from the XML Schema (XML Schema Part 2: Datatypes - http://www.w3.org/TR/xmlschema-2/), where their exact descriptions can be found.

The full details with class and attributes description are attached in a generated file PIM/PAGE-OM.rtf. This file can be used interchangeably with the XMI file because both were generated from a model created by the Enterprise Architect tool - whose main file is also attached as file PIM/PAGE-OM.EAP. Note, however, that the latter file may have some EA specific expressions and extensions that are not normative for this specification.

For convenience, the whole PAGE-OM specification can be seen at http://www.pageom.org/models/omg/v_1.0/.

7.1 ~~Detailed~~ Model Diagrams

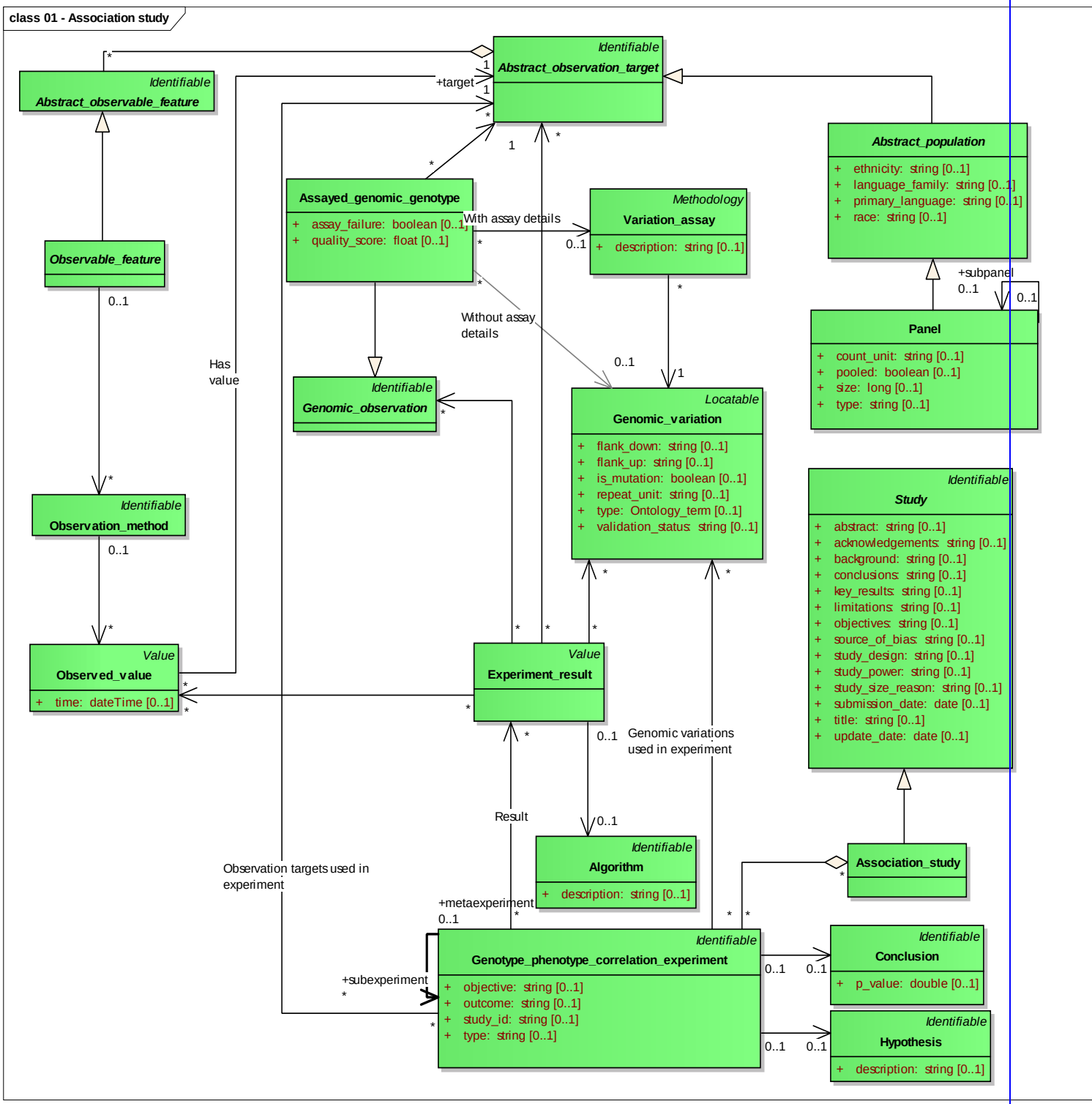

Figure 01 - Association study: Association study is the only kind of study specified in this model. It's parent, Study, can be used as an extension point to add new study types. Association studies are composed from set of experiments (Genotype_phenotype_correlation_experiment) done over observation targets (Panel, Molecular_sample, Individual). These experiments lead to set of results

(Experiment_result) documented by genomic observations and observed phenotype values.

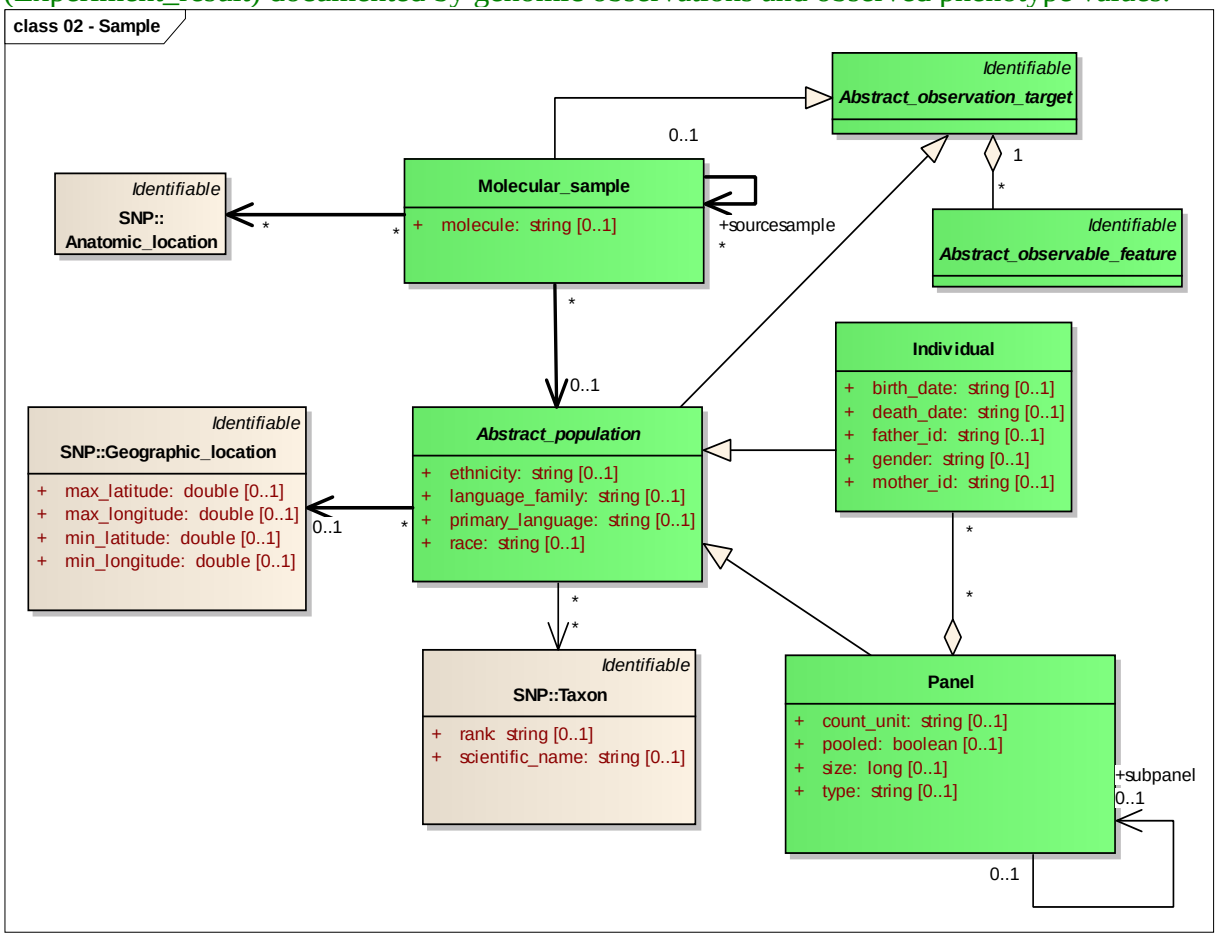

Figure 02 - Sample: An individual (Individual) interacts with the rest of the model by giving one or more samples (Molecular_samples) from different tissues. Mixed samples and more complicated sample organisations are modelled by allowing Molecular_samples and Panels be composed of multiples of themselves. In these cases it is common that individuals can not be identified.

An individual (Individual) has parent IDs as attributes. This information makes it possible to reconstruct pedigrees when needed. A taxon identifies the taxonomic group, typically species, the individual or the population sample belongs to. Further, they can be placed on a geographic map (Geographic_location)

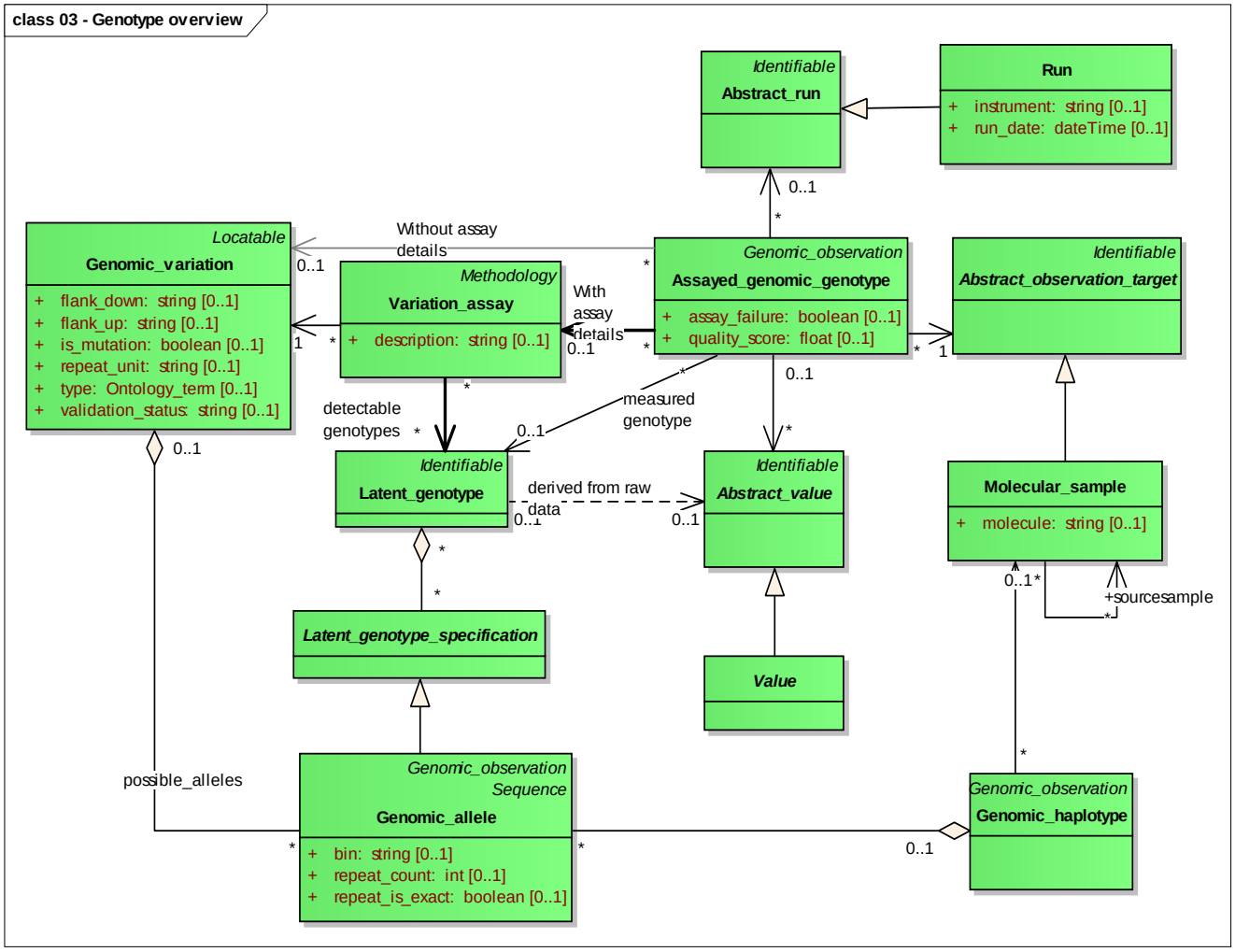

Figure 03 - Genotype overview: Genotypes (Assayed_genomic_genotype) produced over markers (Genomic_variation) and samples (Abstract_ovservation_target) under specific experimental conditions (Run). Latent_genotype has set of alleles (Genomic_allele) depending on ploidy level.

Genotyping assay can detect different genotypes from different individuals. One genotype is measured from one indivudual at given marker site. Genotypes may be stored with or without assay information (Variation assay is PCR (Polymerase Chain Reaction) assay designed for the vartiation site)

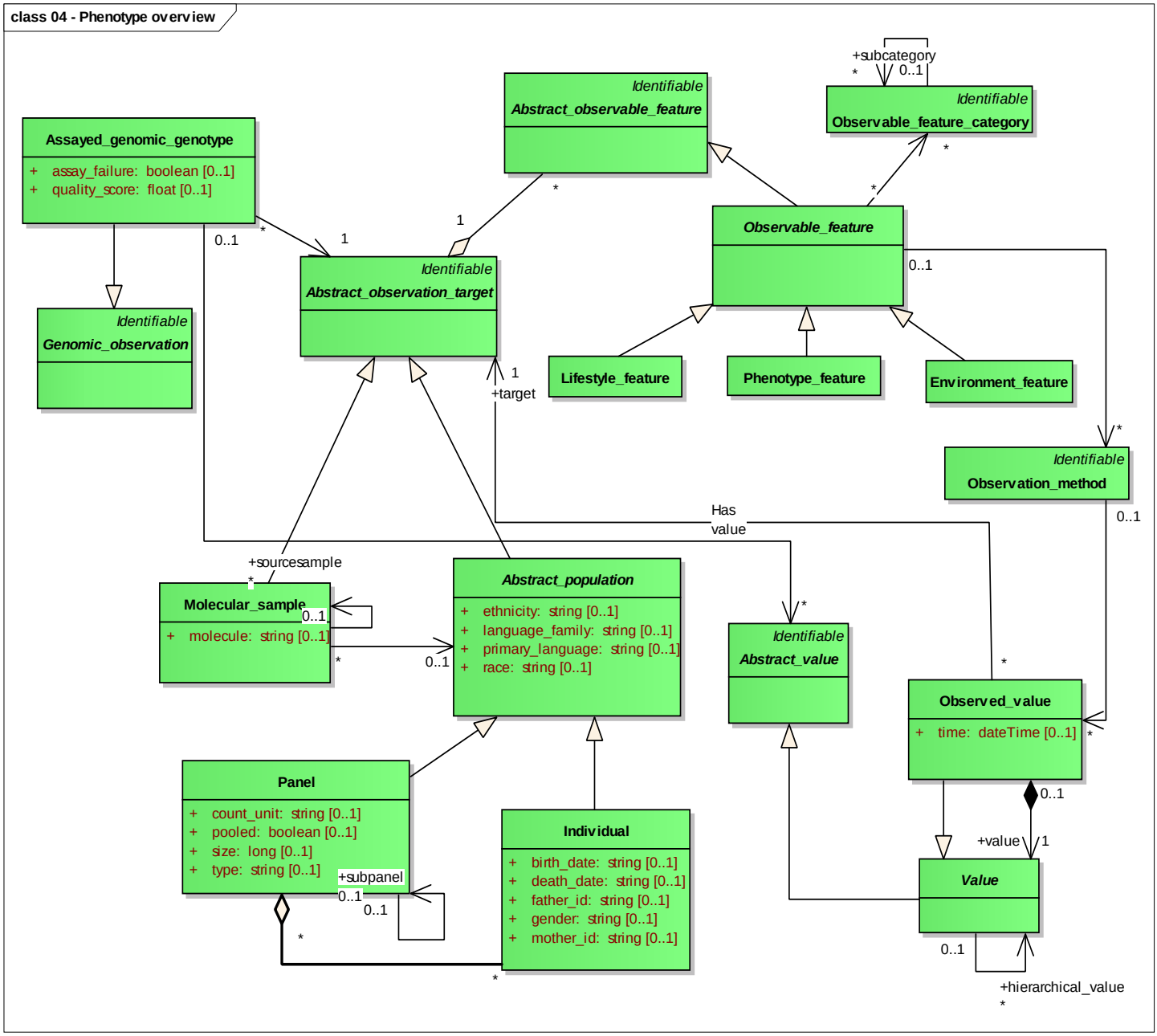

Figure 04 - Phenotype overview: Phenotype, environment and life style features are special cases of observable features from which observations can be made using specific observation methods. These observations lead to observed values obtained at specific time from observation target . Values can be also derived or categorized values (for example, high cholesterol level) in which case supporting values can be obtained from associated values, implemented using a recursion in the value model.

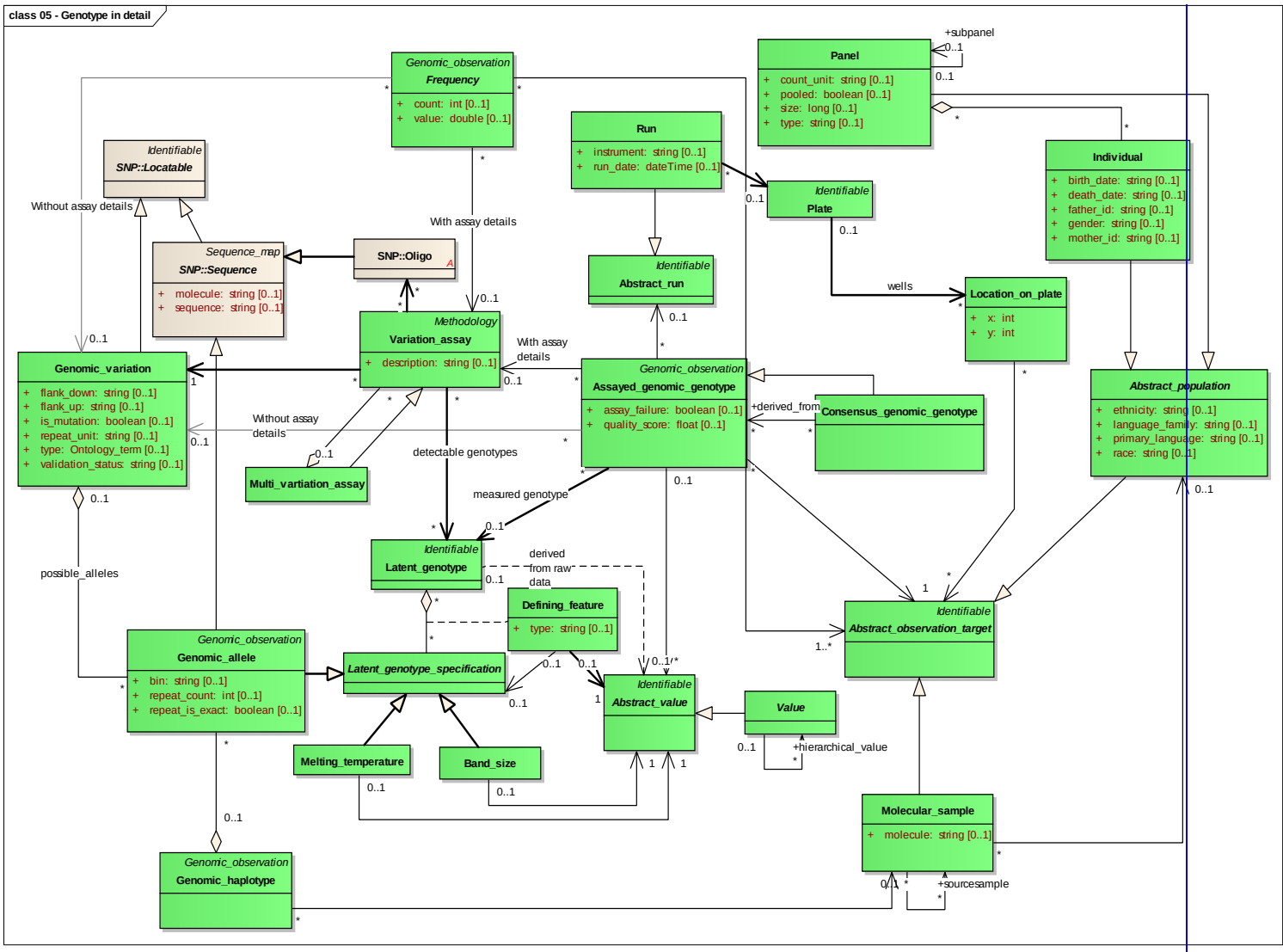

Figure 05 - Genotype in detail: An assay (Variation_assay) is applied to a sample (Molecular_sample) to get a genotype (Assayed_genomic_genotype). One or more Assayed_genomic_genotype are used to reach a consensus (Consensus_genomic_genotype). The assays commonly use oligo primers to detect allelic variations (Genomic_alleles in Genomic_variation).
Assay can have reference to possible detectable latent genotype (combination of alleles depending on ploidy). One of these combinations can be detected in single genotype measurement (Assayed_genomic_genotype) done using the specific assay.

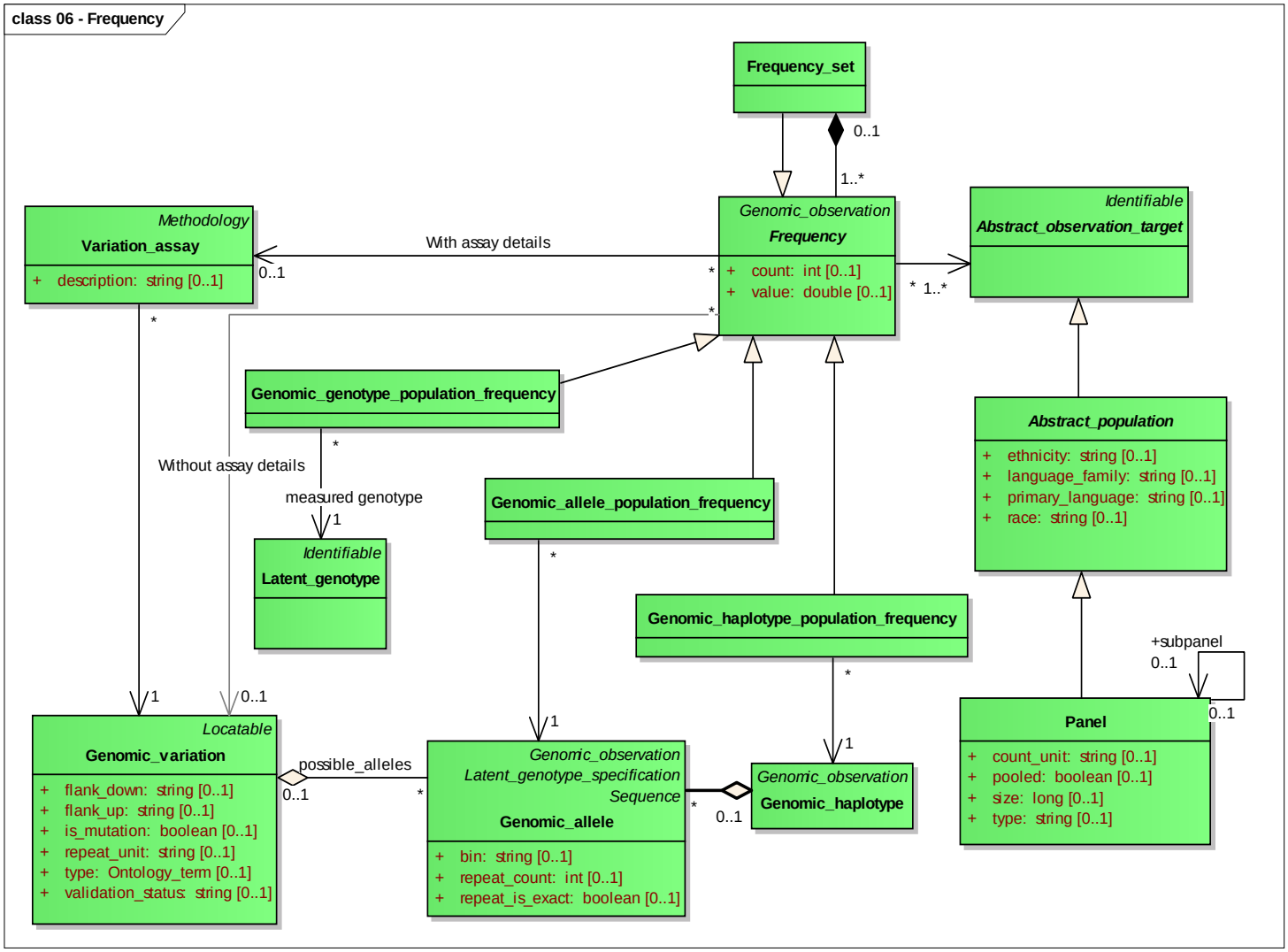

Figure 06 - Frequency: Alleles (Genomic_alleles), genotypes (Latent_genotype) and haplotypes (Genomic_haplotype) can have measured frequencies in population samples (Panels). In addition, heterozygosity (Heterozygosity) is a measure of observed variability of a polymorphic site (Genomic_variation) in a sub-population (Panel).

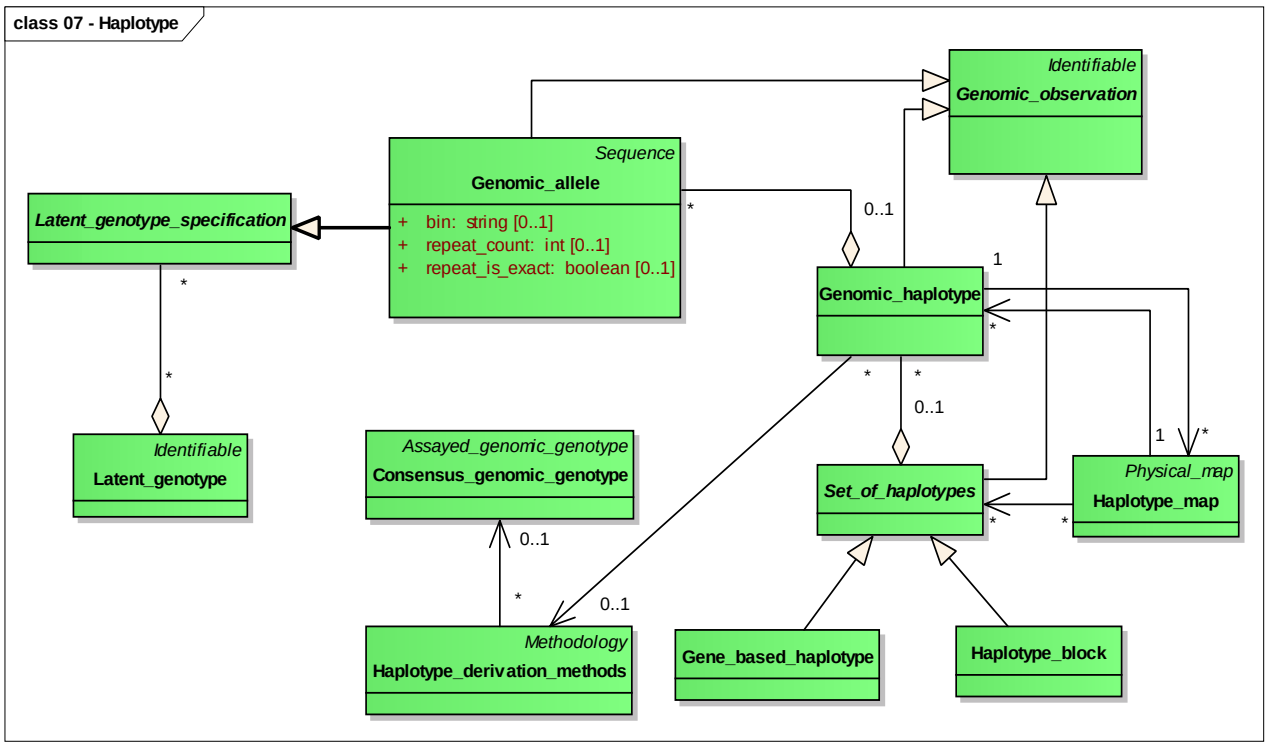

Figure 07 - Haplotype: Haplotype (Genomic_haplotype) is a set of polymorphisms (Genomic alleles) on a single chromosome (chromatid). Haplotypes may be grouped into sets like haplotype blocks (Haplotype_block) separated by recombination regions and gene based haplotype (Gene_based_haplotype), which combines sets of haplotypes on same gene structure.

Depending on an assay used, a genotype (Consensus_genomic_genotype) may contain information about the phase of the detected Genomic_alleles. Alternatively, various additional methods (Haplotype_derivation_methods) can be used to measure or calculate haplotypes (Genomic_haplotypes) from the genotype data. The aim of many haplotype studies is to find haplotypes within a given sequence region that define most of the variation in populations (Set_of_haplotypes). Haplotype blocks (Haplotype_blocks) are clusters determined by linkage disequilibrium-based methods. The other common way to define clusters is use a the sequence region of the gene (Gene_based_haplotype).

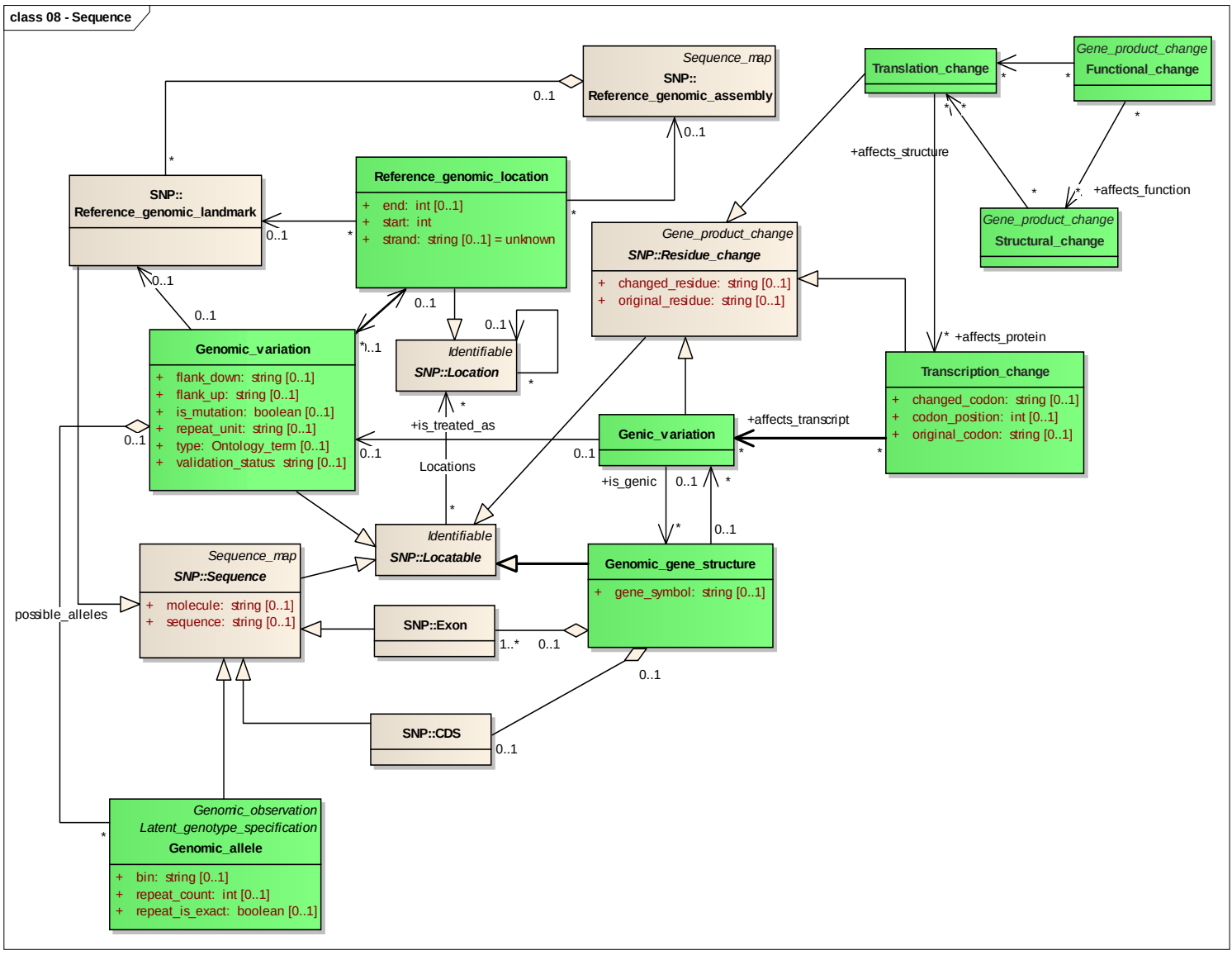

Figure 08 - Sequence: All sequences inherit from a generic Sequence class. All sequences (Sequence) and sequence features (e.g. Genomic_variation) can be located within a genomic sequence (Reference_genomic_sequence). This is modeled by common inheritance from an abstract superclass Locatable that can have several locations (Genomic_reference_location).

A Reference_genomic_landmark is any accessioned sequence within Reference_genomic_assembly. The model allows for multiple assemblies. Any location within a landmark and therefore in an assembly is called Reference_genomic_location. Any variable site in an assembly is a Genomic_polymorphism. The variable sequences within Genomic_variations are called Genomic_alleles.

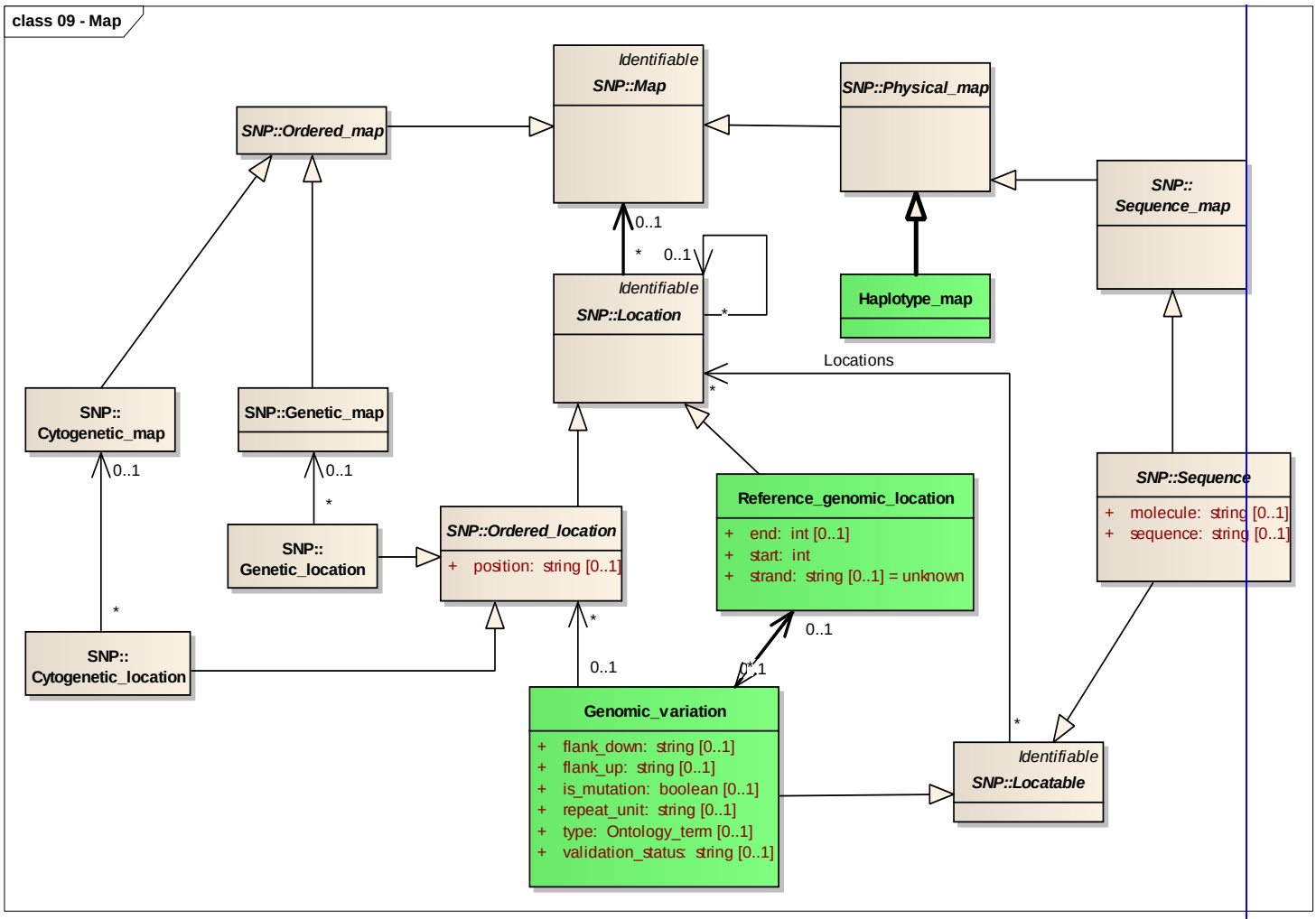

Figure 09 - Map: A map organizes genomic features and assigns them locations. The primary maps (Physical_map, Sequence_map) which have additive sequence residue count distances between markers, are separated here from secondary maps (Ordered_map) which at best can only deal with non-metric distances. For each type a map (Map) there is a corresponding location (Location). Hence a marker (Genetic_variation) can have several locations (Reference_genomic_location) in sequences (Reference_genomic_landmark) and additionally be located within a band (Cytogenetic_location) in a species-specific karyotype (Cytogenetic_map) as well as in (Genetic_location) several genetic maps (Genetic_map), e.g. in maps based on male and female recombination frequencies. Genomic assembly (Reference_genomic_assembly) is a physical sequence map which is combination of genomic sequence(Reference_genomic_landmark). Feature locations can also be expressed in chromosomal locations in an assembly(Reference_genomic_location_in_assembly).

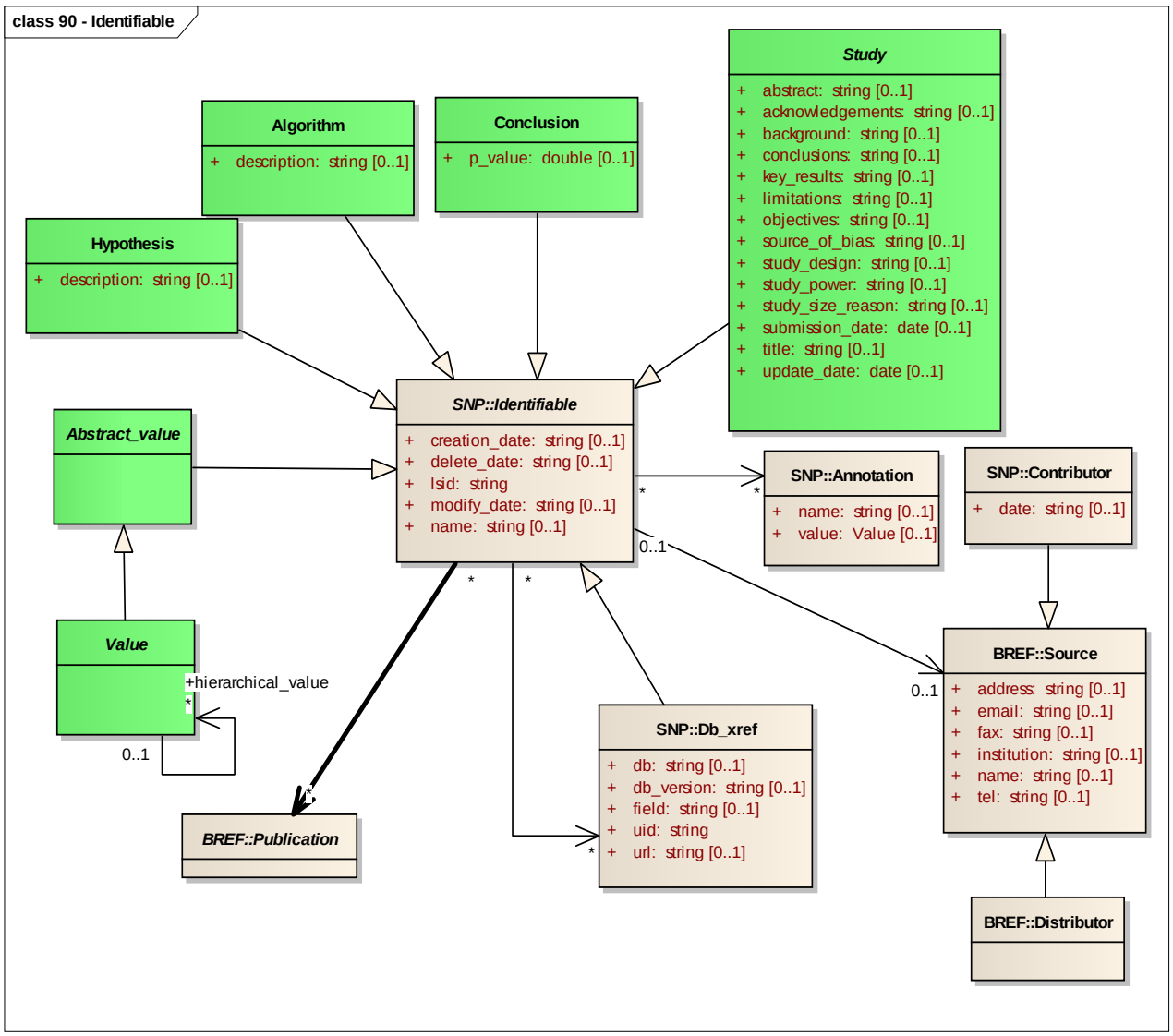

Figure 90 - Identifiable: All classes in the model inherit from Identifiable. In this way, their instances are uniquely identifiable. Any Identifiable instance must use its "lsid attribute". For this attribute, it is recommended to use the OMG Life Sciences Identifier specification. All classes in the model can be linked with annotation (Annotation) and database cross-reference (Db_xref). Special kinds of cross references are Source for source of data, Distributor for the original database of the data, and Contributor for tracking editorial changes to data.

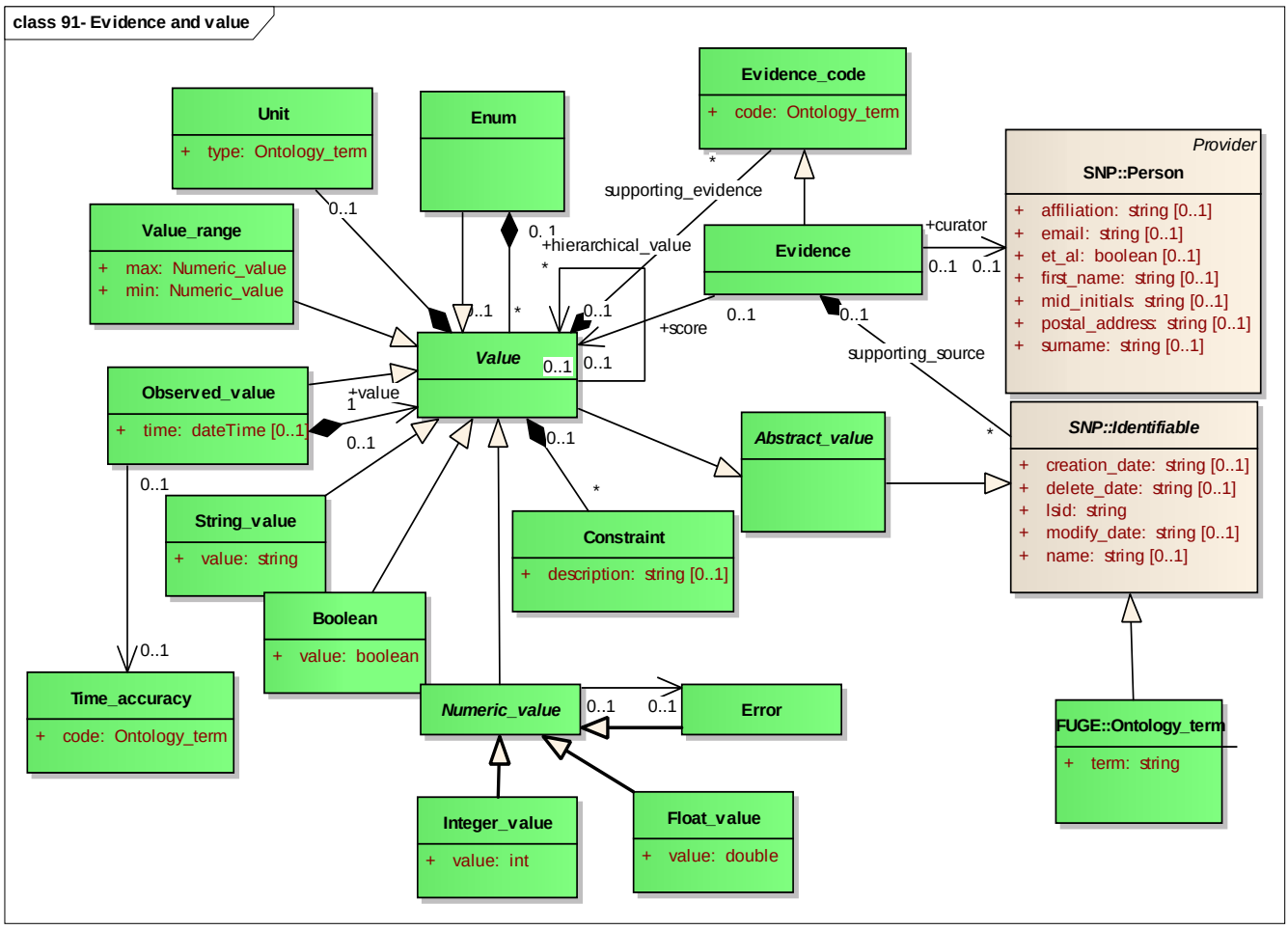

Figure 91- Evidence and value: Core ("simple") string, arithmetic data types and object references are modeled here, in the Value model. The model is based on concept developed in Generation Challenge Program: http://pantheon.generationcp.org/demeter/Values.html
Model Documentation

~~ocumentation~~

7.2 Detailed Model Documentation

Notes

Those cardinalities that are not explicitly given in PIM should be interpreted~~considered~~ as ~~cardinality~~ "0..1~~one~~".

Model Documentation

PAGE-OM

Package: Model

Document:
PAGE-OM model (Brookes) captures information related to genotype and phenotype observations and their relationships. The core conceptual domain is experiment part (See Figure 01 - Association study), which bring in data from phenotype (Figure 04- Phenotype overview) and genotype domains (Figures 03 – Genotype overview, 05 - Genotype in detail and 06 - Frequency) along with experimental result information that elucidates how genetic variations influence phenotypic variation.

Abstract_observable_feature

Type: **Class Identifiable**
Package: PAGE-OM

Document:
An extension point for other kinds of observable features.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Observable_feature	Abstract_observable_feature
Aggregation Target, for example sample or individual.	Card. Role Name	* Abstract_observable_feature	1 Abstract_observation_target
Generalization	Card. Role Name	Abstract_observable_feature	Identifiable

Abstract_observation_target

Type: **Class Identifiable**

Package: PAGE-OM

Document:

It is an abstract class for all entities from which one can make genotype or phenotype measurements or observations. It deals with entities capable of being observed.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Molecular_sample	Abstract_observation_target
Association Observation target (e.g. sample) from which genotype is measured.	Card. Role Name	* Assayed_genomic_genotype	1 Abstract_observation_target
Aggregation Target, for example sample or individual.	Card. Role Name	* Abstract_observable_feature	1 Abstract_observation_target
Generalization	Card. Role Name	Abstract_population	Abstract_observation_target
Association Observation target, sample, on well location.	Card. Role Name	* Location_on_plate	* Abstract_observation_target
Association Target from which frequency is measured.	Card. Role Name	* Frequency	1..* Abstract_observation_target
Association *Has value* Observed values (of observable features)	Card. Role Name	* Observed_value	1 target Abstract_observation_target

Connector		Source	Target
measured on the target			
Association Panel (set of samples or study subjects) over which experiment was done.	Card. Role Name	* Experiment_result	* Abstract_observation_target
Association *Observation targets used in experiment*	Card. Role Name	* Genotype_phenotype_correlation_experiment	* Abstract_observation_target
Generalization	Card. Role Name	Abstract_observation_target	Identifiable

Abstract_population

Type: **Class Abstract_observation_target**

Package: PAGE-OM

Document:

An interbreeding set of individuals, from whom a Panel is drawn. (Population in SNP-PML) . Extends Abstract_observation_target, which is abstract class for all entities from which one can make genotype or phenotype measurements or observations.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Abstract_population	Abstract_observation_target
Association	Card. Role	*	0..1

Connector		Source	Target
Source of sample. For example individual or panel (pool of individuals).	Name	Molecular_sample	Abstract_population
Association Geographic location of individual or group of individuals (panel)	Card. Role Name	* Abstract_population	0..1 Geographic_location
Association Taxonomic name (e.g. species) of individual or population.	Card. Role Name	* Abstract_population	* Taxon
Generalization	Card. Role Name	 Individual	 Abstract_population
Generalization	Card. Role Name	 Panel	 Abstract_population

Attributes

Attribute	Lower bound	Notes
ethnicity	*0*	Additional ethnic category of the population sample or "mixed".
language_family	*0*	Language family name or code, e.g. as in Ethnologue
primary_language	*0*	Language spoken (name or code), e.g. as in Ethnologue
race	*0*	Broad ethnic category of the population sample or "mixed".

Abstract_run

Type: **Class Identifiable**

Package: PAGE-OM

Document:

An extension point for other kinds of runs.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Abstract_run	Identifiable
Association Provides information on experimental conditions (run of experiment).	Card. Role Name	* Assayed_genomic_genotype	0..1 Abstract_run
Generalization	Card. Role Name	Run	Abstract_run

Abstract_value

Type: **Class Identifiable**

Package: PAGE-OM

Document:

An extension point for kinds of values

Connections

Connector		Source	Target
Association Raw data values used in allele calling.	Card. Role Name	0..1 Defining_feature	1 Abstract_value
Generalization	Card. Role Name	 Value	 Abstract_value
Generalization	Card. Role Name	 Abstract_value	 Identifiable
Dependency *derived from raw data* Latent genotypes do depend on actual measured values (raw data).	Card. Role Name	0..1 Latent_genotype	0..1 Abstract_value
Association Raw data. Actual measured values like intensities.	Card. Role Name	0..1 Assayed_genomic_genotype	* Abstract_value
Association Value of melting temperature.	Card. Role Name	0..1 Melting_temperature	1 Abstract_value
Association Value of band-size.	Card. Role Name	0..1 Band_size	1 Abstract_value

Algorithm

Type: **Class Identifiable**

Package: PAGE-OM

Document:

Step-by-step procedure for solving a problem

Connections

Connector		Source	Target
Association	Card. Role Name	0..1 Page	* Algorithm
Association Algorithm used to obtain the result	Card. Role Name	0..1 Experiment_result	0..1 Algorithm
Generalization	Card. Role Name	 Algorithm	 Identifiable

Attributes

Attribute	Lower bound	Notes
description	*0*	Description of algorithm

Assayed_genomic_genotype

Type: **Class Genomic_observation**

Package: PAGE-OM

Document:

The result of applying a variation assay to an individual, to reveal one or more of the genomic alleles carried by that individual. This term applies to the observed data rather than to the inferred state of the individual. Thus the same individual might have several different genotypes at the same site, where the variation might be due to differing assays, experimental error, dominant systems, missing data, and so forth. Synonym: Measurement.

Connections

Connector		Source	Target
Association Observation target (e.g. sample) from which genotype is measured.	Card. Role Name	* Assayed_genomic_genotype	1 Abstract_observation_target
Association Provides information on experimental conditions (run of experiment).	Card. Role Name	* Assayed_genomic_genotype	0..1 Abstract_run
Association *measured genotype*	Card. Role Name	* Assayed_genomic_genotype	0..1 Latent_genotype
Association *Without assay details* Used when assay level information is not know or is not important. There can be more than one genotypes per genomic variation, done over different samples.	Card. Role Name	* Assayed_genomic_genotype	0..1 Genomic_variation
Association *With assay details* Used when assay level information is needed. Specific variation assay designed for genomic site (locus).	Card. Role Name	* Assayed_genomic_genotype	0..1 Variation_assay
Association Source genotypes from which consensus is derived from.	Card. Role Name	* Consensus_genomic_genotype	* derived_from Assayed_genomic_genotype

Connector		Source	Target
Association	Card. Role Name	0..1 Page	* Assayed_genomic_genotype
Generalization	Card. Role Name	 Consensus_genomic_genotype	 Assayed_genomic_genotype
Association Raw data. Actual measured values like intensities.	Card. Role Name	0..1 Assayed_genomic_genotype	* Abstract_value
Generalization	Card. Role Name	 Assayed_genomic_genotype	 Genomic_observation

Attributes

Attribute	Lower bound	Notes
assay_failure	*0*	Failure of assay. Value is true if assay has failed
quality_score	*0*	Quality score of measurement. Depends on the instrument

Association_study

Type: **Class Study**

Package: PAGE-OM

Document:

Association study is core concept of the specification. It captures relationships between phenotypes and genotypes. It is an examination of genetic variation across the genome, designed to identify genetic associations with observable phenotypes. Association studies are results of correlation experiments,

Connections

Connector		Source	Target
Association	Card. Role Name	0..1 Page	* Association_study
Aggregation Correlation experiments which are part (or used in) study.	Card. Role Name	* Genotype_phenotype_correlation_experiment	* Association_study
Generalization	Card. Role Name	 Association_study	 Study

Band_size

Type: **Class Latent_genotype_specification**

Package: PAGE-OM

Document:

DNA fragment length estimated from gel electrophoresis

Connections

Connector		Source	Target
Generalization	Card. Role Name	 Band_size	 Latent_genotype_specification
Association	Card. Role Name	0..1 Page	* Band_size
Association Value of band-size.	Card. Role Name	0..1	1

Connector		Source	Target
		Band_size	Abstract_value

Boolean

Type: **Class Value**

Package: PAGE-OM

Document:

Value of type boolean

Connections

Connector		Source	Target
Generalization	Card. Role Name	Boolean	Value

Attributes

Attribute	Lower bound	Notes
value	*1*	Boolean value

Conclusion

Type: **Class Identifiable**

Package: PAGE-OM

Document:

A reasoned judgment of an experiment

Connections

Connector		Source	Target
Association Conclusion of experiment	Card. Role Name	0..1 Genotype_phenotype_correlation_experiment	0..1 Conclusion
Association	Card. Role Name	0..1 Page	* Conclusion
Generalization	Card. Role Name	 Conclusion	 Identifiable

Attributes

Attribute	Lower bound	Notes
p_value	*0*	Probability value

Consensus_genomic_genotype

Type: **Class Assayed_genomic_genotype**

Package: PAGE-OM

Document:

This class represents consensus from several experiments providing genotypes of the same sample on the same site.

Connections

Connector		Source	Target
Association Consensus genotypes used to derive the haplotype.	Card. Role Name	* Haplotype_derivation_methods	0..1 Consensus_genomic_genotype
Association The consensus genotype whose frequency is given.	Card. Role Name	* Genomic_genotype_population_frequency	0..1 Consensus_genomic_genotype
Association Source genotypes from which consensus is derived from.	Card. Role Name	* Consensus_genomic_genotype	* derived_from Assayed_genomic_genotype
Generalization	Card. Role Name	 Consensus_genomic_genotype	 Assayed_genomic_genotype

Constraint

Type: **Class**

Package: PAGE-OM

Document:

The contents of a Value can be limited by Constraints. Different types of Constraints allow various ways how to limit or validate one or more Value instances. The Constraint superclass only stores a string description of the Constraint. The actual full semantics of a constraint are specified in various subclasses described below.

But there are no subclasses in the PAGE-OM - because they are out of scope of PAGE-OM.

Connections

Connector		Source	Target

Connector		Source	Target
Aggregation Constraints	Card. Role Name	* Constraint	0..1 Value

Attributes

Attribute	Lower bound	Notes
description	*0*	Description

Defining_feature

Type: **AssociationClass**

Package: PAGE-OM

Document:

Is an association class that has list of values, which are used in defining the instance of Latent_genotype_specifications (for example intensity values used in allele calling). The class captures information how alleles are called (observed) from raw measurement values like intensity values

Connections

Connector		Source	Target
Association Raw data values used in allele calling.	Card. Role Name	0..1 Defining_feature	1 Abstract_value
Association Gives information on allele calling.	Card. Role Name	0..1 Defining_feature	0..1 Latent_genotype_specification

Attributes

Attribute	Lower bound	Notes
type	*0*	Type of feature

Enum

Type: **Class Value**

Package: PAGE-OM

Document:

Enumeration contains list of Values

Connections

Connector		Source	Target
Aggregation	Card. Role Name	* Value	0..1 Enum
Generalization	Card. Role Name	 Enum	 Value

Environment_feature

Type: **Class Observable_feature**

Package: PAGE-OM

Document:

Circumstances, objects, or conditions by which one is surrounded

Connections

Connector		Source	Target
Generalization	Card. Role Name	Environment_feature	Observable_feature
Association	Card. Role Name	0..1 Page	* Environment_feature

Error

Type: **Class Numeric_value**

Package: PAGE-OM

Document:

Error value is numeric value of accuracy. Quality score

Connections

Connector		Source	Target
Generalization	Card. Role Name	Error	Numeric_value
Association Error of numeric value.	Card. Role Name	0..1 Numeric_value	0..1 Error

Evidence

Type: **Class Evidence_code**

Package: PAGE-OM

Document:

.Evidence indicates reliability of a feature or simply documents its authoritative origin

Connections

Connector		Source	Target
Aggregation *supporting_source* Supporting evidences: Citations, studies, ontology terms etc	Card. Role Name	* Identifiable	0..1 Evidence
Association Score of evidence (e.g. p-value)	Card. Role Name	0..1 Evidence	0..1 score *score of value* Value
Association	Card. Role Name	0..1 Page	* Evidence
Generalization	Card. Role Name	 Evidence	 Evidence_code
Association Curator of evidence.	Card. Role Name	0..1 Evidence	0..1 curator *Curator of evidence* Person

Evidence_code

Type: **Class**

Package: PAGE-OM

Document:

which is a controlled vocabulary term such as a GO evidence) Evidence can be an EvidenceCode inheriting from)but can be a more fully documented Evidence object (code or ICIS Method code .a curator modeled as a Contact ,lly curated by a specified persongenera (EvidenceCode

but also ,1and 0which is usually a numeric value between)Its strength is expressed by the score .(an ontology term value .g.e -other types of Value are allowed

because it is identified by)is its supporting source which can be anything The core of an evidence .Studies and OntologyTerms ,Usual evidence sources are BiblioReferences .(a SimpleIdentifier (html.Features/demeter/org.generationcp.pantheon//:http -generationcp)Reference

Connections

Connector		Source	Target
Aggregation *supporting_evidence* Values supporting the evidence.	Card. Role Name	* Evidence_code	0..1 Value
Generalization	Card. Role Name	 Evidence	 Evidence_code

Attributes

Attribute	Lower bound	Notes
code	*1*	Evidence code as specified using ontology term

Experiment_result

Type: **Class Value**

Package: PAGE-OM

Document:

The experiment result (for example a single p-value) gathers correlation between genomic observation and phenotypic observed values. A correlation experiment can consist of more than one experiment results.

Connections

Connector		Source	Target
Association	Card. Role Name	0..1 Page	* Experiment_result
Association *Result* Experiments results which are part of the correlation experiment	Card. Role Name	* Genotype_phenotype_correlation_experiment	* Experiment_result
Association Algorithm used to obtain the result	Card. Role Name	0..1 Experiment_result	0..1 Algorithm
Generalization	Card. Role Name	 Experiment_result	 Value
Association Panel (set of samples or study subjects) over which experiment was done.	Card. Role Name	* Experiment_result	* Abstract_observation_target
Association Combination of Genomic_observations determined for each	Card. Role Name	* Experiment_result	* Genomic_observation

Connector		Source	Target
Observation_target. For example actual genotypes over Molecular_samples			
Association Combination of Observable_values determined for each Observation_target. For example phenotype in GWA studies.	Card. Role Name	* Experiment_result	* Observed_value
Association Combination of Genomic_variations studied for each Observation_target. For example marker dimension in GWA studies.	Card. Role Name	* Experiment_result	* Genomic_variation

Float_value

Type: **Class Numeric_value**

Package: PAGE-OM

Document:

Value of type float

Connections

Connector		Source	Target
Generalization	Card. Role Name	Float_value	Numeric_value

Attributes

Attribute	Lower bound	Notes
value	*1*	Value

Frequency

Type: **Class Genomic_observation**

Package: PAGE-OM

Document:

Abstract class for frequencies, expressed in percentages. Alleles (Genomic_alleles), genotypes (Consensus_genomic_genotype) and haplotypes (Genomic_haplotype) can have measured frequencies in population samples (Panels). In addition, heterozygosity (Heterozygosity) is a measure of observed variability of a polymorphic site (Genomic_variation)in a sub-population (Panel).

Connections

Connector		Source	Target
Generalization Frequency of heterozygote alleles.	Card. Role Name	 Heterozygosity	 Frequency
Association Target from which frequency is measured.	Card. Role Name	* Frequency	1..* Abstract_observation_target
Generalization	Card. Role Name	 Frequency	 Genomic_observation
Association *Without assay details* Genomic variation	Card. Role Name	* Frequency	0..1 Genomic_variation

Connector		Source	Target
site (for example marker).			
Association *With assay details* Assay desinged for the variation site. Association is used when assay level information is needed. (Optional with the "direct" association from Frequency to Genomic_variation) .	Card. Role Name	* Frequency	0..1 Variation_assay
Generalization	Card. Role Name	 Frequency_set	 Frequency
Aggregation Set of frequencies.	Card. Role Name	1..* Frequency	0..1 Frequency_set
Generalization	Card. Role Name	 Genomic_allele_population_frequency	 Frequency
Generalization	Card. Role Name	 Genomic_haplotype_population_frequency	 Frequency
Generalization	Card. Role Name	 Genomic_genotype_population_frequency	 Frequency

Attributes

Attribute	Lower bound	Notes
count	*0*	Total number

Attribute	Lower bound	Notes
value	*0*	Value of frequency (%)

Frequency_set

Type: **Class Frequency**

Package: PAGE-OM

Document:

Set of frequencies

Connections

Connector		Source	Target
Association	Card. Role Name	0..1 Page	* Frequency_set
Generalization	Card. Role Name	 Frequency_set	 Frequency
Aggregation Set of frequencies.	Card. Role Name	1..* Frequency	0..1 Frequency_set

Functional_change

Type: **Class Gene_product_change**

Package: PAGE-OM

Document:

Change in the function of the final gene product.

Connections

Connector		Source	Target
Association Translation which affects function of protein.	Card. Role Name	* Functional_change	* Translation_change
Association Structural change which has impact on function of protein.	Card. Role Name	* Functional_change	* affects_function Structural_change
Association	Card. Role Name	0..1 Page	* Functional_change
Generalization	Card. Role Name	 Functional_change	 Gene_product_change

Gene_based_haplotype

Type: **Class Set_of_haplotypes**

Package: PAGE-OM

Document:

Collection of variable nucleotides (Genomic_alleles in Genomic_variations) that define a gene. In older usage synonym locus.

Connections

Connector		Source	Target
Association	Card. Role Name	0..1 Page	* Gene_based_haplotype
Generalization	Card. Role Name	 Gene_based_haplotype	 Set_of_haplotypes

Genic_variation

Type: **Class Residue_change**

Package: PAGE-OM

Document:

Genomic variation with location in genic coordinates. Synonym: mutation (when change from a common allele
affects phenotype)

Connections

Connector		Source	Target
Generalization	Card. Role Name	 Genic_variation	 Residue_change
Association Genioc variation which has impact on gene structure.	Card. Role Name	0..1 Genomic_gene_structure	* Genic_variation
Association Genomic variation site of the genic variation.	Card. Role Name	0..1 Genic_variation	0..1 is_treated_as Genomic_variation
Association Genic variation	Card. Role Name	*	* affects_transcript

Connector		Source	Target
which influences on the transcript change.		Transcription_change	Genic_variation
Association Genomic structure(s) which are associated to the genic variation.	Card. Role Name	0..1 is_genic Genic_variation	* Genomic_gene_structure
Association	Card. Role Name	0..1 Page	* Genic_variation

Genomic_allele

Type: **Class Genomic_observation, Latent_genotype_specification, Sequence**

Package: PAGE-OM

Document:

One of several alternative DNA sequences of a Reference_genomic_location as it appears in the population of organisms. Synonym: variant, allele

Connections

Connector		Source	Target
Association Alleles which are part of the haplotype i.e. are on same phase.	Card. Role Name	* Genomic_allele	0..1 Genomic_haplotype
Generalization	Card. Role Name	 Genomic_reference_allele	 Genomic_allele
Aggregation *possible_alleles* The relationship to	Card. Role Name	* Genomic_allele	0..1 Genomic_variation

Connector		Source	Target
gives information on possible sequence variations attached to the locus (as defined by flanking sequences).			
Generalization	Card. Role Name	Genomic_allele	Latent_genotype_specification
Generalization	Card. Role Name	Genomic_allele	Sequence
Generalization	Card. Role Name	Genomic_allele	Genomic_observation
Association The DNA sequence of a genomic variation as it appears in the .public database	.Card Role Name	0..1 one_of Genomic_reference_allele	0..1 Genomic_allele
Association	Card. Role Name	0..1 Page	* Genomic_allele
Association Allele whose frequency is reported.	Card. Role Name	* Genomic_allele_population_frequency	1 Genomic_allele

Attributes

Attribute	Lower bound	Notes
bin	*0*	Size class for microsatellite alleles when exact size can not be determined. Semantic name.
count_repeat	*0*	t number of ge ,'microsatellite'variation type is _If the Genomic 7 .g.e ,repeat units as value
repeat_is_exact	*0*	The Genomic_variation type is 'microsatellite' and gets value true if the allele region consists of repeat units only.

Genomic_allele_population_frequency

Type: **Class Frequency**

Package: PAGE-OM

Document:

The frequency with which a particular Genomic_allele is seen in a particular Panel. This frequency can be measured from
pooled samples. Synonyms: Genomic_allele_panel_frequency, allele_frequency.

Connections

Connector		Source	Target
Association	Card. Role Name	0..1 Page	* Genomic_allele_population_frequency
Association Allele whose frequency is reported.	Card. Role Name	* Genomic_allele_population_frequency	1 Genomic_allele
Generalization	Card. Role Name	 Genomic_allele_population_frequency	 Frequency

Genomic_gene_structure

Type: **Class Locatable**

Package: PAGE-OM

Document:

A structure of a gene expressed as location of the CDS and exons. Defines genic coordinate

system from start of the CDS downstream.

Connections

Connector		Source	Target
Association Coding region which is part of the gene structure.	Card. Role Name	0..1 CDS	0..1 Genomic_gene_structure
Association Exons which are part of gene structure.	Card. Role Name	1..* Exon	0..1 Genomic_gene_structure
Generalization	Card. Role Name	 Genomic_gene_structure	 Locatable
Association Genioc variation which has impact on gene structure.	Card. Role Name	0..1 Genomic_gene_structure	* Genic_variation
Association Genomic structure(s) which are associated to the genic variation.	Card. Role Name	0..1 is_genic Genic_variation	* Genomic_gene_structure
Association	Card. Role Name	0..1 Page	* Genomic_gene_structure

Attributes

Attribute	Lower bound	Notes
gene_symbol	*0*	gene symbol for the gene e.g. approved by the HUGO nomenclature committee.

Genomic_genotype_population_frequency

Type: **Class Frequency**

Package: PAGE-OM

Document:

Frequency of a Consensus_genomic_genotype in a Panel.

OSAGE-OM
Has many to one relationship to Latent_genotype (Consensus_genomic_genotype in SNP-PML)

Connections

Connector		Source	Target
Association *measured genotype*	Card. Role Name	* Genomic_genotype_population_frequency	1 Latent_genotype
Association The consensus genotype whose frequency is given.	Card. Role Name	* Genomic_genotype_population_frequency	0..1 Consensus_genomic_genotype
Association	Card. Role Name	0..1 Page	* Genomic_genotype_population_frequency
Generalization	Card. Role Name	 Genomic_genotype_population_frequency	 Frequency

Genomic_haplotype

Type: **Class Genomic_observation**

Package: PAGE-OM

Document:

A set of Genomic_alleles across an equal number of Genomic_variations in a single chromosome and in a single individual. The Genomic_haplotype is derived from a set of Consensus_genomic_genotype. For each Genomic_variation, the haplotype contains one and only one Genomic_allele. Furthermore, the Genomic_alleles are required to be in phase on the individual, meaning that they are located on the same contiguous strand of DNA. Synonym: Haplotype.

Connections

Connector		Source	Target
Association Alleles which are part of the haplotype i.e. are on same phase.	Card. Role Name	* Genomic_allele	0..1 Genomic_haplotype
Association Provides information on haplotype derivation method.	Card. Role Name	* Genomic_haplotype	0..1 Haplotype_derivation_methods
Association Haplotype set	Card. Role Name	* Genomic_haplotype	0..1 Set_of_haplotypes
Generalization	Card. Role Name	 Genomic_haplotype	 Genomic_observation
Association Sample from which genomic haplotypes are measured.	Card. Role Name	* Genomic_haplotype	* Molecular_sample
Association Haplotypes which are part of the map.	Card. Role Name	1 Haplotype_map	* Genomic_haplotype

Connector		Source	Target
Association	Card. Role Name	0..1 Page	* Genomic_haplotype
Association Haplotype whose frequency is reported.	Card. Role Name	* Genomic_haplotype_population_frequency	1 Genomic_haplotype
Association Map of haplotype. Provides information on coordinate system.	Card. Role Name	1 Genomic_haplotype	* Haplotype_map

Genomic_haplotype_population_frequency

Type: **Class Frequency**

Package: PAGE-OM

Document:

Frequency of a Genomic_haplotype in a Panel.

Connections

Connector		Source	Target
Association	Card. Role Name	0..1 Page	* Genomic_haplotype_population_frequency
Association Haplotype whose frequency is reported.	Card. Role Name	* Genomic_haplotype_population_frequency	1 Genomic_haplotype

Connector		Source	Target
Generalization	Card. Role Name	Genomic_haplotype_population_frequency	Frequency

Genomic_observation

Type: **Class Identifiable**
Package: PAGE-OM

Document:
Genomic observation

Connections

Connector		Source	Target
Generalization	Card. Role Name	Genomic_haplotype	Genomic_observation
Generalization	Card. Role Name	Genomic_allele	Genomic_observation
Generalization	Card. Role Name	Frequency	Genomic_observation
Association All genomic observations which were used in the experiment.	Card. Role Name	0..1 Genotype_phenotype_correlation_experiment	* Genomic_observation
Association Combination of	Card. Role Name	*	*

Connector		Source	Target
Genomic_observations determined for each Observation_target. For example actual genotypes over Molecular_samples		Experiment_result	Genomic_observation
Generalization	Card. Role Name	Genomic_observation	Identifiable
Generalization	Card. Role Name	Set_of_haplotypes	Genomic_observation
Generalization	Card. Role Name	Assayed_genomic_genotype	Genomic_observation

Genomic_variation

Type: **Class Locatable**

Package: PAGE-OM

Document:

A variable site in a Reference_genomic_landmark sequence. Synonyms: polymorphic site, marker, (Genomic_polymorphism in SNP-PML).

Connections

Connector		Source	Target
Generalization Another genomic variation close enough to affect the	Card. Role Name	Neighbour_variation	Genomic_variation

Connector		Source	Target
primer design.			
Aggregation *possible_alleles* The relationship to gives information on possible sequence variations attached to the locus (as defined by flanking sequences).	Card. Role Name	* Genomic_allele	0..1 Genomic_variation
Association Genomic variations which are associated to the location.	Card. Role Name	0..1 Reference_genomic_location	* Genomic_variation
Association Genomic landmark (part of sequence assembly)	Card. Role Name	0..1 Genomic_variation	0..1 Reference_genomic_landmark
Association Genomic variation site for which assay was desinged	Card. Role Name	* Variation_assay	1 Genomic_variation
Association Location on reference genome.	Card. Role Name	0..1 Genomic_variation	* Reference_genomic_location
Association Genomic variation site of the genic variation.	Card. Role Name	0..1 Genic_variation	0..1 is_treated_as Genomic_variation
Association Ordered location of variation (e.g. cytogenetic location)	Card. Role Name	* Ordered_location	0..1 Genomic_variation
Association *Without assay details* Used when assay level information is not know or is not important. There can be more than	Card. Role Name	* Assayed_genomic_genotype	0..1 Genomic_variation

Connector		Source	Target
one genotypes per genomic variation, done over different samples.			
Association Variability of variation site.	Card. Role Name	0..1 Genomic_variation	* Heterozygosity
Association	Card. Role Name	0..1 Genomic_variation	* Neighbour_variation
Generalization	Card. Role Name	 Genomic_variation	 Locatable
Association	Card. Role Name	0..1 Page	* Genomic_variation
Association *Without assay details* Genomic variation site (for example marker).	Card. Role Name	* Frequency	0..1 Genomic_variation
Association Combination of Genomic_variations studied for each Observation_target. For example marker dimension in GWA studies.	Card. Role Name	* Experiment_result	* Genomic_variation
Association *Genomic variations used in experiment*	Card. Role Name	* Genotype_phenotype_correlation_experiment	* Genomic_variation

Attributes

Attribute	Lower bound	Notes
flank_down	*0*	Downstream flanking sequence (at least 25 residues, if

Attribute	Lower bound	Notes
		possible)
flank_up	*0*	Upstream flanking sequence (at least 25 residues, if possible).
is_mutation	*0*	Proven phenotype change causing mutation.
repeat_unit	*0*	If type is 'microsatellite', gives the repeat unit, e.g. "CA"
type	*0*	The type of the polymorphism. E.g. SNP, microsatellite, indel, translocation,...
validation_status	*0*	Validation status, e.g. "Proven", "Suspected"

Genotype_phenotype_correlation_experiment

Type: **Class Identifiable**

Package: PAGE-OM

Document:

Family or case control based association study experiment. Represents set of experiment sub-sections that would normally be listed in the results section in manuscripts.

Connections

Connector		Source	Target
Association Experiments which are associated to meta experiment.	Card. Role Name	0..1 metaexperiment Genotype_phenotype_correlation_experiment	* subexperiment Genotype_phenotype_correlation_experiment
Association Conclusion of experiment	Card. Role Name	0..1 Genotype_phenotype_correlation_experiment	0..1 Conclusion
Association	Card. Role Name	0..1	*

Connector		Source	Target
		Page	Genotype_phenotype_correlation_experiment
Association *Result* Experiments results which are part of the correlation experiment	Card. Role Name	* Genotype_phenotype_correlation_experiment	* Experiment_result
Aggregation Correlation experiments which are part (or used in) study.	Card. Role Name	* Genotype_phenotype_correlation_experiment	* Association_study
Association All genomic observations which were used in the experiment.	Card. Role Name	0..1 Genotype_phenotype_correlation_experiment	* Genomic_observation
Association *Observation targets used in experiment*	Card. Role Name	* Genotype_phenotype_correlation_experiment	* Abstract_observation_target
Association Optional hypothesis of experiment	Card. Role Name	0..1 Genotype_phenotype_correlation_experiment	0..1 Hypothesis
Association *Genomic variations used in experiment*	Card. Role Name	* Genotype_phenotype_correlation_experiment	* Genomic_variation
Generalization	Card. Role Name	 Genotype_phenotype_correlation_experiment	 Identifiable

Attributes

Attribute	Lower bound	Notes

Attribute	Lower bound	Notes
objective	*0*	Objective of experiment
outcome	*0*	A free text description summarizing outcome of all experiment results in this correlation experiment
study_id	*0*	Identifier of original study. Can be used in cases where experiment was originally done for different study.
type	*0*	Type of experiment

Haplotype_block

Type: **Class Set_of_haplotypes**

Package: PAGE-OM

Document:

Large (spanning a few kb to >100 kb) blocks of Genomic_alleles in linkage disequilibrium (LD) and a few haplotypes per
block, separated by regions of recombination.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Haplotype_block	Set_of_haplotypes
Association	Card. Role Name	0..1 Page	* Haplotype_block

Haplotype_derivation_methods

Type: **Class Methodology**

Package: PAGE-OM

Document:

Association class describing methods used to derive Genomic_haplotypes from Consensus_genomic_genotypes.

Connections

Connector		Source	Target
Association Provides information on haplotype derivation method.	Card. Role Name	* Genomic_haplotype	0..1 Haplotype_derivation_methods
Association Consensus genotypes used to derive the haplotype.	Card. Role Name	* Haplotype_derivation_methods	0..1 Consensus_genomic_genotype
Association	Card. Role Name	0..1 Page	* Haplotype_derivation_methods
Generalization	Card. Role Name	 Haplotype_derivation_methods	 Methodology

Haplotype_map

Type: **Class Physical_map**

Package: PAGE-OM

Document:

Map of haplotypes. Features include: Block length distribution, measures of block variability,

relative proportions of common haplotypes, block coverage of chromosomes and/or genome. LD and other values between haplotypes, markers, alleles.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Haplotype_map	Physical_map
Association Set of haplotypes which are associated with the map.	Card. Role Name	* Haplotype_map	* Set_of_haplotypes
Association Haplotypes which are part of the map.	Card. Role Name	1 Haplotype_map	* Genomic_haplotype
Association	Card. Role Name	0..1 Page	* Haplotype_map
Association Map of haplotype. Provides information on coordinate system.	Card. Role Name	1 Genomic_haplotype	* Haplotype_map

Heterozygosity

Type: **Class Frequency**

Package: PAGE-OM

Document:

Heterozygosity (Heterozygosity) is a measure of observed variability of a polymorphic site (Genomic_variation) in a sub-population (Panel).

Connections

Connector		Source	Target
Generalization Frequency of heterozygote alleles.	Card. Role Name	 Heterozygosity	 Frequency
Association Variability of variation site.	Card. Role Name	0..1 Genomic_variation	* Heterozygosity
Association	Card. Role Name	0..1 Page	* Heterozygosity

Hypothesis

Type: **Class Identifiable**

Package: PAGE-OM

Document:

Free text description of hypothesis of study.

Connections

Connector		Source	Target
Generalization	Card. Role Name	 Hypothesis	 Identifiable
Association	Card. Role Name	0..1	*

Connector		Source	Target
		Page	Hypothesis
Association Optional hypothesis of experiment	Card. Role Name	0..1 Genotype_phenotype_correlation_experiment	0..1 Hypothesis

Attributes

Attribute	Lower bound	Notes
description	*0*	Description of hypothesis

Individual

Type: **Class Abstract_population**

Package: PAGE-OM

Document:

A single member of a species, where a species is an accessioned taxon defined by a public database, and the individual is accessioned in a public or private database. Synonym: "inbred strain" in homozygous lineages.

Connections

Connector		Source	Target
Association Individuals which are part of the Panel	Card. Role Name	* Individual	* Panel
Generalization	Card. Role Name	 Individual	 Abstract_population
Association	Card. Role	0..1	*

Connector		Source	Target
	Name	Page	Individual

Attributes

Attribute	Lower bound	Notes
birth_date	*0*	Date of birth of the individual. May be better abbreviated to birth year to protect the privacy of the individual
death_date	*0*	Date of death of the individual. May be better abbreviated to plain year to protect the privacy of the individual.
father_id	*0*	Id of the father to allow building of pedigrees.
gender	*0*	Recommended values are 'unknown', 'male' and 'female'. Additional values can be used to denote unusual karyotypes.
mother_id	*0*	Id of the mother to allow building of pedigrees.

Integer_value

Type: **Class Numeric_value**

Package: PAGE-OM

Document:

Value of type integer

Connections

Connector		Source	Target
Generalization	Card. Role Name	Integer_value	Numeric_value

Attributes

Attribute	Lower bound	Notes
value	*1*	Integer value

Latent_genotype

Type: **Class Identifiable**

Package: PAGE-OM

Document:

.assays_Potentially existing genotypes on specific site that could be observed by Variation genotype _tsample generates a single Laten_assay on one Molecular_Application of a Variation this depends on ploidy level in case of)specifications _genotype _which has one or more Latent .(allele_Genomic

assay can have only one type of _genotypes associated to one instance of a Variation_Latent .feature_d by Definingas define ,specifications_genotype_Latent

This class is a holder for one or many observable variation objects
genotype is used to attach possible variations to _Latent .(specification_genotype_Latent)
and marker loci (assay_Variation)variation assays ,(genotype_genomic_Assayed)measurements .(variation_Genomic)

Connections

Connector		Source	Target
Generalization	Card. Role Name	Latent_genotype	Identifiable
AssociationClass (Generic) allleles which specify the genotype.	Card. Role Name	* Latent_genotype	* Latent_genotype_specification
Association *measured genotype*	Card. Role Name	* Assayed_genomic_genotype	0..1 Latent_genotype

Connector		Source	Target
Association *detectable genotypes* All genotypes that can be detected. Gives information on all measurable variations measurable (by this assay) on the variation site.	Card. Role Name	* Variation_assay	* Latent_genotype
Association *measured genotype*	Card. Role Name	* Genomic_genotype_population_frequency	1 Latent_genotype
Association	Card. Role Name	0..1 Page	* Latent_genotype
Dependency *derived from raw data* Latent genotypes do depend on actual measured values (raw data).	Card. Role Name	0..1 Latent_genotype	0..1 Abstract_value

Latent_genotype_specification

Type: **Class**

Package: PAGE-OM

Document:

Abstract super class of observable variation objects, like alleles, melting temperatures (Melting_temperature), band sizes (Band_size). The class is an extension point to other kinds of variations.

Connections

Connector		Source	Target
AssociationClass (Generic) allleles which specify the genotype.	Card. Role Name	* Latent_genotype	* Latent_genotype_specification
Generalization	Card. Role Name	 Melting_temperature	 Latent_genotype_specification
Generalization	Card. Role Name	 Band_size	 Latent_genotype_specification
Generalization	Card. Role Name	 Genomic_allele	 Latent_genotype_specification
Association Gives information on allele calling.	Card. Role Name	0..1 Defining_feature	0..1 Latent_genotype_specification

Lifestyle_feature

Type: **Class Observable_feature**

Package: PAGE-OM

Document:

Way of life of an individual or panel

Connections

Connector		Source	Target
Generalization	Card. Role		

Connector		Source	Target
	Name	Lifestyle_feature	Observable_feature
Association	Card. Role Name	0..1 Page	* Lifestyle_feature

Location_on_plate

Type: **Class**

Package: PAGE-OM

Document:

X,Y plate_positions (wells). Numbering starts from one. Each well can contain one or more observation targets (molecular samples) prepared for measurement using one or more variation assays (e.g. assay multiplexing Assay_set). Note: These are optional laboratory specific details (Sample and Assay information is in Assayed_genomic_genotype)

Connections

Connector		Source	Target
Association *wells* Well positions.	Card. Role Name	0..1 Plate	* Location_on_plate
Association Observation target, sample, on well location.	Card. Role Name	* Location_on_plate	* Abstract_observation_target
Association	Card. Role Name	0..1 Page	* Location_on_plate

Attributes

Attribute	Lower bound	Notes
x	*1*	x coordinate of plate.
y	*1*	y coordinate of plate

Melting_temperature

Type: **Class Latent_genotype_specification**

Package: PAGE-OM

Document:

The temperature at which DNA goes from a double-stranded to a single-stranded state. Unit of temperature is Celsius.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Melting_temperature	Latent_genotype_specification
Association	Card. Role Name	0..1 Page	* Melting_temperature
Association Value of melting temperature.	Card. Role Name	0..1 Melting_temperature	1 Abstract_value

Molecular_sample

Type: **Class Abstract_observation_target**

Package: PAGE-OM

Document:

A sample from an Individual or from a Panel defining the molecule and tissue/cell used (Anatomic_locations) in the Variation_assay. Synonym: Sample of individual.

Connections

Connector		Source	Target
Association Source of sample(s) from which sample is derived.	Card. Role Name	0..1 Molecular_sample	* sourcesample Molecular_sample
Generalization	Card. Role Name	 Molecular_sample	 Abstract_observation_target
Association Source of sample. For example individual or panel (pool of individuals).	Card. Role Name	* Molecular_sample	0..1 Abstract_population
Association Sample from which genomic haplotypes are measured.	Card. Role Name	* Genomic_haplotype	* Molecular_sample
Association Anatomic location from which sample was taken.	Card. Role Name	* Molecular_sample	* Anatomic_location
Association	Card. Role Name	0..1 Page	* Molecular_sample

Attributes

Attribute	Lower bound	Notes
molecule	*0*	The molecule (RNA, DNA, protein) used in the assay.

Multi_vartiation_assay

Type: **Class Variation_assay**

Package: PAGE-OM

Document:

Multi_variation_assay is a collection of assays which may be used simultaneously. Examples would be multiplex assays, micro-array based assays, or a panel of single-plex assays that share some common feature or purpose.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Multi_vartiation_assay	Variation_assay
Aggregation Assays which are part of multivariation assay (for example assay which can be applied on different variation sites).	Card. Role Name	* Variation_assay	0..1 Multi_vartiation_assay
Association	Card. Role Name	0..1 Page	* Multi_vartiation_assay

Neighbour_variation

Type: **Class Genomic_variation**

Package: PAGE-OM

Document:

Another Genomic_variation close enough to affect the primer design.

Connections

Connector		Source	Target
Generalization Another genomic variation close enough to affect the primer design.	Card. Role Name	Neighbour_variation	Genomic_variation
Association	Card. Role Name	0..1 Genomic_variation	* Neighbour_variation
Association	Card. Role Name	0..1 Page	* Neighbour_variation

Numeric_value

Type: **Class Value**

Package: PAGE-OM

Document:

Numeric value

Connections

Connector		Source	Target
Generalization	Card. Role Name	Float_value	Numeric_value
Generalization	Card. Role Name	Integer_value	Numeric_value
Generalization	Card. Role Name	Error	Numeric_value
Association Error of numeric value.	Card. Role Name	0..1 Numeric_value	0..1 Error
Generalization	Card. Role Name	Numeric_value	Value

Observable_feature

Type: **Class Abstract_observable_feature**

Package: PAGE-OM

Document:

Measurable feature of observable (e.g. size of nose)

Connections

Connector		Source	Target
Generalization	Card. Role Name	Observable_feature	Abstract_observable_feature

Connector		Source	Target
Generalization	Card. Role Name	Environment_feature	Observable_feature
Association Optional generic categories where observable feature belongs to.	Card. Role Name	* Observable_feature	* Observable_feature_category
Association Observation method(s) used to measure the feature.	Card. Role Name	0..1 Observable_feature	* Observation_method
Generalization	Card. Role Name	Phenotype_feature	Observable_feature
Generalization	Card. Role Name	Lifestyle_feature	Observable_feature

Observable_feature_category

Type: **Class Identifiable**

Package: PAGE-OM

Document:

All features considered by this model can be categorized by using this class. The category should be expressed by an ontology term.

Connections

Connector		Source	Target
Association	Card. Role Name	0..1	* subcategory

Connector		Source	Target
		Observable_feature_category	Observable_feature_category
Generalization	Card. Role Name	 Phenotype_disease_category	 Observable_feature_category
Association Optional generic categories where observable feature belongs to.	Card. Role Name	* Observable_feature	* Observable_feature_category
Association	Card. Role Name	0..1 Page	* Observable_feature_category
Generalization	Card. Role Name	 Observable_feature_category	 Identifiable

Observation_method

Type: **Class Identifiable**

Package: PAGE-OM

Document:

Observable features can be measured by different methods. This class specifies which method has been used. For example, a method can be usage of a ruler or filling a questionnaire.

Connections

Connector		Source	Target
Association Observation method(s) used to measure the feature.	Card. Role Name	0..1 Observable_feature	* Observation_method

Connector		Source	Target
Association	Card. Role Name	0..1 Page	* Observation_method
Association Values of measurements done by the method	Card. Role Name	0..1 Observation_method	* Observed_value
Generalization	Card. Role Name	 Observation_method	 Identifiable

Observed_value

Type: **Class Value**

Package: PAGE-OM

Document:

Observation done at specific point in time.

Connections

Connector		Source	Target
Association Accuracy code for value.	Card. Role Name	0..1 Observed_value	0..1 Time_accuracy
Association	Card. Role Name	0..1 Page	* Observed_value
Association Actual value of observation	Card. Role Name	0..1 Observed_value	1 value Value

Connector		Source	Target
Association *Has value* Observed values (of observable features) measured on the target	Card. Role Name	* Observed_value	1 target Abstract_observation_target
Generalization	Card. Role Name	 Observed_value	 Value
Association Combination of Observable_values determined for each Observation_target. For example phenotype in GWA studies.	Card. Role Name	* Experiment_result	* Observed_value
Association Values of measurements done by the method	Card. Role Name	0..1 Observation_method	* Observed_value

Attributes

Attribute	Lower bound	Notes
time	*0*	Time of observation

Page

Type: **Class**

Package: PAGE-OM

Document:

This class does not contain any scientific meaning. Its main purpose is to be the root element for the situations where this specification is used for data exchange formats (e.g. xml-schema). Therefore, it has optional direct associations to all important classes so that implementations can exchange only relevant data.

Connections

Connector		Source	Target
Association	Card. Role Name	0..1 Page	* Algorithm
Association	Card. Role Name	0..1 Page	* Plate
Association	Card. Role Name	0..1 Page	* Genomic_genotype_population_frequency
Association	Card. Role Name	0..1 Page	* Melting_temperature
Association	Card. Role Name	0..1 Page	* Multi_vartiation_assay
Association	Card. Role Name	0..1 Page	* Observable_feature_category
Association	Card. Role Name	0..1 Page	* Observation_method
Association	Card. Role Name	0..1 Page	* Observed_value
Association	Card. Role Name	0..1 Page	* Phenotype_feature
Association	Card. Role	0..1	*

Connector		Source	Target
	Name	Page	Lifestyle_feature
Association	Card. Role Name	0..1 Page	* Heterozygosity
Association	Card. Role Name	0..1 Page	* Frequency_set
Association	Card. Role Name	0..1 Page	* Run
Association	Card. Role Name	0..1 Page	* Assayed_genomic_genotype
Association	Card. Role Name	0..1 Page	* Functional_change
Association	Card. Role Name	0..1 Page	* Gene_based_haplotype
Association	Card. Role Name	0..1 Page	* Genic_variation
Association	Card. Role Name	0..1 Page	* Genomic_allele
Association	Card. Role Name	0..1 Page	* Genomic_allele_population_frequency
Association	Card. Role Name	0..1 Page	* Genotype_phenotype_correlation_experiment

Connector		Source	Target
Association	Card. Role Name	0..1 Page	* Band_size
Association	Card. Role Name	0..1 Page	* Location_on_plate
Association	Card. Role Name	0..1 Page	* Association_study
Association	Card. Role Name	0..1 Page	* Genomic_haplotype
Association	Card. Role Name	0..1 Page	* Conclusion
Association	Card. Role Name	0..1 Page	* Environment_feature
Association	Card. Role Name	0..1 Page	* Evidence
Association	Card. Role Name	0..1 Page	* Hypothesis
Association	Card. Role Name	0..1 Page	* Experiment_result
Association	Card. Role Name	0..1 Page	* Person
Association	Card. Role Name	0..1 Page	* Genomic_gene_structure

Connector		Source	Target
Association	Card. Role Name	0..1 Page	* Db_xref
Association	Card. Role Name	0..1 Page	* Exon
Association	Card. Role Name	0..1 Page	* Genetic_location
Association	Card. Role Name	0..1 Page	* Genetic_map
Association	Card. Role Name	0..1 Page	* Genomic_reference_allele
Association	Card. Role Name	0..1 Page	* Geographic_location
Association	Card. Role Name	0..1 Page	* Cytogenetic_location
Association	Card. Role Name	0..1 Page	* Organization
Association	Card. Role Name	0..1 Page	* Contributor
Association	Card. Role Name	0..1 Page	* Reference_genomic_assembly
Association	Card. Role Name	0..1 Page	* Reference_genomic_landmark

Connector		Source	Target
Association	Card. Role Name	0..1 Page	* Subject_descriptor
Association	Card. Role Name	0..1 Page	* Taxon
Association	Card. Role Name	0..1 Page	* Bibliographic_reference
Association	Card. Role Name	0..1 Page	* Bibref_description
Association	Card. Role Name	0..1 Page	* Journal
Association	Card. Role Name	0..1 Page	* Service
Association	Card. Role Name	0..1 Page	* Oligo
Association	Card. Role Name	0..1 Page	* Panel
Association	Card. Role Name	0..1 Page	* Genomic_haplotype_population_frequency
Association	Card. Role Name	0..1 Page	* Genomic_variation
Association	Card. Role Name	0..1	*

Connector		Source	Target
		Page	Haplotype_block
Association	Card. Role Name	0..1 Page	* Haplotype_derivation_methods
Association	Card. Role Name	0..1 Page	* Haplotype_map
Association	Card. Role Name	0..1 Page	* Individual
Association	Card. Role Name	0..1 Page	* Latent_genotype
Association	Card. Role Name	0..1 Page	* Cytogenetic_map
Association	Card. Role Name	0..1 Page	* Neighbour_variation
Association	Card. Role Name	0..1 Page	* Reference_genomic_location
Association	Card. Role Name	0..1 Page	* Structural_change
Association	Card. Role Name	0..1 Page	* Transcription_change
Association	Card. Role Name	0..1 Page	* Translation_change
Association	Card. Role Name	0..1	*

Connector		Source	Target
		Page	Variation_assay
Association	Card. Role Name	0..1 Page	* Anatomic_location
Association	Card. Role Name	0..1 Page	* Annotation
Association	Card. Role Name	0..1 Page	* CDS
Association	Card. Role Name	0..1 Page	* Molecular_sample
Association	Card. Role Name	0..1 Page	* Source

Panel

Type: **Class Abstract_population**

Package: PAGE-OM

Document:

A set of samples from individuals drawn from the same species and used for genetic studies. A panel must be identifiable with a list of accessioned individuals, if possible. Panel can have subpanels. Synonym: SampleSet, Sample from population(s), “Plate” in Coriel sense.

Connections

Connector		Source	Target

Connector		Source	Target
Association Individuals which are part of the Panel	Card. Role Name	* Individual	* Panel
Association Panel can be made from subpanels	Card. Role Name	0..1 Panel	0..1 subpanel Panel
Association	Card. Role Name	0..1 Page	* Panel
Generalization	Card. Role Name	 Panel	 Abstract_population

Attributes

Attribute	Lower bound	Notes
count_unit	*0*	Values are 'chromosome' and 'individual'. Default is 'individual'.
pooled	*0*	True if accessioned individuals are not available.
size	*0*	The size of the sample. Note that the count_unit field affects how this value is interpreted.
type	*0*	Optional identifier of the panel category: e.g. plate, family, population sample

Phenotype_disease_category

Type: **Class Observable_feature_category**

Package: PAGE-OM

Document:

Specialized category of features representing diseases.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Phenotype_disease_category	Observable_feature_category

Phenotype_feature

Type: **Class Observable_feature**

Package: PAGE-OM

Document:

Observable part of the structure, function or behavior of a living organism.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Phenotype_feature	Observable_feature
Association	Card. Role Name	0..1 Page	* Phenotype_feature

Plate

Type: **Class Identifiable**

Package: PAGE-OM

Document:

A sample holder, for example a microtiter plate used in one or many runs, represented by instances of Run. Samples, represented by instances of Molecular_sample, are positioned on the plate using instances of Location_on_plate.

Connections

Connector		Source	Target
Association *wells* Well positions.	Card. Role Name	0..1 Plate	* Location_on_plate
Generalization	Card. Role Name	 Plate	 Identifiable
Association Plate used in experiment run.	Card. Role Name	* Run	0..1 Plate
Association	Card. Role Name	0..1 Page	* Plate

Reference_genomic_location

Type: **Class Location**

Package: PAGE-OM

Document:

A location within a Reference_genomic_landmark. Attributes of the location are the Reference_genomic_assembly and/or
the Reference_genomic_landmark, the start and end range and strand of the feature relative to the Reference_genomic_landmark.

Connections

Connector		Source	Target
Association Landmark of the location. Identified region on assembly.	Card. Role Name	0..1 Reference_genomic_landmark	* Reference_genomic_location
Generalization	Card. Role Name	 Reference_genomic_location_in_assembly	 Reference_genomic_location
Generalization	Card. Role Name	 Reference_genomic_location	 Location
Association Reference assembly on which locations are defined (coordinate system)	Card. Role Name	0..1 Reference_genomic_assembly	* Reference_genomic_location
Association Genomic variations which are associated to the location.	Card. Role Name	0..1 Reference_genomic_location	* Genomic_variation
Association Location on reference genome.	Card. Role Name	0..1 Genomic_variation	* Reference_genomic_location
Association	Card. Role Name	0..1 Page	* Reference_genomic_location

Attributes

Attribute	Lower bound	Notes
end	*0*	End of the location in the reference sequence.
start	*1*	Start of the location in the reference sequence.

Attribute	Lower bound	Notes
strand unknown	*0*	Orientation of the feature in the reference sequence. One of 'forward', 'reverse', 'unknown'. Defaults to 'unknown'.

Reference_genomic_location_in_assembly

Type: **Class Reference_genomic_location**

Package: PAGE-OM

Document:

A location in one chromosome of a reference genomic assembly. Instead of the reference sequence being an accessioned
sequence, it is a versioned assembly.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Reference_genomic_location_in_assembly	Reference_genomic_location

Attributes

Attribute	Lower bound	Notes
chromosome_name	*0*	Name of the chromosome in the assembly.

Run

Type: **Class Abstract_run**

Package: PAGE-OM

Document:

The class contains information on measurement of samples, done on a physical device connected to plate.. This information includes time of execution, name of instrument, etc.

Connections

Connector		Source	Target
Association Plate used in experiment run.	Card. Role Name	* Run	0..1 Plate
Association	Card. Role Name	0..1 Page	* Run
Generalization	Card. Role Name	 Run	 Abstract_run

Attributes

Attribute	Lower bound	Notes
instrument	*0*	Name of the instrument
run_date	*0*	Date of run.

Set_of_haplotypes

Type: **Class Genomic_observation**

Package: PAGE-OM

Document:

An extension point for collections of haplotypes.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Haplotype_block	Set_of_haplotypes
Association Haplotype set	Card. Role Name	* Genomic_haplotype	0..1 Set_of_haplotypes
Association Set of haplotypes which are associated with the map.	Card. Role Name	* Haplotype_map	* Set_of_haplotypes
Generalization	Card. Role Name	Gene_based_haplotype	Set_of_haplotypes
Generalization	Card. Role Name	Set_of_haplotypes	Genomic_observation

String_value

Type: **Class Value**

Package: PAGE-OM

Document:

Value of type string

Connections

Connector		Source	Target
Generalization	Card. Role Name	String_value	Value

Attributes

Attribute	Lower bound	Notes
value	*1*	Actual value

Structural_change

Type: **Class Gene_product_change**

Package: PAGE-OM

Document:

Change in the 3D structure of the polypeptide chain.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Structural_change	Gene_product_change
Association Structural change which has impact on function of protein.	Card. Role Name	* Functional_change	* affects_function Structural_change
Association Translational change which leads to a structural change.	Card. Role Name	* Structural_change	* affects_structure Translation_change
Association	Card.	0..1	*

Connector		Source	Target
	Role Name	Page	Structural_change

Study

Type: **Class Identifiable**

Package: PAGE-OM

Document:

An extension point for adding other kind of studies in the future.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Association_study	Study
Generalization	Card. Role Name	Study	Identifiable

Attributes

Attribute	Lower bound	Notes
abstract	*0*	Abstract
acknowledgements	*0*	Acknowledgements
background	*0*	Background information
conclusions	*0*	Summarizing conclusion for all experiments in this study

Attribute	Lower bound	Notes
key_results	*0*	Key findings
limitations	*0*	Limitations
objectives	*0*	Summarizing objective for all experiments in this study
source_of_bias	*0*	Possible source of bias
study_design	*0*	Study design
study_power	*0*	Power of study
study_size_reason	*0*	Reason for study size
submission_date	*0*	Submission date of study
title	*0*	Title of study
update_date	*0*	Date when study is updated

Time_accuracy

Type: **Class**

Package: PAGE-OM

Document:

Accuracy code contains information on incompleteness of time of measurement or information on reason why the time of measurement is unknown or incomplete.

Connections

Connector		Source	Target
Association Accuracy code for value.	Card. Role Name	0..1 Observed_value	0..1 Time_accuracy

Attributes

Attribute	Lower bound	Notes
code	*1*	Accuracy code as defined in specific ontology

Transcription_change

Type: **Class Residue_change**

Package: PAGE-OM

Document:

Change in the quality or quantity of the mature RNA product.

Connections

Connector		Source	Target
Association Transcription change which affects translation.	Card. Role Name	* Translation_change	* affects_protein Transcription_change
Association Genic variation which influences on the transcript change.	Card. Role Name	* Transcription_change	* affects_transcript Genic_variation
Association	Card. Role Name	0..1 Page	* Transcription_change
Generalization	Card. Role Name	 Transcription_change	 Residue_change

Attributes

Attribute	Lower bound	Notes
changed_codon	*0*	The new codon in the transcript, if applicable.
codon_position	*0*	The first affected nucleotide in the codon. Values are: 1, 2 or 3.
original_codon	*0*	The affected codon in the transcript.

Translation_change

Type: **Class Residue_change**

Package: PAGE-OM

Document:

Change in the quality or quantity of (predicted) polypeptide chain (2D).

Connections

Connector		Source	Target
Generalization	Card. Role Name	Translation_change	Residue_change
Association Translation which affects function of protein.	Card. Role Name	* Functional_change	* Translation_change
Association Translational change which leads to a structural change.	Card. Role Name	* Structural_change	* affects_structure Translation_change
Association Transcription change which affects translation.	Card. Role Name	* Translation_change	* affects_protein Transcription_change
Association	Card.	0..1	*

Connector		Source	Target
	Role Name	Page	Translation_change

Unit

Type: **Class**

Package: PAGE-OM

Document:

Unit of value. Unit is defined using ontology term

Connections

Connector		Source	Target
Aggregation Unit of value	Card. Role Name	0..1 Unit	0..1 Value

Attributes

Attribute	Lower bound	Notes
type	*1*	Type of unit

Value

Type: **Class Abstract_value**

Package: PAGE-OM

Document:

Abstract class. Extension point for Value implementations. Value model is based on concept developed in Generation Challenge Program:
http://pantheon.generationcp.org/demeter/Values.html

Connections

Connector		Source	Target
Aggregation *supporting_evidence* Values supporting the evidence.	Card. Role Name	* Evidence_code	0..1 Value
Generalization	Card. Role Name	 Value	 Abstract_value
Aggregation	Card. Role Name	* Value	0..1 Enum
Generalization	Card. Role Name	 Value_range	 Value
Generalization	Card. Role Name	 String_value	 Value
Generalization	Card. Role Name	 Boolean	 Value
Aggregation Constraints	Card. Role Name	* Constraint	0..1 Value
Association Score of evidence (e.g. p-value)	Card. Role Name	0..1 Evidence	0..1 score *score of value* Value
Association Actual value of	Card. Role	0..1	1

Connector		Source	Target
observation	Name	Observed_value	value Value
Generalization	Card. Role Name	Experiment_result	Value
Aggregation Unit of value	Card. Role Name	0..1 Unit	0..1 Value
Generalization	Card. Role Name	Numeric_value	Value
Generalization	Card. Role Name	Observed_value	Value
Generalization	Card. Role Name	Enum	Value
Association Value can be composed of inferred from other values.	Card. Role Name	0..1 Value	* hierarchical_value Value

Value_range

Type: **Class Value**

Package: PAGE-OM

Document:

Inclusive value range

Connections

Connector		Source	Target
Generalization	Card. Role Name	Value_range	Value

Attributes

Attribute	Lower bound	Notes
max	*1*	maximum value
min	*1*	minimum value

Variation_assay

Type: **Class Methodology**

Package: PAGE-OM

Document:

An experimental lab protocol and set of reagents for detecting the Genomic_alleles of Genomic_variations carried by an individual or a panel of individuals. Synonym: Assay. Non instrument part of the experiment - same assay can be used in different instrument runs

Connections

Connector		Source	Target
Generalization	Card. Role Name	Multi_vartiation_assay	Variation_assay
Generalization	Card. Role Name	Variation_assay	Methodology

Connector		Source	Target
Association *detectable genotypes* All genotypes that can be detected. Gives information on all measurable variations measurable (by this assay) on the variation site.	Card. Role Name	* Variation_assay	* Latent_genotype
Association Genomic variation site for which assay was desinged	Card. Role Name	* Variation_assay	1 Genomic_variation
Aggregation Assays which are part of multivariation assay (for example assay which can be applied on different variation sites).	Card. Role Name	* Variation_assay	0..1 Multi_vartiation_assay
Association PCR primers.	Card. Role Name	* Variation_assay	* Oligo
Association *With assay details* Used when assay level information is needed. Specific variation assay designed for genomic site (locus).	Card. Role Name	* Assayed_genomic_genotype	0..1 Variation_assay
Association	Card. Role Name	0..1 Page	* Variation_assay
Association *With assay details* Assay desinged for the variation site.	Card. Role Name	* Frequency	0..1 Variation_assay

Connector		Source	Target
Association is used when assay level information is needed. (Optional with the "direct" association from Frequency to Genomic_variation) .			

Attributes

Attribute	Lower bound	Notes
description	*0*	Free text description of the assay protocol.

FUGE

Package: PAGE-OM

Document:
FuGE (http://fuge.sourceforge.net/) is a model of the shared components in different functional genomics domains.

Ontology_source

Type: **Class Identifiable**

Package: FUGE

Document:
The source ontology or controlled vocabulary list that ontology terms have been obtained from

Connections

Connector		Source	Target

Connector		Source	Target
Generalization	Card. Role Name	Ontology_source	Identifiable
Association Source or name space of the term.	Card. Role Name	* Ontology_term	0..1 Ontology_source

Attributes

Attribute	Lower bound	Notes
ontology_URI	*0*	A URI is short for Uniform Resource Identifier. A URI is a compact sequence of characters that identifies an abstract or physical resource.

Ontology_term

Type: **Class Identifiable**

Package: FUGE

Document:

Ontology term, defined in more detail in FuGE (http://fuge.sourceforge.net/). This is just a place holder

Connections

Connector		Source	Target
Generalization	Card. Role Name	Ontology_term	Identifiable
Association Source or name space of the term.	Card. Role Name	* Ontology_term	0..1 Ontology_source

Attributes

Attribute	Lower bound	Notes
term	*1*	Ontology term

BasicTypes

Package: PAGE-OM

Document:
XML basic types (http://www.w3.org/TR/xmlschema-2).

anySimpleType

Type: **Class anyType**
Package: BasicTypes

Document:

Connections

Connector		Source	Target
Generalization	Card. Role Name	decimal	anySimpleType
Generalization	Card. Role Name	float	anySimpleType
Generalization	Card. Role Name	anyURI	anySimpleType

Connector		Source	Target
Generalization	Card. Role Name	dateTime	anySimpleType
Generalization	Card. Role Name	time	anySimpleType
Generalization	Card. Role Name	double	anySimpleType
Generalization	Card. Role Name	anySimpleType	anyType
Generalization	Card. Role Name	string	anySimpleType
Generalization	Card. Role Name	date	anySimpleType
Generalization	Card. Role Name	boolean	anySimpleType

anyType

Type: **Class**

Package: BasicTypes

Document:

Connections

Connector		Source	Target
Generalization	Card. Role Name	anySimpleType	anyType

anyURI

Type: **Class anySimpleType**

Package: BasicTypes

Document:

Connections

Connector		Source	Target
Generalization	Card. Role Name	anyURI	anySimpleType

boolean

Type: **Class anySimpleType**

Package: BasicTypes

Document:

Connections

Connector		Source	Target
Generalization	Card. Role Name	boolean	anySimpleType

date

Type: **Class anySimpleType**

Package: BasicTypes

Document:

Connections

Connector		Source	Target
Generalization	Card. Role Name	date	anySimpleType

dateTime

Type: **Class anySimpleType**

Package: BasicTypes

Document:

Connections

Connector		Source	Target
Generalization	Card. Role Name	dateTime	anySimpleType

decimal

Type: **Class anySimpleType**

Package: BasicTypes

Document:

Connections

Connector		Source	Target
Generalization	Card. Role Name	decimal	anySimpleType
Generalization	Card. Role Name	integer	decimal

double

Type: **Class anySimpleType**

Package: BasicTypes

Document:

Connections

Connector		Source	Target
Generalization	Card. Role Name	double	anySimpleType

float

Type: **Class anySimpleType**

Package: BasicTypes

Document:

Connections

Connector		Source	Target
Generalization	Card. Role Name	float	anySimpleType

int

Type: **Class long**

Package: BasicTypes

Document:

Connections

Connector		Source	Target
Generalization	Card. Role Name	int	long
Generalization	Card. Role Name	short	int

integer

Type: **Class decimal**

Package: BasicTypes

Document:

Connections

Connector		Source	Target

Connector		Source	Target
Generalization	Card. Role Name	long	integer
Generalization	Card. Role Name	integer	decimal

long

Type: **Class integer**

Package: BasicTypes

Document:

Connections

Connector		Source	Target
Generalization	Card. Role Name	int	long
Generalization	Card. Role Name	long	integer

short

Type: **Class int**

Package: BasicTypes

Document:

Connections

Connector		Source	Target
Generalization	Card. Role Name	short	int

string

Type: **Class anySimpleType**

Package: BasicTypes

Document:

Connections

Connector		Source	Target
Generalization	Card. Role Name	string	anySimpleType

time

Type: **Class anySimpleType**

Package: BasicTypes

Document:

Connections

Connector		Source	Target
Generalization	Card. Role Name	time	anySimpleType

SNP

Package: PAGE-OM

Document:
Module replicating part of OMG specification SNP (Single Nucleotide Polymorphism).

Anatomic_location

Type: **Class Identifiable**

Package: SNP

Document:

Description of the anatomical location the molecular sample is derived from. Best expressed as a controlled vocabulary of anatomical terms.

Connections

Connector		Source	Target
Association Anatomic location from which sample was taken.	Card. Role Name	* Molecular_sample	* Anatomic_location
Association	Card. Role Name	0..1 Page	* Anatomic_location
Generalization	Card. Role Name	 Anatomic_location	 Identifiable

Annotation

Type: **Class**

Package: SNP

Document:

A free text of name and value pair. Note: name should be Ontology_term

Connections

Connector		Source	Target
Association	Card. Role Name	* Identifiable	* Annotation
Association	Card. Role Name	0..1 Page	* Annotation

Attributes

Attribute	Lower bound	Notes
name	*0*	Name of the annotation feature
value	*0*	Value of annotation feature

CDS

Type: **Class Sequence**

Package: SNP

Document:

Region in a reference sequence that determines the start, end and strand of the gene coding region.

Connections

Connector		Source	Target
Association Coding region which is part of the gene structure.	Card. Role Name	0..1 CDS	0..1 Genomic_gene_structure
Generalization	Card. Role Name	 CDS	 Sequence
Association	Card. Role Name	0..1 Page	* CDS

Contributor

Type: **Class Source**

Package: SNP

Document:

Person who has edited data in the data set. Synonym: curator.

Connections

Connector		Source	Target
Association	Card. Role Name	0..1 Page	* Contributor
Generalization	Card. Role Name	 Contributor	 Source

Attributes

Attribute	Lower bound	Notes
date	*0*	Date of the editing.

Cytogenetic_location

Type: **Class Ordered_location**

Package: SNP

Document:

A location in a cytogenetic map expressed as a band or a range of bands.

Connections

Connector		Source	Target

Connector		Source	Target
Association	Card. Role Name	0..1 Cytogenetic_map	* Cytogenetic_location
Generalization	Card. Role Name	 Cytogenetic_location	 Ordered_location
Association	Card. Role Name	0..1 Page	* Cytogenetic_location

Cytogenetic_map

Type: **Class Ordered_map**

Package: SNP

Document:

Map of cytogenetic bands describing their relative order.

Connections

Connector		Source	Target
Association	Card. Role Name	0..1 Cytogenetic_map	* Cytogenetic_location
Generalization	Card. Role Name	 Cytogenetic_map	 Ordered_map
Association	Card. Role Name	0..1 Page	* Cytogenetic_map

Db_xref

Type: **Class Identifiable**

Package: SNP

Document:

Cross reference between two Identifiables in different databases.

Connections

Connector		Source	Target
Association	Card. Role Name	* Identifiable	* Db_xref
Association	Card. Role Name	0..1 identifier Db_xref	0..1 Bibliographic_reference
Association	Card. Role Name	0..1 Page	* Db_xref
Generalization	Card. Role Name	 Db_xref	 Identifiable

Attributes

Attribute	Lower bound	Notes
db	*0*	The remote database name or abbreviation.
db_version	*0*	Version of the database

Attribute	Lower bound	Notes
field	*0*	Field in the remote database where the UID value is found. E.g. ID or AC in EMBL. Optional
uid	*1*	Unique identifier in the remote database.
url	*0*	Full URL to the cross-referenced entry

Exon

Type: **Class Sequence**

Package: SNP

Document:

Region in a reference sequence that determines the start, end and strand of the transcript that is not spliced away.

Connections

Connector		Source	Target
Association Exons which are part of gene structure.	Card. Role Name	1..* Exon	0..1 Genomic_gene_structure
Association	Card. Role Name	0..1 Page	* Exon
Generalization	Card. Role Name	 Exon	 Sequence

Gene_product_change

Type: **Class Identifiable**

Package: SNP

Document:

Free form description of change in gene product

Connections

Connector		Source	Target
Generalization	Card. Role Name	Structural_change	Gene_product_change
Generalization	Card. Role Name	Gene_product_change	Identifiable
Generalization	Card. Role Name	Residue_change	Gene_product_change
Generalization	Card. Role Name	Functional_change	Gene_product_change

Attributes

Attribute	Lower bound	Notes
label	*0*	Human understandable term for the change. E.g. missense, truncation.

Genetic_location

Type: **Class Ordered_location**

Package: SNP

Document:

A relative location of a marker in a genetic map based on recombination.

Connections

Connector		Source	Target
Association	Card. Role Name	0..1 Genetic_map	* Genetic_location
Generalization	Card. Role Name	 Genetic_location	 Ordered_location
Association	Card. Role Name	0..1 Page	* Genetic_location

Genetic_map

Type: **Class Ordered_map**

Package: SNP

Document:

A map containing the order and non-metric distance between genetic markers for identifiable sequence features. Unit:
centiMorgan.

Connections

Connector		Source	Target
Association	Card. Role	0..1	*

Connector		Source	Target
	Name	Genetic_map	Genetic_location
Generalization	Card. Role Name	Genetic_map	Ordered_map
Association	Card. Role Name	0..1 Page	* Genetic_map

Genomic_reference_allele

Type: **Class Genomic_allele**

Package: SNP

Document:

The DNA sequence of a Genomic_variation as it appears in the public database.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Genomic_reference_allele	Genomic_allele
Association	Card. Role Name	* Reference_genomic_assembly	* present_in Genomic_reference_allele
Association The DNA sequence of a genomic variation as it appears in the .public database	.Card Role Name	0..1 one_of Genomic_reference_allele	0..1 Genomic_allele

Connector		Source	Target
Association	Card. Role Name	0..1 Page	* Genomic_reference_allele

Geographic_location

Type: **Class Identifiable**

Package: SNP

Document:

Location of an individual or population in a geographic map. Locations are expressed in decimal degrees. Northern
latitudes (if north of the equator) and eastern longitudes (if east of GM) have positive values by convention.

Connections

Connector		Source	Target
Association Geographic location of individual or group of individuals (panel)	Card. Role Name	* Abstract_population	0..1 Geographic_location
Association	Card. Role Name	0..1 Page	* Geographic_location
Generalization	Card. Role Name	 Geographic_location	 Identifiable

Attributes

Attribute	Lower bound	Notes
max_latitude	*0*	Upper bound of a geographical location expressed in degrees from the equator.
max_longitude	*0*	Upper bound of a geographical location expressed in degrees from the Greenwich meridian.
min_latitude	*0*	Lower bound of a geographical location expressed in degrees from the equator
min_longitude	*0*	Lower bound of a geographical location expressed in degrees from the Greenwich meridian

Identifiable

Type: **Class**

Package: SNP

Document:

Anything with a unique identifier.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Latent_genotype	Identifiable
Generalization	Card. Role Name	Plate	Identifiable
Aggregation *supporting_source* Supporting evidences: Citations, studies, ontology terms etc	Card. Role Name	* Identifiable	0..1 Evidence
Generalization	Card. Role Name	Subject_descriptor	Identifiable

Connector		Source	Target
Generalization	Card. Role Name	Hypothesis	Identifiable
Generalization	Card. Role Name	Abstract_run	Identifiable
Association	Card. Role Name	* Identifiable	* Annotation
Association	Card. Role Name	* Identifiable	* Publication
Association	Card. Role Name	* Identifiable	* Db_xref
Generalization	Card. Role Name	Bibref_scope	Identifiable
Generalization	Card. Role Name	Journal	Identifiable
Association Source of identifiable.	Card. Role Name	0..1 Identifiable	0..1 Source
Generalization	Card. Role Name	Gene_product_change	Identifiable
Generalization	Card. Role Name	Ontology_source	Identifiable
Generalization	Card. Role Name	Algorithm	Identifiable

Connector		Source	Target
Generalization	Card. Role Name	Abstract_value	Identifiable
Generalization	Card. Role Name	Ontology_term	Identifiable
Generalization	Card. Role Name	Genomic_observation	Identifiable
Generalization	Card. Role Name	Provider	Identifiable
Generalization	Card. Role Name	Map	Identifiable
Generalization	Card. Role Name	Taxon	Identifiable
Generalization	Card. Role Name	Anatomic_location	Identifiable
Generalization	Card. Role Name	Study	Identifiable
Generalization	Card. Role Name	Conclusion	Identifiable
Generalization	Card. Role Name	Observation_method	Identifiable
Generalization	Card. Role Name	Methodology	Identifiable

Connector		Source	Target
Generalization	Card. Role Name	Abstract_observation_target	Identifiable
Generalization	Card. Role Name	Genotype_phenotype_correlation_experiment	Identifiable
Generalization	Card. Role Name	Observable_feature_category	Identifiable
Generalization	Card. Role Name	Abstract_observable_feature	Identifiable
Generalization	Card. Role Name	Geographic_location	Identifiable
Generalization	Card. Role Name	Locatable	Identifiable
Generalization	Card. Role Name	Location	Identifiable
Generalization	Card. Role Name	Db_xref	Identifiable

Attributes

Attribute	Lower bound	Notes
creation_date	*0*	Date of creation of the object.
delete_date	*0*	Date of deletion of the object.

Attribute	Lower bound	Notes
lsid	*1*	Life Science Identifier.
modify_date	*0*	Date of last modification of the object.
name	*0*	Name can be non unique. Display name

Locatable

Type: **Class Identifiable**

Package: SNP

Document:

Abstract class for anything that can be placed on a sequence, i.e. can act as a sequence feature.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Genomic_gene_structure	Locatable
Generalization	Card. Role Name	Genomic_variation	Locatable
Association *Locations* Locations	Card. Role Name	* Locatable	* Location
Generalization	Card. Role Name	Residue_change	Locatable
Generalization	Card. Role Name	Sequence	Locatable

Connector		Source	Target
Generalization	Card. Role Name	Locatable	Identifiable

Location

Type: **Class Identifiable**

Package: SNP

Document:

A position in an abstract map.

Connections

Connector		Source	Target
Association Source location.	Card. Role Name	* Location	0..1 Location
Generalization	Card. Role Name	Reference_genomic_location	Location
Association	Card. Role Name	0..1 Map	* Location
Generalization	Card. Role Name	Ordered_location	Location
Association *Locations* Locations	Card. Role Name	* Locatable	* Location

Connector		Source	Target
Generalization	Card. Role Name	Location	Identifiable

Map

Type: **Class Identifiable**

Package: SNP

Document:

An abstract map.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Physical_map	Map
Association	Card. Role Name	0..1 Map	* Location
Generalization	Card. Role Name	Ordered_map	Map
Generalization	Card. Role Name	Map	Identifiable

Methodology

Type: **Class Identifiable**

Package: SNP

Document:

Abstract class for a laboratory method or a computational protocol.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Variation_assay	Methodology
Generalization	Card. Role Name	Haplotype_derivation_methods	Methodology
Generalization	Card. Role Name	Methodology	Identifiable

Oligo

Type: **Class Sequence**

Package: SNP

Document:

A DNA oligonucleotide used for detection and assay of Genomic_variations in Variation_assays.
Synonym: primer

Connections

Connector		Source	Target
Generalization	Card. Role Name	 Oligo	 Sequence
Association PCR primers.	Card. Role Name	* Variation_assay	* Oligo
Association	Card. Role Name	0..1 Page	* Oligo

Ordered_location

Type: **Class Location**

Package: SNP

Document:

A location on a secondary map that gives ordering information. That map can be a genetic or a cytogenetic map.

Connections

Connector		Source	Target
Generalization	Card. Role Name	 Genetic_location	 Ordered_location
Generalization	Card. Role Name	 Ordered_location	 Location

Connector		Source	Target
Association Ordered location of variation (e.g. cytogenetic location)	Card. Role Name	* Ordered_location	0..1 Genomic_variation
Generalization	Card. Role Name	 Cytogenetic_location	 Ordered_location

Attributes

Attribute	Lower bound	Notes
position	*0*	Location description

Ordered_map

Type: **Class Map**

Package: SNP

Document:

A non-metric map category.

Connections

Connector		Source	Target
Generalization	Card. Role Name	 Cytogenetic_map	 Ordered_map
Generalization	Card. Role Name	 Genetic_map	 Ordered_map

Connector		Source	Target
Generalization	Card. Role Name	Ordered_map	Map

Organization

Type: **Class Provider**

Package: SNP

Document:

Affiliation of a group of unidentified persons.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Organization	Provider
Association	Card. Role Name	0..1 Page	* Organization

Person

Type: **Class Provider**

Package: SNP

Document:

An individual identified by name who is involved with creating or disseminating publications.

Possible roles are author,
editor, contributor, publisher, or applicant.

Connections

Connector		Source	Target
Generalization	Card. Role Name	 Person	 Provider
Association	Card. Role Name	* editor Person	0..1 Book
Association	Card. Role Name	0..1 Page	* Person
Association Curator of evidence.	Card. Role Name	0..1 Evidence	0..1 curator *Curator of evidence* Person

Attributes

Attribute	Lower bound	Notes
affiliation	*0*	Name of the organization the person is affiliated with.
email	*0*	Email address
et_al	*0*	Indicator that there are more persons in the list of persons; used typically for the bibliographic reference that contains only an abbreviated author list. Last listed author should have et_al set to true.
first_name	*0*	First name
mid_initials	*0*	Initials of any subsequent personal names.
postal_address	*0*	Full postal address

Attribute	Lower bound	Notes
surname	*0*	Last name.

Physical_map

Type: **Class Map**

Package: SNP

Document:

Map of physical entities, e.g. clones, contigs.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Physical_map	Map
Generalization	Card. Role Name	Haplotype_map	Physical_map
Generalization	Card. Role Name	Sequence_map	Physical_map

Reference_genomic_assembly

Type: **Class Sequence_map**

Package: SNP

Document:

Set of reference sequences (Reference_genomic_landmarks) with an algorithm to define the genome and its constituents.

Connections

Connector		Source	Target
Association Set of genomic landmarks which are part of the reference genomic assembly.	Card. Role Name	* Reference_genomic_landmark	0..1 Reference_genomic_assembly
Association Reference assembly on which locations are defined (coordinate system)	Card. Role Name	0..1 Reference_genomic_assembly	* Reference_genomic_location
Association	Card. Role Name	* Reference_genomic_assembly	0..1 Taxon
Generalization	Card. Role Name	 Reference_genomic_assembly	 Sequence_map
Association	Card. Role Name	* Reference_genomic_assembly	* present_in Genomic_reference_allele
Association	Card. Role Name	0..1 Page	* Reference_genomic_assembly

Reference_genomic_landmark

Type: **Class Sequence**

Package: SNP

Document:

An accessioned genomic sequence which can be unambiguously assigned to a location in an assembly based on information
in public databases.

Connections

Connector		Source	Target
Association Set of genomic landmarks which are part of the reference genomic assembly.	Card. Role Name	* Reference_genomic_landmark	0..1 Reference_genomic_assembly
Association Landmark of the location. Identified region on assembly.	Card. Role Name	0..1 Reference_genomic_landmark	* Reference_genomic_location
Association Genomic landmark (part of sequence assembly)	Card. Role Name	0..1 Genomic_variation	0..1 Reference_genomic_landmark
Generalization	Card. Role Name	 Reference_genomic_landmark	 Sequence
Association	Card. Role Name	0..1 Page	* Reference_genomic_landmark

Residue_change

Type: **Class Gene_product_change, Locatable**

Package: SNP

Document:

Abstract superclass for classes describing residue changes in gene-related sequences.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Translation_change	Residue_change
Generalization	Card. Role Name	Genic_variation	Residue_change
Generalization	Card. Role Name	Residue_change	Gene_product_change
Generalization	Card. Role Name	Residue_change	Locatable
Generalization	Card. Role Name	Transcription_change	Residue_change

Attributes

Attribute	Lower bound	Notes
changed_residue	*0*	The new residue in the reference sequence.
original_residue	*0*	The affected residue in the reference sequence.

Sequence

Type: **Class Locatable, Sequence_map**

Package: SNP

Document:

Biological sequence. A sequence accessioned by a public database and associated with an accessioned taxon, a version number and a release date.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Oligo	Sequence
Generalization	Card. Role Name	CDS	Sequence
Generalization	Card. Role Name	Genomic_allele	Sequence
Generalization	Card. Role Name	Reference_genomic_landmark	Sequence
Generalization	Card. Role Name	Sequence	Locatable
Generalization	Card. Role Name	Exon	Sequence
Generalization	Card. Role Name	Sequence	Sequence_map

Attributes

Attribute	Lower bound	Notes
molecule	*0*	Indicates the alphabet of the sequence molecule e.g. 'DNA', 'RNA', 'protein'
sequence	*0*	The residue sequence string.

Sequence_map

Type: **Class Physical_map**

Package: SNP

Document:

A map composed of one or more biological sequences.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Reference_genomic_assembly	Sequence_map
Generalization	Card. Role Name	Sequence	Sequence_map
Generalization	Card. Role Name	Sequence_map	Physical_map

Subject_descriptor

Type: **Class Identifiable**

Package: SNP

Document:

Subject term and vocabulary name container.

Connections

Connector		Source	Target
Generalization	Card. Role Name	 Subject_descriptor	 Identifiable
Association	Card. Role Name	* subject_heading Subject_descriptor	0..1 Bibref_subject
Association	Card. Role Name	* code Subject_descriptor	0..1 Bibref_subject
Association	Card. Role Name	* keyword Subject_descriptor	0..1 Bibref_subject
Association	Card. Role Name	0..1 Page	* Subject_descriptor

Attributes

Attribute	Lower bound	Notes
term	*0*	The keyword, code or subject heading value string
vocabulary_name	*0*	Name of the controlled vocabulary. E.g. Sears List of Subject Headings (SEARS), Library of Congress Subject Headings (LCSH), or Medical Subject Headings (MeSH).

Taxon

Type: **Class Identifiable**

Package: SNP

Document:

Taxonomic unit. The UID is typically a species TaxID from the Taxonomy Database (http://www.ncbi.nih.gov/
Taxonomy/)

Connections

Connector		Source	Target
Association Taxonomic name (e.g. species) of individual or population.	Card. Role Name	* Abstract_population	* Taxon
Association	Card. Role Name	* Reference_genomic_assembly	0..1 Taxon
Association	Card. Role Name	0..1 Page	* Taxon
Generalization	Card. Role Name	 Taxon	 Identifiable

Attributes

Attribute	Lower bound	Notes
rank	*0*	Rank (taxonomic category) of the taxon in the hierarchy. Typically 'species' or 'family'.
scientific_name	*0*	Scientific name, e.g. Homo sapiens.

BREF

Package: PAGE-OM

Document:
Module covering bibliographic references as defined in OMG specification SNP (Single Nucleotide Polymorphism).

Article

Type: **Class Bibliographic_reference**

Package: BREF

Document:
Non fictional prose forming an independent part of a publication.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Journal_article	Article
Generalization	Card. Role Name	Article	Bibliographic_reference
Generalization	Card. Role Name	Book_article	Article

Attributes

Attribute	Lower bound	Notes
first_page	*0*	First page in a article

Attribute	Lower bound	Notes
last_page	*0*	Last page in a article

Bibliographic_reference

Type: **Class Publication**

Package: BREF

Document:

The BibliographicReference class is the core of the data model. It is a super-class for all specialized citation types, but it can also be instantiated and represent an additional type not specifically defined in this specification.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Article	Bibliographic_reference
Generalization	Card. Role Name	Book	Bibliographic_reference
Generalization	Card. Role Name	Patent	Bibliographic_reference
Generalization	Card. Role Name	Thesis	Bibliographic_reference
Generalization	Card. Role Name	Web_resource	Bibliographic_reference
Association	Card. Role	0..1	0..1

Connector		Source	Target
	Name	desciption Bibref_description	Bibliographic_reference
Generalization	Card. Role Name	 Tech_report	 Bibliographic_reference
Association	Card. Role Name	0..1 status Entry_status	0..1 Bibliographic_reference
Generalization	Card. Role Name	 Proceeding	 Bibliographic_reference
Generalization	Card. Role Name	 Bibliographic_reference	 Publication
Association	Card. Role Name	0..1 identifier Db_xref	0..1 Bibliographic_reference
Association	Card. Role Name	0..1 publisher Provider	0..1 Bibliographic_reference
Association	Card. Role Name	* Provider	0..1 Bibliographic_reference
Association	Card. Role Name	* author Provider	0..1 Bibliographic_reference
Association	Card. Role Name	0..1 subject Bibref_subject	0..1 Bibliographic_reference
Association	Card. Role Name	0..1 coverage Bibref_scope	0..1 Bibliographic_reference
Association	Card. Role	*	0..1

Connector		Source	Target
	Name	contributor Provider	Bibliographic_reference
Association	Card. Role Name	0..1 Page	* Bibliographic_reference

Attributes

Attribute	Lower bound	Notes
date	*0*	Defines a date associated with an event in the life cycle of the cited resource, when this resource became available. Usually, it is a date of publishing, however, for not yet published resources, it can be a date of creation.
format	*0*	Describes the physical or digital manifestation of the cited resource. It can have very different content depending on the citation type.
language	*0*	Defines a language of the intellectual contents of the cited resource. The recommendation is to use values as defined by RFC1766 which includes a two-letter Language Code (taken from the ISO639 standard, followed optionally by a two-letter Country Code (taken from the ISO3166 standard). For example, "en" for English, "fr" for French, or "en-uk" for English used in the United Kingdom. Another possibility is to use MARC List of Languages.
rights	*0*	Specifies information about rights over the cited resource. Typically, it contains a rights management statement for the resource, or it refers to a service providing such information. Rights information often encompasses Intellectual Property Rights [IPR], Copyrights, and various Property Rights.
title	*0*	a name by which the resource is)A title given to the cited resource (formally known
type	*0*	It defines the nature or genre of the cited resource. Although a working draft of Dublin Core Types recommends a type classification, the proposed types are mostly out of scope of this specification. The majority of cited resources would fall in the same category "text". For the future, however, Dublin Core is considering the addition of subtypes to the high level types, or other ways of making sub-categories

Bibref_description

Type: **Class**

Package: BREF

Document:

A brief account of the content of the cited resource. It can be either the abstract, or the table of contents, or both. It can be written in a language different from the language of the cited resource.

Connections

Connector		Source	Target
Association	Card. Role Name	0..1 desciption Bibref_description	0..1 Bibliographic_reference
Association	Card. Role Name	0..1 Page	* Bibref_description

Attributes

Attribute	Lower bound	Notes
abstract_type	*0*	Format of the abstract expressed in MIME (Multipurpose Internet Mail Extensions)
language	*0*	Language of the description
table_of_contents	*0*	A list of divisions (chapters or articles) and the pages on which they start
the_abstract	*0*	Summary of the main points of the publication.
toc_type	*0*	Format of the table of contents expressed in MIME (Multipurpose Internet Mail Extensions).

Bibref_scope

Type: **Class Identifiable**

Package: BREF

Document:

It defines the extent or scope of the content of the cited resource. It can include spatial location (a place name or
geographic co-ordinates), temporal period (a period label, date, or date range), or both. Finally, it can have additional
dynamic properties such as jurisdiction).

Connections

Connector		Source	Target
Generalization	Card. Role Name	Bibref_scope	Identifiable
Association	Card. Role Name	0..1 coverage Bibref_scope	0..1 Bibliographic_reference

Attributes

Attribute	Lower bound	Notes
spatial_location	*0*	A place or area name or geographic coordinates in string format.
temporal_period	*0*	A period label, date, or date range.

Bibref_subject

Type: **Class**

Package: BREF

Document:

It defines the topic of the content of the cited resource.

Connections

Connector		Source	Target
Association	Card. Role Name	* subject_heading Subject_descriptor	0..1 Bibref_subject
Association	Card. Role Name	* code Subject_descriptor	0..1 Bibref_subject
Association	Card. Role Name	* keyword Subject_descriptor	0..1 Bibref_subject
Association	Card. Role Name	0..1 subject Bibref_subject	0..1 Bibliographic_reference

Book

Type: **Class Bibliographic_reference**

Package: BREF

Document:

A written work or composition that has been published; usually printed on pages bound together.

Connections

Connector		Source	Target

Connector		Source	Target
Generalization	Card. Role Name	 Book	 Bibliographic_reference
Association	Card. Role Name	1 from_book Book	0..1 Book_article
Association	Card. Role Name	* editor Person	0..1 Book

Attributes

Attribute	Lower bound	Notes
edition	*0*	Name of the edition.
isbn	*0*	International Standard Book Number
series	*0*	Name of a collection of books.
volume	*0*	Volume number of the book in a series.

Book_article

Type: **Class Article**

Package: BREF

Document:

Article that forms part of a book.

Connections

Connector		Source	Target

Connector		Source	Target
Generalization	Card. Role Name	 Book_article	 Article
Association	Card. Role Name	1 from_book Book	0..1 Book_article

Distributor

Type: **Class Source**

Package: BREF

Document:

The original source of the data.

Connections

Connector		Source	Target
Generalization	Card. Role Name	 Distributor	 Source

Entry_status

Type: **Class Publication**

Package: BREF

Document:

Defines information related to the citation itself rather than to the cited resource.

Connections

Connector		Source	Target
Association	Card. Role Name	0..1 status Entry_status	0..1 Bibliographic_reference
Generalization	Card. Role Name	 Entry_status	 Publication

Attributes

Attribute	Lower bound	Notes
last_modified_date	*0*	last date the citation entry was changed
subset	*0*	Some bibliographic repositories consist of several, or even many, databases. The subset helps to locate the citation.

Journal

Type: **Class Identifiable**

Package: BREF

Document:

A class describing a journal; a periodical dedicated to a particular subject. The citations referring to the journal articles
have a reference to this class. There are only few explicit attributes defined, the rest are accessible using dynamic
properties.

Connections

Connector		Source	Target
Generalization	Card. Role Name	 Journal	 Identifiable
Association	Card. Role Name	1 from_journal Journal	0..1 Journal_article
Association	Card. Role Name	0..1 Page	* Journal

Attributes

Attribute	Lower bound	Notes
abbreviation	*0*	Accepted abbreviated form of the journal name.
issn	*0*	The ISSN (International Standard Serial Number) is an eight-digit number which identifies periodical publications as such.

Journal_article

Type: **Class Article**

Package: BREF

Document:

Article, especially a scientific one that forms part of a journal.

Connections

Connector		Source	Target
Generalization	Card. Role Name	 Journal_article	 Article

Connector		Source	Target
Association	Card. Role Name	1 from_journal Journal	0..1 Journal_article

Attributes

Attribute	Lower bound	Notes
issue	*0*	Integer indicating the ordering of the issue within a volume.
issue_supplement	*0*	Additional identifier for a separate supplement of the issue.
volume	*0*	Volume number of the journal. Typically all issues of a journal published within a year belong to the same volume

Patent

Type: **Class Bibliographic_reference**

Package: BREF

Document:

A document granting an inventor sole rights to an invention.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Patent	Bibliographic_reference
Association	Card. Role Name	0..1 applicant Provider	0..1 Patent

Attributes

Attribute	Lower bound	Notes
doc_number	*0*	Patent identifier
doc_office	*0*	Name of the patent office. E.g. European Patent Office
doc_type	*0*	Type of a patent: e.g. 'utility', 'design' and 'plant'

Proceeding

Type: **Class Bibliographic_reference**

Package: BREF

Document:

Printed communication from a scientific organization.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Proceeding	Bibliographic_reference

Provider

Type: **Class Identifiable**

Package: BREF

Document:

Bibliographic resources. The most obvious examples are authors, but it includes also publishers and other contributors.

These attributes define the active participants. They may be persons, organizations, or even services. A publisher is
responsible for making the resource available. The authors and contributors are in ordered lists. The authors and
contributors are responsible for creating the contents of the cited resource. There is no formal definition of how this
60 Single Nucleotide Polymorphisms Specification
responsibility is divided between them. However, the authors are usually primary creators while contributors may be
illustrators, translators, or other creative providers. Their role may be specified in a separate attribute in dynamic
properties

Connections

Connector		Source	Target
Generalization	Card. Role Name	Organization	Provider
Generalization	Card. Role Name	Person	Provider
Association	Card. Role Name	0..1 publisher Provider	0..1 Bibliographic_reference
Association	Card. Role Name	* Provider	0..1 Bibliographic_reference
Association	Card. Role Name	* author Provider	0..1 Bibliographic_reference
Association	Card. Role Name	0..1 applicant Provider	0..1 Patent
Association	Card. Role	*	0..1

Connector		Source	Target
	Name	contributor Provider	Bibliographic_reference
Generalization	Card. Role Name	Service	Provider
Generalization	Card. Role Name	Provider	Identifiable

Publication

Type: **Class**

Package: BREF

Document:

A Citation. REF BibRef OMG standard. Not identifiable (Defined in model)

Connections

Connector		Source	Target
Association	Card. Role Name	* Identifiable	* Publication
Generalization	Card. Role Name	Bibliographic_reference	Publication
Generalization	Card. Role Name	Entry_status	Publication

Service

Type: **Class Provider**

Package: BREF

Document:

Provider of software service

Connections

Connector		Source	Target
Generalization	Card. Role Name	 Service	 Provider
Association	Card. Role Name	0..1 Page	* Service

Source

Type: **Class**

Package: BREF

Document:

Source of the data. Individual or institute.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Distributor	Source
Association Source of identifiable.	Card. Role Name	0..1 Identifiable	0..1 Source
Generalization	Card. Role Name	Contributor	Source
Association	Card. Role Name	0..1 Page	* Source

Attributes

Attribute	Lower bound	Notes
address	*0*	Street address.
email	*0*	Contact e-mail.
fax	*0*	Fax number.
institution	*0*	Name of the institution
name	*0*	Name of the Source.
tel	*0*	Telephone number.

Tech_report

Type: **Class Bibliographic_reference**

Package: BREF

Document:

Technical report

Connections

Connector		Source	Target
Generalization	Card. Role Name	Tech_report	Bibliographic_reference

Thesis

Type: **Class Bibliographic_reference**

Package: BREF

Document:

A treatise advancing a new point of view resulting from original research; usually a requirement for an advanced academic degree.

Connections

Connector		Source	Target
Generalization	Card. Role Name	Thesis	Bibliographic_reference

Web_resource

Type: **Class Bibliographic_reference**

Package: BREF

Document:

A referred web address

Connections

Connector		Source	Target
Generalization	Card. Role Name	Web_resource	Bibliographic_reference

Attributes

Attribute	Lower bound	Notes
cost	*0*	Cost of accessing the information
estimated_size	*0*	Rough size of the retrieved reference entry.
url	*0*	Uniform Resource Locator

8 Platform Specific Model

The previous chapter defines a platform independent model. The real implementations, however, are expected to depend on a more specific platform. This chapter shows an XML-based platform specific model expressing exchange format for data that conforms to platform independent model described above.

The platform specific model for XML derives its architecture according to the XML as suggested by W3C document "Extensible Markup Language (XML) 1.0 (http://www.w3.org/TR/REC-xml/)".

The model is defined by XML Schema as suggested by following W3C documents.

- XML Schema Part 0: Primer (http://www.w3.org/TR/xmlschema-0/)
- XML Schema Part 1: Structure (http://www.w3.org/TR/xmlschema-1/)
- XML Schema Part 2: Datatypes (http://www.w3.org/TR/xmlschema-2/)

The normative XML schema was generated from the modelling tool Enterprise Architect (http://www.sparxsystems.com/~~EA~~), version 7.5.847~~7.0.817.1~~. The target namespaces were propagated by the EA tool into the resulting XML Schema. However, because these target namespaces are not part of the XMI, when using a different tool, the namespaces must be changed in the resulting XML Schema manually as follows:

UML Package name	**Target namespace**
SNP	http://www.omg.org/spec/PAGE-OM/20090722/snp~~http://www.openpml.org/page-om/snp~~
~~SN P2~~	~~http://www.openpml.org/page-om/snp2~~
PAGE	http://www.omg.org/spec/PAGE-OM/20090722/page~~http://www.openpml.org/page-~~
BREF	http://www.omg.org/spec/PAGE-OM/20090722/bref~~http://www.openpml.org/page-~~
FUGE	http://www.omg.org/spec/PAGE-OM/20090722/fuge~~http://www.openpml.org/page-~~

Also following manual modifications were done into the XML schema files:
SchemaLocations attributes were added into the import elements for all packages (snp.xsd~~, snp2.xsd~~, page-om.xsd, bref.xsd and fuge.xsd) respectively.

Duplicate attributes coming from diamond inheritance are removed manually from snp.xsd file by editing out first occurrences of attributes from lsid to Source in Sequence and Residue elements.

The normative XML schema also uses consistently constructs `xs : choice minOccurs=" 0" maxOccurs="unbounded` instead of constructs `xs: sequence`.

The normative XML Schema is attached in the "~~Accompanied~~Accompanying files" as file *PSM/PAGE-OM.xsd*.

Annex C
~~Accompanied~~Accompanying Files

(Informative)

This annex is a set of the ~~accompanied~~accompanying files. Some of these files are normative and some of them contain examples and convenient images. If there is a discrepancy between the contents of the normative files and this document, then the normative files take precedence.

The ~~accompanied~~accompanying files are divided into two sections, PIM and PSM. The PIM section includes files related directly to the "Platform Independent Model," the PSM section contains the files used in the "Platform Specific Model."

Platform Independent Model

PAGE-OM_uml_2. 1_xmi_2. 1 .xml

A normative representation of the PAGE-OM specification.

PAGE-OM.EAP

This is the main file created by the Enterprise Architect (http://www.sparxsystems.com/) modeling tool. It served as a base for generating the XMI and description files, mentioned above.

01_Association_study.tiff

02_Sample.tiff

03_Genotype_overview.tiff

04_Phenotype_overview.tiff

05_Genotype_in_details.tiff

06_Frequency.tiff

07_Haplotype.tiff

08_Sequence.tiff

09_Map.tiff

90_Identifiable.tiff

91_Evidence_and_Value.tiff~~BREF_Bib.png~~

~~BasicTypes.png~~

~~FUGE_Ontology.png~~~~PAGE_Evidence_and_value.png~~

~~PAGE_Frequency.png~~~~PAGE_Genotype.png~~~~PAGE_Genotype2.png~~~~PAGE_Identifiable.png~~~~PAGE_Phenotype.png~~
~~PAGE_Phenotype2.png~~~~Page_study.png~~

~~Root_bref.png~~

~~Root_snp.png~~

~~Root_snp2.png~~

Annex C

~~SNP2_Haplotype.pngSNP2_Map.png~~

~~SNP2_Sample.pngSNP2_Sequence.pngSNP2_Sequence2.pngSNP_Map.png~~

~~SNP_Sample.png~~

~~SNP_Sequence.pngSNP_Sequence2.png~~

Diagrams of the PAGE-OM specification. These are the same diagrams as in this document but perhaps with a better resolution.

Platform Specific Model

~~all.xsd~~

basictypes.xsd
bref.xsd
fuge.xsd
page-om.xsd
snp.xsd
~~snp2.xsd~~

Annex C
References

(Informative)

- Brookes et. al., The Phenotype and Genotype Experiment Object Model (PaGE-OM): A robust data structure for information related to DNA variation, Human Mutation (2009), 6, p. 968-977

- Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. (2001) dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 29(1):308-3 11; see also http://www.ncbi.nlm.nih.gov/SNP/index.html

- Thorisson GA, Stein LD. (2003) The SNP Consortium website: past, present and future. Nucleic Acids Res. 31(1): 124- 127; see also http://snp.cshl.org/

- Fredman D, Siegfried M, Yuan YP, Bork P, Lehvaslaiho H, Brookes AJ. (2002) HGVbase: a human sequence variation database emphasizing data quality and a broad spectrum of data sources. Nucleic Acids Res. 30(1):387-391; see also http://www.hgvbase.org/

- Mika Hirakawa, Toshihiro Tanaka, Yoichi Hashimoto, Masako Kuroda, Toshihisa Takagi and Yusuke Nakamura (2002) JSNP: a database of common gene variations in the Japanese population. Nucleic Acids Res. 30(1):158-162; see also http://snp.ims.u-tokyo.ac.jp/

- Baker P G, Goble C A, Bechhofer S, Paton N W, Stevens R and Brass A (1999) An Ontology for Bioinformatics Applications. Bioinformatics 15, 510-520

- The Gene Ontology Consortium. Gene Ontology: tool for the unification of biology. Nature Genet. (2000) 25: 25-29.; see also http://www.geneontology.org/

- T.E. Klein, J.T. Chang, M.K. Cho, K.L. Easton, R. Fergerson, M. Hewett, Z. Lin, Y. Liu, S. Liu, D.E. Oliver, D.L. Rubin, F. Shafa, J.M. Stuart and R.B. Altman, "Integrating Genotype and Phenotype Information: An Overview of the PharmGKB Project", The Pharmacogenomics Journal (2001) 1, 167-170.; see also http://www.pharmgkb.org/

- Rajeevan H, Osier MV, Cheung KH, Deng H, Druskin L, Heinzen R, Kidd JR, Stein S, Pakstis AJ, Tosches NP, Yeh CC, Miller PL, Kidd KK. "ALFRED the ALlele FREquency Database update." Nucleic Acids Research..31(1):270- 271 .(2003); see also http://alfred.med.yale.edu/

- Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R, Hao L, Kiang A, Paschall J, Phan L, Popova N, Pretel S, Ziyabari L, Lee M, Shao Y, Wang ZY, Sirotkin K, Ward M, Kholodov M, Zbicz K, Beck J, Kimelman M, Shevelev S, Preuss D, Yaschenko E, Graeff A, Ostell J, Sherry ST. The NCBI dbGaP database of genotypes and phenotypes. Nat Genet. 2007 Oct;39(1 0): 1181-1186.

- Yamada Y, Kato K, Hibino T, Yokoi K, Matsuo H, Segawa T, Watanabe S, Ichihara S, Yoshida H, Satoh K, Nozawa Y. Prediction of genetic risk for metabolic syndrome. Atherosclerosis 2007; 191:298-304.

Annex C
Glossary

(Informative)

Term	Definition
Abstract_observable_feature [in PAGE-OM]	An extension point for other kinds of observable features.
Abstract_observation_target [in PAGE-OM]	It is an abstract class for all entities from which one can make genotype or phenotype measurements or observations. It deals with entities capable of being observed.
Abstract_population [in PAGE-OM]	An interbreeding set of individuals, from whom a Panel is drawn. (Population in SNP-PML) . Extends Abstract_observation_target, which is abstract class for all entities from which one can make genotype or phenotype measurements or observations.
Abstract_run [in PAGE-OM]	An extension point for other kinds of runs.
Abstract_value [in PAGE-OM]	An extension point for kinds of values
Algorithm [in PAGE-OM]	Step-by-step procedure for solving a problem
Anatomic_location [in SNP]	Description of the anatomical location the molecular sample is derived from. Best expressed as a controlled vocabulary of anatomical terms.
Annotation [in SNP]	A free text of name and value pair. Note: name should be Ontology_term
Article [in BREF]	Non fictional prose forming an independent part of a publication.
Assayed_genomic_genotype [in PAGE-OM]	The result of applying a variation assay to an individual, to reveal one or more of the genomic alleles carried by that individual. This term applies to the observed data rather than to the inferred state of the individual. Thus the same individual might have several different

	genotypes at the same site, where the variation might be due to differing assays, experimental error, dominant systems, missing data, and so forth. Synonym: Measurement.
Association_study [in PAGE-OM]	Association study is core concept of the specification. It captures relationships between phenotypes and genotypes. It is an examination of genetic variation across the genome, designed to identify genetic associations with observable phenotypes. Association studies are results of correlation experiments,
Band_size [in PAGE-OM]	DNA fragment length estimated from gel electrophoresis
Bibliographic_reference [in BREF]	The BibliographicReference class is the core of the data model. It is a super-class for all specialized citation types, but it can also be instantiated and represent an additional type not specifically defined in this specification.
Bibref_description [in BREF]	A brief account of the content of the cited resource. It can be either the abstract, or the table of contents, or both. It can be written in a language different from the language of the cited resource.
Bibref_scope [in BREF]	It defines the extent or scope of the content of the cited resource. It can include spatial location (a place name or geographic co-ordinates), temporal period (a period label, date, or date range), or both. Finally, it can have additional dynamic properties such as jurisdiction).
Bibref_subject [in BREF]	It defines the topic of the content of the cited resource.
Book [in BREF]	A written work or composition that has been published; usually printed on pages bound together.
Book_article [in BREF]	Article that forms part of a book.
Boolean [in PAGE-OM]	Value of type boolean
CDS [in SNP]	Region in a reference sequence that determines the start, end and strand of the gene coding region.

Annex C

Conclusion [in PAGE-OM]	A reasoned judgment of an experiment
Consensus_genomic_genotype [in PAGE-OM]	This class represents consensus from several experiments providing genotypes of the same sample on the same site.
Constraint [in PAGE-OM]	The contents of a Value can be limited by Constraints. Different types of Constraints allow various ways how to limit or validate one or more Value instances. The Constraint superclass only stores a string description of the Constraint. The actual full semantics of a constraint are specified in various subclasses described below. But there are no subclasses in the PAGE-OM - because they are out of scope of PAGE-OM.
Contributor [in SNP]	Person who has edited data in the data set. Synonym: curator.
Cytogenetic_location [in SNP]	A location in a cytogenetic map expressed as a band or a range of bands.
Cytogenetic_map [in SNP]	Map of cytogenetic bands describing their relative order.
Db_xref [in SNP]	Cross reference between two Identifiables in different databases.
Defining_feature [in PAGE-OM]	Is an association class that has list of values, which are used in defining the instance of Latent_genotype_specifications (for example intensity values used in allele calling). The class captures information how alleles are called (observed) from raw measurement values like intensity values
Distributor [in BREF]	The original source of the data.
Entry_status [in BREF]	Defines information related to the citation itself rather than to the cited resource.
Enum [in PAGE-OM]	Enumeration contains list of Values
Environment_feature [in PAGE-OM]	Circumstances, objects, or conditions by which one is surrounded

Error [in PAGE-OM]	Error value is numeric value of accuracy. Quality score
Evidence [in PAGE-OM]	Evidence indicates reliability of a feature or simply documents its authoritative origin.
Evidence_code [in PAGE-OM]	Evidence can be an EvidenceCode (which is a controlled vocabulary term such as a GO evidence code or ICIS Method code) but can be a more fully documented Evidence object (inheriting from EvidenceCode) generally curated by a specified person, a curator modeled as a Contact. Its strength is expressed by the score (which is usually a numeric value between 0 and 1, but also other types of Value are allowed - e.g. an ontology term value). The core of an evidence is its supporting source which can be anything (because it is identified by a SimpleIdentifier). Usual evidence sources are BiblioReferences, Studies and OntologyTerms. Reference (generationcp - http://pantheon.generationcp.org/demeter/Features.html)
Exon [in SNP]	Region in a reference sequence that determines the start, end and strand of the transcript that is not spliced away.
Experiment_result [in PAGE-OM]	The experiment result (for example a single p-value) gathers correlation between genomic observation and phenotypic observed values. A correlation experiment can consist of more than one experiment results.
Float_value [in PAGE-OM]	Value of type float
Frequency [in PAGE-OM]	Abstract class for frequencies, expressed in percentages. Alleles (Genomic_alleles), genotypes (Consensus_genomic_genotype) and haplotypes (Genomic_haplotype) can have measured frequencies in population samples (Panels). In addition, heterozygosity (Heterozygosity) is a measure of observed variability of a polymorphic site (Genomic_variation)in a sub-population (Panel).
Frequency_set [in PAGE-OM]	Set of frequencies
Functional_change [in PAGE-OM]	Change in the function of the final gene product.
Gene_based_haplotype [in PAGE-OM]	Collection of variable nucleotides (Genomic_alleles in

Annex C

	Genomic_variations) that define a gene. In older usage synonym locus.
Gene_product_change [in SNP]	Free form description of change in gene product
Genetic_location [in SNP]	A relative location of a marker in a genetic map based on recombination.
Genetic_map [in SNP]	A map containing the order and non-metric distance between genetic markers for identifiable sequence features. Unit: centiMorgan.
Genic_variation [in PAGE-OM]	Genomic variation with location in genic coordinates. Synonym: mutation (when change from a common allele affects phenotype)
Genomic_allele [in PAGE-OM]	One of several alternative DNA sequences of a Reference_genomic_location as it appears in the population of organisms. Synonym: variant, allele
Genomic_allele_population_frequency [in PAGE-OM]	The frequency with which a particular Genomic_allele is seen in a particular Panel. This frequency can be measured from pooled samples. Synonyms: Genomic_allele_panel_frequency, allele_frequency.
Genomic_gene_structure [in PAGE-OM]	A structure of a gene expressed as location of the CDS and exons. Defines genic coordinate system from start of the CDS downstream.
Genomic_genotype_population_frequency [in PAGE-OM]	Frequency of a Consensus_genomic_genotype in a Panel. OSAGE-OM Has many to one relationship to Latent_genotype (Consensus_genomic_genotype in SNP-PML)
Genomic_haplotype [in PAGE-OM]	A set of Genomic_alleles across an equal number of Genomic_variations in a single chromosome and in a single individual. The Genomic_haplotype is derived from a set of Consensus_genomic_genotype. For each Genomic_variation, the haplotype contains one and only one Genomic_allele. Furthermore, the Genomic_alleles are required to be in phase on the individual, meaning that they are located on the same contiguous strand of DNA. Synonym: Haplotype.

Genomic_haplotype_population_frequency [in PAGE-OM]	Frequency of a Genomic_haplotype in a Panel.
Genomic_observation [in PAGE-OM]	Genomic observation
Genomic_reference_allele [in SNP]	The DNA sequence of a Genomic_variation as it appears in the public database.
Genomic_variation [in PAGE-OM]	A variable site in a Reference_genomic_landmark sequence. Synonyms: polymorphic site, marker, (Genomic_polymorphism in SNP-PML).
Genotype_phenotype_correlation_experiment [in PAGE-OM]	Family or case control based association study experiment. Represents set of experiment sub-sections that would normally be listed in the results section in manurscripts.
Geographic_location [in SNP]	Location of an individual or population in a geographic map. Locations are expressed in decimal degrees. Northern latitudes (if north of the equator) and eastern longitudes (if east of GM) have positive values by convention.
Haplotype_block [in PAGE-OM]	Large (spanning a few kb to >100 kb) blocks of Genomic_alleles in linkage disequilibrium (LD) and a few haplotypes per block, separated by regions of recombination.
Haplotype_derivation_methods [in PAGE-OM]	Association class describing methods used to derive Genomic_haplotypes from Consensus_genomic_genotypes.
Haplotype_map [in PAGE-OM]	Map of haplotypes. Features include: Block length distribution, measures of block variability, relative proportions of common haplotypes, block coverage of chromosomes and/or genome. LD and other values between haplotypes, markers, alleles.
Heterozygosity [in PAGE-OM]	Heterozygosity (Heterozygosity) is a measure of observed variability of a polymorphic site (Genomic_variation) in a sub-population (Panel).
Hypothesis [in PAGE-OM]	Free text description of hypothesis of study.

Annex C

Identifiable [in SNP]	Anything with a unique identifier.
Individual [in PAGE-OM]	A single member of a species, where a species is an accessioned taxon defined by a public database, and the individual is accessioned in a public or private database. Synonym: "inbred strain" in homozygous lineages.
Integer_value [in PAGE-OM]	Value of type integer
Journal [in BREF]	A class describing a journal; a periodical dedicated to a particular subject. The citations referring to the journal articles have a reference to this class. There are only few explicit attributes defined, the rest are accessible using dynamic properties.
Journal_article [in BREF]	Article, especially a scientific one that forms part of a journal.
Latent_genotype [in PAGE-OM]	Potentially existing genotypes on specific site that could be observed by Variation_assays. Application of a Variation _assay on one Molecular_sample generates a single Latent_genotype which has one or more Latent_genotype _specifications (this depends on ploidy level in case of Genomic_allele). Latent_genotypes associated to one instance of a Variation_assay can have only one type of Latent_genotype_specifications, as defined by Defining_feature. This class is a holder for one or many observable variation objects (Latent_genotype_specification). Latent_genotype is used to attach possible variations to measurements (Assayed_genomic_genotype), variation assays (Variation_assay) and marker loci (Genomic_variation).
Latent_genotype_specification [in PAGE-OM]	Abstract super class of observable variation objects, like alleles, melting temperatures (Melting_temperature), band sizes (Band_size). The class is an extension point to other kinds of variations.
Lifestyle_feature [in PAGE-OM]	Way of life of an individual or panel
Locatable [in SNP]	Abstract class for anything that can be placed on a sequence, i.e. can act as a sequence feature.

Location [in SNP]	A position in an abstract map.
Location_on_plate [in PAGE-OM]	X,Y plate_positions (wells). Numbering starts from one. Each well can contain one or more observation targets (molecular samples) prepared for measurement using one or more variation assays (e.g. assay multiplexing Assay_set). Note: These are optional laboratory specific details (Sample and Assay information is in Assayed_genomic_genotype)
Map [in SNP]	An abstract map.
Melting_temperature [in PAGE-OM]	The temperature at which DNA goes from a double-stranded to a single-stranded state. Unit of temperature is Celsius.
Methodology [in SNP]	Abstract class for a laboratory method or a computational protocol.
Molecular_sample [in PAGE-OM]	A sample from an Individual or from a Panel defining the molecule and tissue/cell used (Anatomic_locations) in the Variation_assay. Synonym: Sample of individual.
Multi_vartiation_assay [in PAGE-OM]	Multi_variation_assay is a collection of assays which may be used simultaneously. Examples would be multiplex assays, micro-array based assays, or a panel of single-plex assays that share some common feature or purpose.
Neighbour_variation [in PAGE-OM]	Another Genomic_variation close enough to affect the primer design.
Numeric_value [in PAGE-OM]	Numeric value
Observable_feature [in PAGE-OM]	Measurable feature of observable (e.g. size of nose)
Observable_feature_category [in PAGE-OM]	All features considered by this model can be categorized by using this class. The category should be expressed by an ontology term.
Observation_method [in PAGE-OM]	Observable features can be measured by different methods. This class specifies which method has been used. For example, a method can be usage of a ruler or filling a questionnaire.

Annex C

Observed_value [in PAGE-OM]	Observation done at specific point in time.
Oligo [in SNP]	A DNA oligonucleotide used for detection and assay of Genomic_variations in Variation_assays. Synonym: primer
Ontology_source [in FUGE]	The source ontology or controlled vocabulary list that ontology terms have been obtained from
Ontology_term [in FUGE]	Ontology term, defined in more detail in FuGE (http://fuge.sourceforge.net/). This is just a place holder
Ordered_location [in SNP]	A location on a secondary map that gives ordering information. That map can be a genetic or a cytogenetic map.
Ordered_map [in SNP]	A non-metric map category.
Organization [in SNP]	Affiliation of a group of unidentified persons.
Page [in PAGE-OM]	This class does not contain any scientific meaning. Its main purpose is to be the root element for the situations where this specification is used for data exchange formats (e.g. xml-schema). Therefore, it has optional direct associations to all important classes so that implementations can exchange only relevant data.
Panel [in PAGE-OM]	A set of samples from individuals drawn from the same species and used for genetic studies. A panel must be identifiable with a list of accessioned individuals, if possible. Panel can have subpanels. Synonym: SampleSet, Sample from population(s), “Plate” in Coriel sense.
Patent [in BREF]	A document granting an inventor sole rights to an invention.
Person [in SNP]	An individual identified by name who is involved with creating or disseminating publications. Possible roles are author, editor, contributor, publisher, or applicant.
Phenotype_disease_category [in PAGE-OM]	Specialized category of features representing diseases.

Phenotype_feature [in PAGE-OM]	Observable part of the structure, function or behavior of a living organism.
Physical_map [in SNP]	Map of physical entities, e.g. clones, contigs.
Plate [in PAGE-OM]	A sample holder, for example a microtiter plate used in one or many runs, represented by instances of Run. Samples, represented by instances of Molecular_sample, are positioned on the plate using instances of Location_on_plate.
Proceeding [in BREF]	Printed communication from a scientific organization.
Provider [in BREF]	Bibliographic resources. The most obvious examples are authors, but it includes also publishers and other contributors. These attributes define the active participants. They may be persons, organizations, or even services. A publisher is responsible for making the resource available. The authors and contributors are in ordered lists. The authors and contributors are responsible for creating the contents of the cited resource. There is no formal definition of how this 60 Single Nucleotide Polymorphisms Specification responsibility is divided between them. However, the authors are usually primary creators while contributors may be illustrators, translators, or other creative providers. Their role may be specified in a separate attribute in dynamic properties
Publication [in BREF]	A Citation. REF BibRef OMG standard. Not identifiable (Defined in model)
Reference_genomic_assembly [in SNP]	Set of reference sequences (Reference_genomic_landmarks) with an algorithm to define the genome and its constituents.
Reference_genomic_landmark [in SNP]	An accessioned genomic sequence which can be unambiguously assigned to a location in an assembly based on information in public databases.
Reference_genomic_location [in PAGE-OM]	A location within a Reference_genomic_landmark. Attributes of the location are the Reference_genomic_assembly and/or the Reference_genomic_landmark, the start and end range

	and strand of the feature relative to the Reference_genomic_landmark.
Reference_genomic_location_in_assembly [in PAGE-OM]	A location in one chromosome of a reference genomic assembly. Instead of the reference sequence being an accessioned sequence, it is a versioned assembly.
Residue_change [in SNP]	Abstract superclass for classes describing residue changes in gene-related sequences.
Run [in PAGE-OM]	The class contains information on measurement of samples, done on a physical device connected to plate.. This information includes time of execution, name of instrument, etc.
Sequence [in SNP]	Biological sequence. A sequence accessioned by a public database and associated with an accessioned taxon, a version number and a release date.
Sequence_map [in SNP]	A map composed of one or more biological sequences.
Service [in BREF]	Provider of software service
Set_of_haplotypes [in PAGE-OM]	An extension point for collections of haplotypes.
Source [in BREF]	Source of the data. Individual or institute.
String_value [in PAGE-OM]	Value of type string
Structural_change [in PAGE-OM]	Change in the 3D structure of the polypeptide chain.
Study [in PAGE-OM]	An extension point for adding other kind of studies in the future.
Subject_descriptor [in SNP]	Subject term and vocabulary name container.
Taxon [in SNP]	Taxonomic unit. The UID is typically a species TaxID from the Taxonomy Database (http://www.ncbi.nih.gov/Taxonomy/)
Tech_report [in BREF]	Technical report
Thesis [in BREF]	A treatise advancing a new point of view resulting from original research; usually a requirement for an advanced

	academic degree.
Time_accuracy [in PAGE-OM]	Accuracy code contains information on incompleteness of time of measurement or information on reason why the time of measurement is unknown or incomplete.
Transcription_change [in PAGE-OM]	Change in the quality or quantity of the mature RNA product.
Translation_change [in PAGE-OM]	Change in the quality or quantity of (predicted) polypeptide chain (2D).
Unit [in PAGE-OM]	Unit of value. Unit is defined using ontology term
Value [in PAGE-OM]	Abstract class. Extension point for Value implementations. Value model is based on concept developed in Generation Challenge Program: http://pantheon.generationcp.org/demeter/Values.html
Value_range [in PAGE-OM]	Inclusive value range
Variation_assay [in PAGE-OM]	An experimental lab protocol and set of reagents for detecting the Genomic_alleles of Genomic_variations carried by an individual or a panel of individuals. Synonym: Assay. Non instrument part of the experiment - same assay can be used in different instrument runs
Web_resource [in BREF]	A referred web address

~~Tem Definition~~

~~Anatomic location Description of the anatomical location the molecular sample is derived~~

~~from. Best expressed as a controlled vocabulary of anatomical terms.~~

Annex C

Assay Set Set of assays which are group experimentally. (e.g., microarray)

Assayed genomic genotype The result of applying a Variation assay to an individual, to reveal one or more of the Genomic alleles carried by that individual. This term applies to the observed data rather than to the inferred state of the individual (see Consensus genomic genotype). Thus the same individual might have several different Assayed genomic genotypes at the same Genomic variation, where the variation might be due to differing assays, experimental error, dominant systems, missing data, and so forth. Synonym: Measurement.

Association Study Examination of genetic variation across the human genome, designed to identify genetic associations with observable phenotypes

CDS Region in a reference sequence that determines the start, end and strand of the gene coding region.

Consensus genomic genotype A list of one or more Genomic alleles carried by a single individual for a single Genomic variation. Unlike the Assayed genomic genotype, the genomic genotype is intended to be used solely for assertions as to the true value of the alleles, after experimental error and other factors are taken into account. The number of Genomic alleles is equal to the ploidy (copy-number) of the individual's genomic DNA at the Reference genomic location of the Genomic variation. Synonym: Inferred genomic genotype, Genomic genotype.

~~Contributor — Person who has edited data in the data set. Synonym: curator.~~

~~Cytogenetic location — A location in a cytogenetic map expressed as a band or a range of bands.~~

~~Cytogenetic map — Map of cytogenetic bands describing their relative order.~~

~~Db xref—Cross-reference between two Identifiable in different databases.~~

~~Disease—Special case of a phenotype describing a disease state.~~

~~Environment — Circumstances, objects, or conditions by which one is surrounded.~~

~~Exon — Region in a reference sequence that determines the start, end and strand~~

~~of the transcript that is not spliced away.~~

~~Frequency Set — Set of frequencies for one genomic variation in one panel using one~~

~~variation assay~~

~~Functional change Change in the function of the final gene product.~~

~~Gene based haplotype — Collection of nucleotides (Genomic alleles in Genomic variations) that~~

~~define a gene. In older usage same as locus.~~

~~Gene product change — Abstract superclass for a different genic change descriptors.~~

~~Genetic location — A relative location of a marker in a genetic map based on recombination.~~

~~Genetic map — A map containing the order and non-metric distance between genetic~~

markers for identifiable sequence features. Unit: centi-morgan.

Genic variation — Genomic variation with location in coordinates. Synonym: mutation

(when change from a common allele affects phenotype).

Genomic allele – One of several alternative DNA sequences of a Reference genomic

location as it appears in the population of organisms. Synonym: variant, allele.

Genomic allele population frequency — The frequency with which a particular Genomic allele is seen in a

particular Panel. This can be measured from pooled samples. Synonyms: Genomic allele panel frequency, allele frequency.

Genomic gene structure — A of a gene expressed as location of the CDS and exons. Defines genic

coordinate system from start of the CDS downstream.

Genomic genotype population frequency – Frequency of a Consensus genomic genotype in a Panel.

Genomic haplotype A set of Genomic alleles across an equal number of Genomic variations

in a single chromosome and in a single individual. The Genomic haplotype is derived from a set of genomic genotype. For each Genomic variation, the haplotype contains one and only one Genomic allele. Furthermore, the Genomic alleles are required to be in phase on the individual, meaning that they are located on the same contiguous strand of DNA. Synonym: Haplotype.

Genomic haplotype population frequency Frequency of a Genomic haplotype in a Panel.

Genomic reference allele The DNA sequence of a Genomic variation as it appears in the public

database.

Genomic Variation A variable site in a Reference genomic landmark sequence. Synonyms:

polymorphic site, marker

Genomic variation A site in a Reference genomic sequence. Synonyms: polymorphic site,

polymorphism, SNP.

Genotype-Phenotype Correlation Experiment Study to detect associations between genotype and phenotype. Can be

~~family, population, or case/control based. Synonyms: Association study, Linkage Study.~~

~~Haplotype block — Large (a few kb to >100 kb) blocks of Genomic alleles in linkage~~

~~disequilibrium (LD) and a few haplotypes per block, separated by region of recombination.~~

~~Haplotype derivation methods — Association class describing methods used to derive Genomic haplotypes~~

~~from Consensus genomic genotypes.~~

~~Heterozygosity — A numerical measure of observed abundance of different Genomic alleles~~

~~in a Genomic variation.~~

~~Haplotype map — Map of haplotypes. Features include: 1.~~

~~Individual — A single member of a species, where a species is an accessioned taxon~~

~~defined by a public database, and the individual is accessioned in a public or private database. Synonym: "inbred strain" in homozygous lineages.~~

~~Methodology — Abstract class for a laboratory method or a computational protocol.~~

~~Molecular sample A sample from Individual from a Panel defining the molecule and tissue/~~

~~cell used (Anatomic locations) in the Variation assay. Synonym: Sample of individual.~~

~~Neighbour variation Another Genomic variation close enough to affect the Variation assay.~~

~~Observable Feature Measurable feature of observable (e.g., size of nose)~~

~~Observable Feature Category Observable category ideally expressed using an appropriate ontology~~

~~Observation Method Method of observation (e.g., ruler). Can be also questionnaire.~~

~~Observation Target Capable of being observed. For example nose size or eye color.~~

~~Observation Value Observation done at specific point in time. Time is expressed according~~

~~to the W3C profile version of an ISO 8601 date-time string. http:// www.w3.org/TR/NOTE-datetime~~

~~Oligo A DNA oligo-nucleotide used for detection and assay of Genomic~~

~~variations in Variation assay. Synonym: primer.~~

~~Ordered location — A location on a secondary map that gives ordering information. That map~~

~~can be a genetic or a cytogenetic map.~~

~~Ordered map — A non-metric map category.~~

~~Panel — A set of samples from individuals drawn from the same species and used~~

~~for genetic studies. A panel must be identifiable with a list of accessioned individuals, if possible. (Note: do not use "population" here!) Synonym: Sample Set, Sample from population(s). Plate in Coriel sense.~~

~~Phenotype — Observable part of the structure, function or behavior of a living~~

~~organism.~~

~~Phenotype Disease Category — Category associated with disease phenotype. Best expressed using a~~

~~disease ontology.~~

~~Physical map Map of physical entities, e.g., clones, contigs.~~

~~Population A interbreeding set of individuals, from whom a Panel is drawn.~~

~~Reference genomic assembly Set of reference sequences (Reference genomic landmarks) with an algorithm to define the genome and its constituents.~~

~~Reference genomic landmark An accessioned sequence which can be unambiguously assigned a location in an assembly based on information in public databases.~~

~~Reference genomic locationA location within a Reference genomic. Attributes of the location are the Reference genomic assembly and/or the Reference genomic landmark, the start and end range strand of the feature relative to the Reference genomic landmark.~~

~~Reference genomic location in assembly A location in one chromosome of a reference genomic assembly. Instead of the reference sequence being an accessioned sequence, it is a versioned assembly.~~

~~Residue change Abstract superclass for classes describing residue changes in gene-related sequences.~~

~~Sequence Biological sequence. A sequence accessioned by a public database and~~

~~associated with an accessioned taxon, a version number and a release date.~~

~~Sequence map A map composed of one or more biological sequences.~~

~~Set of haplotypes An abstract representation of a collection of haplotypes.~~

~~Structural change Change in the 3D structure of the polypeptide chain.~~

~~Taxon Taxonomic unit. The UID is typically a TaxID from the Taxonomy~~

~~Database (http://www.ncbi.nih.gov/Taxonomy/).~~