DE GRUYTER Pure Appl. Chem. 2018; 90(7): 1199–1220

IUPAC Technical Report

Ulla M. Petersen*, Ariadna Padró-Miquel, Graham Taylor, Jens Michael Hertz, Rebecca Ceder, Xavier Fuentes-Arderiu and Johan T. den Dunnen

Properties and units in the clinical laboratory sciences part XXIV. Properties and units in clinical molecular genetics (IUPAC Technical Report)

https://doi.org/10.1515/pac-2017-1008
Received October 31, 2017; accepted March 7, 2018

Abstract: This document describes the application of the syntax, semantic rules, and format of the Nomenclature for Properties and Units (NPU) terminology for coded dedicated kinds-of-property in the subject field of clinical molecular genetics. A vocabulary for NPU definitions in this field, based on international terminology and nomenclature, is introduced and examples of actual NPU definitions for different types of investigations are given and explained.

Keywords: DNA; nomenclature; nucleic acids; nucleotides; terminology.

CONTENTS

Article note: Sponsoring bodies: IUPAC Chemistry and Human Health Division and International Federation of Clinical Chemistry and Laboratory Medicine (IFCC); see more details on page 1219.

***Corresponding author: Ulla M. Petersen,** Department of Data Quality and Content, The Danish Health Data Authority, Copenhagen, Denmark, e-mail: UMP@sundhedsdata.dk
Ariadna Padró-Miquel: Clinical Laboratory, IDIBELL – Hospital, Universitari de Bellvitge, L'Hospitalet de Llobregat, Barcelona, Spain
Graham Taylor: Clinical Genomics at ViaPath, Guy's, St Thomas's and King's College Hospitals, London, Great Britain
Jens Michael Hertz: Department of Clinical Genetics, Odense University Hospital, Odense, Denmark
Rebecca Ceder: Equalis, Uppsala, Sweden
Xavier Fuentes-Arderiu: Barcelona, Spain
Johan T. den Dunnen: Medical Center, Leiden University, Leiden, Netherlands

1 Preface

The present document is Part XXIV of a series on properties and units in the clinical laboratory sciences initiated in 1987.

The series currently comprises:

I Syntax and semantic rules [1]
II Kinds-of-property [2]
III Elements (of properties) and their code values [3]
IV Properties and their code values [4]
V Properties and units in thrombosis and haemostasis [5]
VI Properties and units in IOC-prohibited drugs [6]
VIII Properties and units in clinical microbiology [7]
IX Properties and units in trace elements [8]
X Properties and units in general clinical chemistry [9]
XI Coding systems: structure and guidelines [10]
XII Properties and units in clinical pharmacology and toxicology [11]
XIII Properties and units in reproduction and fertility [12]
XVI Properties and units in clinical allergology [13]
XVIII Nomenclature, properties and units in clinical molecular biology [14]
XIX Properties and units for transfusion medicine and immunohematology [15]
XX Properties and units in clinical and environmental human toxicology [16]
XXIII The NPU terminology, principles, and implementation – A user's guide [17]

2 Foreword and scope

The use of informatics in the health care field, and especially within clinical laboratory sciences, has enabled the communication and storage of clinical laboratory information in a structured manner. The demand for actionable data for clinical purposes, *e.g.* for enhancement of clinical workflows, data integration to improve detection and treatment of disease, as well as for research and statistics, is progressively increasing. Laboratory information may be used far from its source of origin in both time and space, and agreement on a common, stable, and unequivocal set of terms and definitions is necessary in order to prevent the loss or distortion of their meaning. For two decades, the NPU Terminology has supplied code values and definitions based on international standards, nomenclatures, and terminologies.

This document continues the work published in Part XVIII. Properties and units in clinical molecular biology (IUPAC Technical Report). Its primary purpose is to apply the IUPAC–IFCC recommended syntax structures for

requests and reports, providing formats and names of properties observed in the domain of clinical molecular genetics to facilitate unequivocal written or electronic communication among health care professionals.

The scope of this document is limited to human molecular genetics; other medical aspects of genetics, such as biochemical or physiological indicators of genetic variants or molecular genetics of clinically relevant microorganisms, are excluded.

This document is concerned with identifying the properties investigated in a standardized way, but it is equally important that the result values reported are filed and communicated in a standardized way. Any report of a variation of a nucleic acid sequence should include the identity of the sequence referred to and be expressed in a structured manner. For this purpose, the Human Genome Variation Society (HGVS) [18] recommendations for the description of sequence variants [19] are strongly recommended.

3 Concepts, terms and definitions

3.1 General concepts for properties in the clinical laboratory sciences

system: a demarcated part or phenomenon of the perceivable or conceivable universe, material or immaterial, that may be regarded as a set of entities, together with a set of relations or processes between these entities [20].

component: part of a system [20].

EXAMPLE

The CFTR gene is a component of the system (human) DNA

property: state- or process-descriptive feature of a system including any pertinent components [20].

kind-of-property: aspect common to mutually comparable properties [20].

dedicated kind-of-property: kind-of-property with given sort of system and any pertinent sorts of component [20].

EXAMPLE

DNA(specification)—A1BG gene; sequence variation

scale type (or **value set type**)**:** a kind-of-property may be qualified to narrative values, nominal value set, ordinal scale, differential scale, or rational scale [21]. Kinds-of-property qualified by the last two scale types are are also called kind-of-quantity.

narrative values: values consisting of an unlimited set of possible descriptions of the kind-of-property. A narrative description may specify in detail the actual property or properties investigated and may include information on technique and clinical relevance.

nominal value set: value set with a set of possible values for a given kind-of-property that are each a word or symbol without any relation to magnitude

NOTE: The values may be listed in any order according to practical considerations and convention.

category: nominal kind-of-property indicating a class among nominal properties according to values of a specified value set [22].

entitic: numerator kind-of-quantity, usually extensive, divided by number of entities [20].

NOTE: may be combined with kinds-of-property to refer to a single entity (particle) of a system consisting of many like entities.

EXAMPLES

Blood—Erythrocytes; volume = 2.3 litre
The total volume of all erythrocytes in the patient's blood is 2.3 litre

Blood—Erythrocytes; entitic volume = 87 femtolitre
The (mean) volume of one erythrocyte in the patient's blood is 87 femtolitre

arbitrary: modifier to kinds-of-property used to denote lack of a common reference for the result value set.

NOTE: May be combined with kinds of property to indicate that result values are not comparable unless performed by identical procedures. Arbitrary kinds-of-property are used with all definitions specifying ordinal scale types. They may also be used with ratio scale types, where a SI unit for technical reasons cannot be assigned. In these cases the unit is replaced with the text '(procedure defined unit)' or '(p.d.u)'.

EXAMPLES

Patient—Body; arbitrary mass = 11 (p.d.u)
Ratio scale type
Result values are given using a local unit of mass, in this example 'Stone'. Values are proportional to the mass expressed in kg.

Patient—Body; arbitrary mass = high
Ordinal scale type
The locally defined value set might be *e.g.* {low, normal, high}

3.2 Concepts in dedicated kinds-of-property in the field of clinical molecular genetics

The object of a clinical investigation is a human patient, a part of the patient or a part of the patient's close environment. This document is concerned solely with the terminology for human molecular genetics. Some of the external definitions or references to concepts used in NPU definitions may apply to all kinds of living organisms, but within this document they should be considered only in the context of a human object of investigation.

Terms used in NPU definitions are established with reference to authoritative external definitions, terminologies and nomenclatures where possible (see Terminological references). References are filed in the international NPU data repository. Where no valid reference exists, a local definition or explanation of the term is filed.

Some concepts are central in the field of clinical molecular genetics, and these terms will make up a large part of the semantic content of NPU definitions within this field. A selection of the most important concepts are listed below, with term, abbreviation in parenthesis where relevant, and with a defining reference where possible. Clinical molecular genetics is still an emerging field of science, and in some cases, it has proved difficult to locate a valid external terminological reference. For these concepts a definition or a description of intended meaning is introduced for the purpose of this document only. The list is not final or exclusive; concepts, references and definitions will be continuously added to the NPU data repository as development in the field of molecular genetics continues.

alleles: variant forms of the same *gene*, occupying the same locus on homologous *chromosomes*, and governing the variants in production of the same gene product [23].

clonality: kind-of-property indicating the property of cells related by descent from a single progenitor cell [24]. Value sets are ordinal.

chromosomes (chroms): In a prokaryotic cell or in the nucleus of a eukaryotic cell, a structure consisting of or containing *DNA* which carries the genetic information essential to the cell [23]. In this context the chromosomes contained in a human individual, or part of the individual, tissue or cell line [25], including protein components, *e.g.* histones.

copy number variation (cnv): kind-of-property designating *structural variation* in the number of copies of one or more specific sequences from a *reference sequence* of the same kind. The results are narrative.

NOTE: This includes deletion of sequences (number of copies = 0). 'Loss of Heterozygosity' (LOH) and 'gene amplification' are subtypes of specific copy number variations. Used to describe large copy segments, of thousands to millions of DNA bases, but there is no international agreement as to size limits. The reference sequence must be stated either via the NPU definition or as part of the report. Result values reported with the same NPU code cannot be assumed comparable unless investigated and reported by identical procedures.

disomy type: classification describing the parental origin of genetic material, *e.g.* maternal, paternal, uniparental or biparental disomy, isodisomy, heterodisomy.

DNA: a deoxyribonucleotide polymer that is the primary genetic material of all cells [23]. In this context DNA contained in a human individual, or part of the individual – tissue, cell line [25] or fluid of a patient. May be specified as to origin (for example leukocytes from blood, tumoral cells from renal tissue, or (cell-free) plasma).

DNA fragment: A sequence of *DNA*, shorter than a whole *chromosome*.

NOTE: Free DNA fragments are present in blood plasma. They represent genetic material from the person, including tumors and transplants, and, in pregnant women, genetic material from the fetus.

entitic number (entitic num.): entitic kind-of property designating the number of a component in a single entity of the system, *e.g.* the number of specific repeats of a sequence in the *genome*. The value set is a ratio scale of natural numbers.

exome: That part of the *genome* that corresponds to the complete complement of *exons* of a cell [23].

gene: A category of nucleic acid sequences that function as units of heredity and which code for the basic instructions for the development, reproduction, and maintenance of organisms [23].

genome: The genetic complement of an organism, including all of its genes, as represented in its *DNA*, or in some cases, its *RNA* [23]. In this context an entity of an individual's complete set of *DNA*, including all of its *genes*.

gene rearrangement: Any *DNA* sequence rearrangement that results in the creation of a novel protein-coding capacity [24].

immunoglobulin gene rearrangement: *gene rearrangement* that creates a novel junction among the immunoglobulin-coding variable, diversity, joining or constant segments [24].

haplotype: a set of *genes* that are closely linked and tend to be inherited together [23].

karyotype: An entity of the complete set of *chromosomes* in an individual, tissue or cell line [25].

methylation: The covalent chemical or biochemical addition of a methyl group(s) to a compound [24].

reference sequence: A sequence of like entities (nucleotides, amino acids) adopted by convention, with which any other sequence of the same kind can be compared to identify deviations from the sequence

NOTE: a reference sequence may be stated by reference to an internationally recognized source, *e.g.* NCBI Reference Sequence Database [26] for nucleotides. A reference sequence may have any length.

RNA: a polynucleotide consisting essentially of chains with a repeating backbone of phosphate and ribose units to which nitrogenous bases are attached. In this context the RNA in the body of a human individual. May be specified as to origin. Includes ribosomal RNA (rRNA), messenger RNA (mRNA), transfer RNA (tRNA), micro RNA (miRNA), and other non-coding RNA (ncRNA)

sequence variation (seq.var.): kind-of property designating *variation* consisting in deviations from a specified part of a *reference sequence* of the same kind. The results are narrative.

NOTE: The sequences are assumed to be fully investigated, even if only clinically relevant deviations are reported. Results reported with the same NPU code cannot be assumed comparable unless investigated and reported by identical procedures.

structural variation (struc.var.): kind-of-property designating *sequence variation* consisting of deviations in the position of one or more parts of the sequence from a *reference sequence* of the same kind. The results are narrative.

NOTE: Used to describe position changes (inversions, translocations, insertions, deletions) of large regions of DNA (approximately 1 kilobase and larger in size). The reference sequence must be stated either via the NPU definition or as part of the report. Results reported with the same NPU code cannot be assumed comparable unless investigated and reported by identical procedures.

variation: kind-of-property of entitic sequences (of *DNA*, *chromosomes*, *genes*, *RNA*, proteins) consisting in deviations from parts of a *reference sequence* of the same kind. The results are narrative.

NOTE: The particular sequences investigated must be stated as part of the report. The reference sequence must be stated either via the NPU definition or as part of the report. Results reported with the same NPU code cannot be assumed comparable unless investigated and reported by identical procedures. The abbreviation 'var.' for this term is avoided, to keep the expression in daily use of this very unspecific kind-of-property distinctly different from abbreviations of the more specific kinds-of-property 'seq.var.' and 'struc.var'.

transcriptome: An entity of the complete set of *RNA* transcripts from *DNA* in a cell or tissue. The transcriptome includes ribosomal RNA (rRNA), messenger RNA (mRNA), transfer RNA (tRNA), micro RNA (miRNA), and other non-coding RNA (ncRNA).

uniparental disomy (upd): *disomy type* indicating the presence in a cell of two paired *chromosomes* from the same parent, with no chromosome of that pair from the other parent [23].

4 Syntax and semantic content

4.1 The NPU concept model

An examination as a scientific process investigates one or more properties of a system. In human clinical sciences, the object of investigation may be a certain patient, a part of the patient, or part of his close environment; and in the field of clinical laboratory sciences the actual patient system investigated is often, but not always, represented by a physical sample.

The result of a clinical laboratory investigation, in *e.g.* a message or health care record concerning a certain patient, may be identified by stating the system investigated, the component studied in that system and the property estimated of that component in that system, with a measurement unit specified where relevant. The NPU terminology supplies definitions of such types of properties dedicated to specific systems and components, according to specific syntax and semantic rules [1]. Such a definition is also referred to as a 'dedicated kind-of-property'.

The semantic content is established with references to international nomenclatures, terminologies and classifications of the relevant scientific field where possible. The sources for semantic content in the field of human molecular genetics are listed in 'Terminological References'

The general structure of a NPU definition is:

System(system specification)—Component(component specification); kind-of-property (procedure specification) = ? unit of measurement

The parts System, Component and kind-of-property are mandatory for all definitions, as is a unit where relevant. Specifications may be added as needed. For everyday use an abbreviated string is employed, where the term of any part except the Component may be replaced by an abbreviation. This representation is used for most examples in this document.

A marker in place of a specification (written 'specification', abbreviated 'spec.', or 'procedure', abbreviated 'proc.'), indicates that the NPU definition does not supply all details needed to define the patient property investigated. The investigating laboratory must supply the clinically relevant missing information.

Each definition is assigned an identifier (NPUxxxxx) intended to represent that definition in health informatics applications. The identifier enables laboratory data to be transmitted, used, and safely stored in *e.g.* laboratory information management systems, electronic health care records, and research databases, and allows for variation in naming conventions. The connection between NPU identifiers and NPU definitions will always be available, and will not change over time or geography.

In this document established NPU definitions are shown with their actual NPU identifier, as filed in the NPU database. Where examples of definitiones are not (yet) in actual use, the potential identifier is represented by 'NPUxxxxx'.

EXAMPLE

Full definition: DNA(specification)—AIP gene; sequence variation
Abbreviated definition: DNA(spec.)—AIP gene; seq.var. = ?
NPU identifier: NPU28397

4.1.1 Naming investigations and their results

There is no demand that the NPU definition or its abbreviated expression serves as the only name or the common name for investigations or results in local systems, or in digital laboratory reports. The abbreviated NPU expressions have proved generally useful as primary names, but because of the formalized structure and wording, they can in some cases be very different from commonly used terms or be very long and unwieldy.

It is recommended that systems presenting laboratory data make the NPU abbreviated expressions and identifiers, together with local terms, available to users. Whether to use local terms or the NPU expressions as primary names is a matter of local preference.

4.1.2 System and system specification

Common systems in clinical molecular genetics are chromosomes, DNA and RNA. The system may be specified, partly specified or unspecified.

A marker in place of a system specification (written 'specification', abbreviated 'spec.'), indicates that the NPU definition does not necessarily supply all details on the system. Supplementary information on the exact system, *e.g.* that the investigation concerns a specific tissue of the patient, may need to be communicated by other means. If no such extra information is supplied, it may be assumed that the system is the germline system of the patient in question.

EXAMPLES

Chromosomes (Amniotic fluid cell)—
(Abbreviated: Chroms(Amf cell)—)
The system investigated is chromosomes of cells from amniotic fluid (of a pregnant patient).

DNA(Leukocytes; Blood)—
(Abbreviated: DNA(Lkcs; B)—
The system investigated is DNA of leukocytes in blood.

DNA(specification)—
(Abbreviated: DNA(spec.)—
The system investigated is DNA. As no origin of the DNA is specified with the result, it is assumed to be germline DNA.

DNA (Plasma)—
DNA fragment(Plasma) —
The systems investigated are DNA or DNA fragments in blood plasma. Fragments of DNA present in plasma may derive from genomic DNA of the person in question, from tumors or other specific cell clones of that person, or from a fetus carried by the person.

RNA(Tumor, specification)—
(Abbreviated: RNA(Tumor; spec.)—
The system investigated is RNA from a tumor. Further specification of *e.g.* the origin or nature of the tumor must be communicated by other means.

4.1.3 Component and component specification

As a general rule, for genes and chromosomes the component is the representative (entitic) instance, not the total number or mass of like elements in the system.

Specifications to the component may indicate a specific type, variant or part of the general concept. Specific variants may be described using a Locus Reference Genomic (LRG) reference sequence [27] followed by the variant description (according to the general recommendation of HGVS Variant Description Nomenclature). If a LRG reference is not available, the variant investigated may be described using a RefSeqGene record from the RefSeq database [26], listing both database accession and version number. For reasons of terminological consistency the genomic (NG_) sequences are preferred.

EXAMPLES

— BRCA1 gene;
The component investigated is the BRCA1 gene or unspecified parts of it.

— BRCA1 gene(LRG_292:g.124140)
The component investigated is a specific part of the BRCA1 gene, specified according to a reference sequence.

— BRCA1 gene(LRG_292:g.124140C > T)
The component investigated is a specific variant in a specific position of the BRCA1 gene, specified according to a reference sequence.

— JAK2 gene(LRG_612t1:exon 12);
The component investigated is a specific part of the JAK2 gene, exon 12, specified according to a transcript reference sequence.

— MT-ATP6 gene(NC_012920.1:m.8993T > G);
The component investigated is a specific variant in a specific position of the mitochondrial MT-ATP6 gene. The position is specified according to a reference sequence.

—HLA-B gene(*27);
The component investigated is a specific allele type of the HLA-B gene.

— Genome (16p2-p3);
The component investigated is a specific part of the DNA of chromosome 16.

—Genome (methylated);
The component investigated is methylated DNA.

— Genome (15q11-q13; methylated);
The component investigated is methylated DNA in the region 15q11-q13 of chromosome 15.

4.1.4 Kinds-of-property and procedure specification

The most frequent kinds-of-property in the field of human molecular genetics will be 'variation', 'sequence variation' (seq.var.), 'structure variation' (struc.var) and 'copy number variation' (cnv), all used with narrative result values.

Apart from 'sequence variation', which indicates that the full sequence in question has been investigated, all other kinds-of-property are used with the marker 'procedure' (proc.) to indicate that result values identified with the same NPU code may not be assumed comparable unless produced by identical procedures.

The technique employed and the variants investigated should be reported as a structured part of the result if clinically relevant.

To describe the variations recommendations from HGVS should be followed.

[gene or sequence]; variation (proc.) = ?
A set of variants has been investigated; possibly only clinically relevant findings are reported

[gene or sequence]; seq.var. = ?
Sequencing of the gene or sequence has been performed; possibly only clinically relevant findings are reported

[gene or sequence]; struc.var.(proc.) = ?
A set of structural variants has been investigated; possibly only clinically relevant findings are reported

[gene or sequence]; cnv (proc.) = ?
A set of copy number variants have been investigated; possibly only clinically relevant findings are reported
EXAMPLES

NPUxxxxx DNA(spec.)—BRCA1 gene; variation(proc.) = ?
The result describes unspecified variations in the BRCA1 gene, identified by *e.g.* a Multiplex Ligation-dependent Probe Amplification (MLPA) procedure. The variants investigated should be reported as part of the result value.

NPUxxxxx DNA(spec.)—BRCA1 gene (LRG_292:g.124140C > T); variation(proc.) = LRG_292:g.124140 [C > T];[?]
Narrative result value describes a patient having the specified sequence variation in at least one allele, presence of other variants have not been investigated.

NPU36073 DNA(spec.)—JAK2 gene; seq.var. = ?
Result describes mutations in the JAK2 gene, identified by sequencing. The report may contain only clinically relevant findings, although the sequencing has covered the whole gene.

NPUxxxxx DNA(spec.)—JAK2 gene (LRG_612:g.93526G > T); seq.var. =
LRG_612:g.[93526G > T];[=]
Narrative result value describes a patient heterozygote for the specified sequence variation, and both alleles have been fully investigated.

NPUxxxxx DNA(spec.)— F12 gene(LRG_145:c.983); seq.var. = LRG_145:c.983[C > G];[=]
Several nucleotide substitutions related to Angioedema Hereditary Type III are expected in this position. In the example, a C to G substitution in one allele is found, but the test performed also has assured that no other substitution (C > A, or C > T) has been found in that position in the other allele.

NPUxxxxx DNA(spec.)—FMR1 gene(NG_007529.1:g.5102GGC[size]); seq.var. =
(NG_007529.1:g.5102GGC [20];[40])

NOTE: This variant is generally referred to as "FMR1, (CGG)n EXPANSION". This is not according to HGVS nomenclature [18], which has the so called "3′ rule". This rule states that:

> for all descriptions the most 3′ position possible of the reference sequence is arbitrarily assigned to have been changed

Applied to the FMR1 gene and the repeat in the 5′UTR the unit of this repeat is GGC (it ends with …cggcg-gcggcggcggc ggc tgggcctc) and not CGG.

NPUxxxxx DNA(spec.)—PMP2 gene; struc.var.(proc.) = ?
The result describes an unspecified set of structural variants in or of the PMP2 gene. The variants investigated should be reported as part of the result value.

NPUxxxxx DNA(spec.)—PMP22 gene (duplication); struc.var.(proc.) = LRG_263:c.(?_-1)_(*1_?)dup
Where there is a need to discern between results of several investigations of related kind, a specification of the actual structural variation investigated, *e.g.* duplications, may be added.
This result value describes a duplication in the PMP22 gene that extends at least from first to last exon.

NPU56545 DNA(Plasma)—EGFR gene; cnv(proc.) = ?
The result describes a set of copy number variations in the EGFR gene. The variants investigated should be reported as part of the result value.

A specification of the actual copy number variation investigated may be added:
NPUxxxxx DNA(spec.)— ERBB2 gene(amplification); cnv(proc.) = ?
The result will describe a selective increase in the number of copies of the ERBB2 gene without a proportional increase in other genes.

NPUxxxxx DNA(B)—LDLR gene (LRG_274t1);cnv = LRG_274t1:c.(2311 + 1_2312-1)_(*807_?)del
The reference sequence is a transcript sequence and the narrative result value expresses deletion of exons

NPUxxxxx DNA(spec.)—NF1 gene (17q11.2); cnv(proc.) = ?
The narrative result will express the copy number variation in the 17q11.2 region in the NF1 gene.

The modifier 'entitic' is not used with the four kinds-of-property mentioned above, as it is implicitly assumed. For other kinds-of-property in general use in NPU definitions, with result value sets of ratio scale type and ordinal scale type, the modifier 'entitic' is needed if the object of investigation is a single instance of the system.

Ratio scale type. If a number of a variant in a single instance of the genome is stated as result, the kind-of-property is **'entitic number' (entitic num.)** with a value set of natural numbers stating the number of copies of the gene or the variant.

EXAMPLE

NPU54181 DNA(spec.)—FLG gene(LRG_1028:g.16819C > T); entitic num. = 1
The patient has the specified variant of the FLG-gene in exactly one allele (heterozygote)

Quantitative results based on counting totals in body fluids or tissues (*e.g.* DNA fragments in plasma) may be described using **'number concentration' (num.c.)** for fluid systems, **'number content'** for solid systems, or **'number fraction' (num.fr.)**, such as in NPU definitions describing cell counts in body fluids or tissues.

Ordinal scale type. When the result is a value on an ordinal scale, the term 'arbitrary' is added to the kind-of-property to indicate the lack of a reference measurement unit. For value sets describing the presence or absence, or some other sort of arbitrary count in a genetic entity, the kind-of-property will be **arbitrary entitic number (arb.entitic num.)**.

For value sets expressing amounts of substances or sequences in the whole system, the kind-of-property will be **'arbitrary content' (arb.cont.)** for solid systems and **'arbitrary concentration' (arb.c.)** for fluid systems.

EXAMPLES

NPUxxxxx DNA(spec.)—KRAS gene (NG_007524.1:g.10571G > T) arb.entitic num.(proc.) = 1
Only presence or absence of the sequence variation has been tested. The laboratory uses an arbitrary ordinal scale of 0 – not present, 1 – present. This patient has the sequence variation in at least one allele.

NPU53657 RNA(spec.)—SOX11 gene; arb.cont.(proc.) = Increased expression
The result expresses the amount of a specific gene sequence in the total RNA system. The value set used by the laboratory is {Decreased expression, Normal, Increased expression}. As RNA is a solid, the kind-of-property is 'arbitrary content'.

Nominal scale type. If the result is a type or classification with a limited and known value set, the kind-of-property will be '**category**'. The modifier 'entitic' is not needed for nominal value sets. The value set or classification system used should be specified by the laboratory.

EXAMPLES

NPUxxxxx DNA(B)—HLA-DQB1 gene; allele type; category(proc.) =
HLA-DQB1*[0101];[0302]
The laboratory classifies the HLA-DQB1 gene according to the allele type found.

NPUxxxxx DNA(Mola hydatidosa)—Ploidy type; category(proc.) = diploid
The laboratory classifies the ploidy as either diploid or polyploid.

NPUxxxxx DNA(lymphoblasts)—Ploidy type; category(proc.) = hyperdiploid
Investigation of the ploidy of blast cells in a patient with acute lymphoblastic leukemia. The result value set represents types of unbalanced chromosome complements (aneuploidy).
NOTE: An open-ended value set of {haploid, diploid, triploid, tetraploid, pentaploid …} expresses a number of copies, with the terms representing natural numbers on a ratio scale. The kind-of-property will be 'entitic number'.

4.1.5 Measurement unit

Measurement units are rare in this field. Where relevant, SI units or SI accepted units are used [1]. The unit 'one' (symbol '1') of dimensionless quantities is not stated.

5 Types of Dedicated kinds-of-property in human molecular genetics

5.1 Properties of specified DNA/RNA sequences

5.1.1 Single genes, or unspecified variants or sequences of single genes

The general model suggested in Part XVIII [14] is continued. The reference for gene names is the HUGO Gene Nomenclature Committee (HGNC) [28] and the result scale type is generally narrative, with four kinds-of-property of increasing specificity available (variation, sequence variation, structure variation, and copy number variation), depending on the scope of the investigation.

Result values that are narrative should include the specification of the actual sequence(s) investigated and reported on. Use of the HGVS [18] Nomenclature is strongly recommended.

EXAMPLES

NPUxxxxx DNA(spec.)—TYR gene; variation(proc.) = ?
The result is a narrative description of unspecified mutations identified by MLPA.

NPU19254 DNA(spec.)—ABCC8 gene; seq.var. = NG_008867.1:lg.86256 [A > G];[=]
The narrative description of a mutation identified in heterozygosis by the sequencing of the ABCC8 gene

NPU29749 RNA(spec.)—BTK gene; seq.var. = [LRG_128:g.34380 [A > G]
The narrative description of a mutation identified in the X chromosome of a male patient by sequencing of the BTK gene

NPUxxxxx DNA(spec.)—PAX6 gene(deletion); struc.var.(proc.) = ?
The result is a narrative description of unspecified deletions identified by MLPA.

NPUxxxxx DNA(Tumor; spec.)—MYCN gene (gene amplification); struc.var.(proc.) = ?
The result is a narrative description of a selective increase in the number of copies of the MYCN gene without a proportional increase in other genes

NPUxxxxx DNA(spec.)—RHD gene; entitic num. (0 1 2) = ? 0
The result is the number of instances of the RHD gene in the genome. No copy of the gene is found (Rhesus negative)

NPUxxxxx DNA(spec.)—RHD gene; entitic num.(0 1 2) = ? 1
The result is a number as above. One copy of the gene is found (heterozygote)

5.1.2 Specified variants or sequences of single genes

The model from Part XVIII [14] is adjusted; OMIM references to variants are replaced by **descriptions** of the actual sequences.

The choice of kind-of-property may vary depending on the communication needs and the specificity of the result.

If the result is based on the investigation of variants of a shorter specified sequence or a single nucleotide position, the kind-of-property is either **'variation'** or **'sequence variation'**, depending on whether the sequence or variant has been partly or fully investigated.

If the result describes structural changes of larger size, **'structure variation'** or **'copy number variation'** may be relevant.

If the result is numeric (the variant is present in 0, 1 or 2 copies, *i.e.* not present, heterozygote, or homozygote for the variant), the kind-of-property is **'entitic number'**.

If a specific sequence is indicated, and the result is given as not present or present (0, 1) *i.e.* a binary ordinal scale, the kind-of-property is **'arbitrary entitic number' (arb. entitic num.)**

EXAMPLES

NPUxxxxx DNA(spec.)— HFE gene(LRG_748:g.8671C > G); variation(proc.) = LRG_748;c.187[C > G];[?]
The component investigated is the specific variant of the HFE gene that has a G (guanine) nucleotide in position 8671 of the LRG_748 genomic reference sequence. Other variants have not been investigated. The result is narrative and describes that the variant is found in one copy. The laboratory reports the variant in terms of the coding string, not the genomic string, but the two expressions are equivalent and both are correct HGVS expressions.

NPUxxxxx DNA(spec.)—ATXN1 gene(LRG_863:g.438804_438806[size]; seq.var = NG_011571.1:g.438804_438806[dupGCA];[=]
The component investigated is (the number of) repetitions in the ATXN1 gene of the trinucleotide in position 438804_438806 of the LRG_863 genomic reference sequence. The result is narrative.

NPUxxxxx RNA(spec.)—ATXN2 gene(NM_002973.3:r.495_497[size]); seq.var. = ?
The component investigated is (the number of) repetitions in the expressed ATXN2 gene of the trinucleotide in position 495–497 of the stated mRNA reference sequence.

NPUxxxxx DNA(spec.)—CFTR gene(LRG_663:g.87846[size]); seq.var = ?
The component investigated is (the number of) repetitions in the CFTR gene of the nucleotide in position 87846 of the LRG_663 genomic reference sequence.

NPUxxxxx DNA(spec.)—DMPK gene(NG_009784.1:g.17294_17296CTG[size]; seq.var. = ?
The component investigated is (the number of) repetitions in the DMPK gene of the CTG trinucleotide in position 17294_17296 of the stated genomic DNA reference sequence.

NPUxxxxx DNA(spec.)— MT-ATP6 gene(NC_012920.1:m.8993T > G); arb.entitic num. = 1
The component investigated is the variant of the mitochondrial MT-ATP6 gene that has a G nucleotide in position 8993 of the stated mitochondrial reference sequence. The variant is present.

NPU54270 JAK2 gene(DNA; Marrow)—JAK2 gene(LRG_612: g.93526G > T); subst.fr. = ?
The component investigated is a somatic mutation of the JAK2 gene. The result value expresses the 'allelic load' in the bone marrow—the fraction of all instances of the gene that have the variant.

The sequencing of whole exons, especially of more than one exon at the same time, can be specified using a transcript reference sequence and the actual exon numbers (check if other NPU definitions specifying the same sequences or variants are already established with genomic references).

EXAMPLES

NPUxxxxx DNA(B) —JAK2 gene(LRG_612t1:exon 14);seq.var. = ?
Variation of the sequence of exon 14 of the JAK2 gene in DNA from blood

NPUxxxxx DNA(B) —CSF3R gene(LRG_144t1:exon 14-17);seq.var. = ?
Variation of the sequence of exon 14-17 of the CSF3R gene in DNA from blood

NPUxxxxx DNA(B) —SETBP1 gene(NM_015559.2:exon 6);seq.var. = ?
Variation of the sequence of exon 6 of the SETBP1 gene in DNA from blood

NPUxxxxx DNA(spec.)—SLC26A4 gene (NM_000441.1: exon 2,4,6,8,10); seq.var. = ?
Variation of the sequence of a specific non-consecutive set of exons of the SLC26A4 gene

DNA(spec.)—ZRSR2 gene(LRG_618t1); seq.var. = ?
Variation of the sequence of all exons of the ZRSR2 gene

5.1.3 DNA sequences larger than a single gene

Variants of DNA over a larger part of a chromosome than a single gene may be located via their chromosome location using the International System for Human Cytogenetic Nomenclature (ISCN) [25]. The component here is DNA, *i.e.* a specific part of the genome, not of the chromosome.

EXAMPLE

NPUxxxxx DNA(spec.)—Genome(17p13.3; deletion);cnv(proc.) = ?
The result is a narrative description of deletions identified by *e.g.* FISH technique in the 17p13.3 area. The report must specify the set of deletions investigated.

5.1.4 DNA fragments in blood plasma

The presence of specific DNA fragments in plasma may be estimated using an ordinal scale, or the actual number concentration of specific fragments may be counted.

EXAMPLES

NPUxxxxx DNA(P)—SRY gene; arb.cont.(proc.) = 1

Ordinal scale. Presence of the SRY gene in the plasma of a pregnant woman suggests a male fetus. The laboratory uses a binary scale {0 1} to express 'no presence' or 'presence'.

NPU53991 DNA-fragment(P)—DNA-fragment(Chromosome 13); arb.fr.(proc.) = low
Ordinal scale. The number of DNA fragments from Chromosome 13 in maternal plasma is estimated as a fraction of the total number of fragments. The laboratory uses an ordinal scale reflecting the fraction size, {low, high} to indicate fetal 'risk of trisomy 13'.

NPU53921 DNA-fragment(P)—DNA-fragment(Chromosome 13); arb.num.fr.(actual-norm; proc.) = ? (p.d.u.)
Ratio scale. The result value reflects the actual proportion of Chromosome 13 DNA fragments to the total number of DNA fragments in plasma, compared to a local normative reference.

NPU54354 DNA(P)—KEL gene(LRG_799:g.9496C > T); arb.cont.(proc.) = 1
Ordinal scale. The investigation aims to determine the presence of KEL blood type variant KEL*0101 in DNA fragments in a pregnant woman's plasma. The variant is found (and assumed to originate from the fetus, if it is not present in the maternal genome).

5.1.5 Fused genes and other structural changes

Fused genes may be identified with reference to the NCI Thesaurus [24]. If no reference is found there, a fusion may be described using ISCN nomenclature [25].

EXAMPLES

NPUxxxxx DNA(spec.)— ABL1/BCR fusion gene; struc.var.(proc.) = ?
Description of the structure of the fused ABL1/BCR gene

NPU56065 RNA(Tumor; spec.)—EML4/ALK fusion gene; arb.cont.(proc.) = 1
The laboratory reports detection of the expressed EML4/ALK fusion gene in RNA of a tumor on an ordinal scale, value set {0, 1}

NPU54794 B—ABL1/BCR Fusion Gene (RNA); num.c. = ? × 10^3/L
The result value is the number concentration of expressed fusion gene in blood

NPUxxxxx TCR-gene(B) – Gene-rearrangement; clonality (proc.) = ?
Ordinal scale type
The result value expresses the presence of genes belonging to a single clone of cells.

5.1.6 HLA and CYP genes

Allele nomenclatures for HLA [29] and CYP [30] gene systems are currently being aligned with the HGVS system. The specific HLA and CYP nomenclatures may be used to specify alleles until the alignment is completed.

EXAMPLES

NPUxxxxx DNA(spec.)—HLA-DQB1 gene; category(proc.) = DQB1[∗01:01];[∗05:02]
NPUxxxxx DNA(spec.)—CYP2D6 gene; category(proc.) = CYP2D6[∗4];[∗2A]
Nominal value sets. The result value sets are classifications consisting of all possible combinations of allelic variants of the genes in question.

NPUxxxxx DNA(spec.)—HLA-B gene(∗27); arb.entitic num.(proc.) = 0
NPUxxxxx DNA(spec.)— HLA-DQB1 gene (∗03:02); arb.entitic num.(proc.) = 1
NPUxxxxx DNA(spec)—HLA-DQA1 gene(∗05); arb.entitic num.(proc.) = 0
Here the laboratory uses the ordinal result scale {0,1} to indicate whether the variant is absent or present in at least one instance.

5.1.7 Haplotypes

Haplotypes are sets of loci that are closely linked and tend to be inherited together [23]. A haplotype motif is a commonly observed structural component consisting of such loci.

EXAMPLE

NPU xxxxx DNA(spec.)—KIR haplotype B motifs; entitic number = ?
The result describes the entitic number of haplotype motifs of Killer-cell Immunoglobulin-like Receptors (KIR)

5.2 One examination producing information on groups of genes

5.2.1 Grouped investigations ordered and reported together

Investigations may concern a whole group of related genes or variants. Results may be single values (classifications, patterns) or sets of values, *e.g.* one for each individual gene or variant.

If several result values are reported, the NPU list structure may be used. The NPU list header defines the complete property investigated and reports a set of individual results, each with their own NPU identifier.

List content examples may be filed with the NPU list entry as an illustration of the meaning of the investigation, or they may be omitted.

In human molecular genetics, sets of genes related to specific syndromes are often examined and reported on together. The reference for phenotypes (syndromes) should be OMIM [31] where possible. Where OMIM lists several related phenotypes corresponding to different genotypes, a phenotypic series ID may serve as a reference. If no phenotypic series ID is available, the reference should be to the OMIM phenotype numbered '1', provided it holds an overview of the related phenotypes.

WHO classifications ICD-10 or ICD-11 may be considered as references for genetic phenotypes if no OMIM reference is available.

Proprietary terms for commercially available panels or arrays are not sufficient to define a set of genes related to a patient property, as they are not based on international nomenclature or terminology and are not unique.

EXAMPLES

NPU54737 DNA(spec.)—Oculocutaneous albinism related gene; variation(list; proc.)
NPU56097 DNA(spec.)—C10orf11 gene; seq.var. = ?
NPU37622 DNA(spec.)—OCA2 gene; seq.var. = ?
NPU56095 DNA(spec.)—OCA2 gene; variation(proc.) = ?
NPU48317 DNA(spec.)—SLC45A2 gene; seq.var. = ?
NPU56096 DNA(spec.)—SLC45A2 gene; variation(proc.) = ?
NPU41365 DNA(spec.)—TYR gene; seq.var. = ?
NPU56090 DNA(spec.)—TYR gene; variation(proc.) = ?
NPU41370 DNA(spec.)—TYRP1 gene; seq.var. = ?

As the investigation may comprise result values with different kinds-of-property, *e.g.* 'cnv', 'seq.var', and 'variation', the list definition must have the least specific kind-of-property relevant (here 'variation').

NPUxxxxx DNA(spec.)— Hypercholesterolemia, familial-related gene;variation (list; proc.)

The result contains detailed information on an unspecified set of genes related to familial hypercholesterolemia. Results are reported for individual NPU-identified genes or variants.

NPUxxxxx RNA(spec.) —Transcriptome; variation(list; proc.) = ?

The result contains detailed information on the transcriptome. Results are reported for individually NPU-identified genes or variants.

The reported investigation result may technically be based on a group of individual investigations, but reduced to a single classification value, *e.g.* a gene signature, when reported. If only one result value is reported, a single NPU definition, rather than a list structure, is used.

EXAMPLES

NPUxxxxx DNA(Marrow)—T Cell Receptor genes; clonality(proc.) = ?
Ordinal scale. The result is based on examination of the genes TRB, TRG and TRD, but only one result value is given, concerning the presence of clones of these 3 genes

NPUxxxxx DNA(spec.)—Brugada syndrome related gene; category(pattern;proc.) = ?
The result expresses a constitutional type

5.2.2 Large gene domains

Large domains may be identified by specific domain terms.

EXAMPLES

NPUxxxxx DNA(spec.)—Genome; seq.var. = ?
Variation of the sequence of the full genome

NPUxxxxx RNA(spec.)—Transcriptome; seq.var. = ?
Variation of the sequence of the full RNA transcriptome

NPU56901 DNA(spec.)—Exome; seq.var. = ?
Variation of the sequence of a full set of coding DNA

NPUxxxxx DNA(spec.)—Mitochondrial DNA; seq.var. = ?
Variation of the sequence of the mitochondrial genome

NPUxxxxx DNA(spec.)—Genome (7q11.23); cnv(deletion; proc.) = ?
The ISCN nomenclature may be used to indicate larger sequences of DNA. The result describes deletions in a specific DNA area of chromosome 7.

5.3 Chromosomes and parts of chromosomes

Where the object of investigation is one or more chromosomes or parts of chromosomes, rather than specific genes, the component may be identified according to the ISCN system [25], and reporting should follow the ISCN nomenclature.

EXAMPLES

NPUxxxxx Chroms(spec.)—Karyotype; variation(proc.) = 47,XXX
The result indicates a karyotype with three X chromosomes

NPUxxxxx Chroms(Amf cell)—Karyotype; variation (proc.) = 46,XY
The amniotic fluid cells show a normal male karyotype

NPUxxxxx Chroms(Marrow)—Chromosomes (9, 22); struc.var. = t(9;22)(q34;q11)
An investigation for an unspecified structural rearrangement between chromosomes 9 and 22 has identified a translocation

NPUxxxxx Chroms(Tissue; spec.)—Chromosomes (t(11;22)(p13;q12)); arb.num. = 1
An investigation of two chromosomes for a specific translocation has found that the translocation is present

NPUxxxxx Chroms(spec.) —Chromosome 7; sequence variation = del(7)(p11.2)
The chromosome sequence of Chromosome 7 shows a deletion in the short arm

5.4 Modifications of DNA

Methylation of DNA/RNA [32] may be estimated using different scale types, kinds-of-property, and units:

EXAMPLE

NPUxxxxx DNA(spec.)—DNA(15q11.2-q13; methylated); arb.cont.(proc.) = |lom
Laboratory reports degree of methylation on an ordinal scale:{ |lom, normal, |gom}, where '|lom' indicates 'loss of methylation' and '|gom' indicates 'gain of methylation'

For quantitative results, specification of the examined part of the genome is necessary, or the result value will be ambiguous. If this specification is not part of the NPU definition, the information must be supplied as part of the report.

EXAMPLE

DNA(Ntrcs; spec.)— DNA(methylated); rel.subst.fr.(actual/norm; proc.) = 0.3

Parental imprinting is a class of DNA modification that may be described using HGVS nomenclature, ISCN nomenclature (with narrative value sets), or a nominal value set of 'disomy types', *e.g.* {uniparental, biparental}.

EXAMPLES

NPUxxxxx Chroms(spec.)—Karyotype, seq.var.(proc.) = 46,XY,upd(15)mat
The result value identifies uniparental disomy of chromosome 15, of maternal origin, in narrative form according to ISCN nomenclature

NPUxxxxx Chroms(spec.)—Chromosome 11(11p15.5); category(disomy type;proc.) = maternal
The result value classifies a part of chromosome 11 as having an uniparental disomy of maternal origin. The result value set is nominal {parental, maternal, biparental}

6 Elements of a data entry

A full data entry for a NPU definition consists of a set of data elements specifying:

- Code qualifier: "NPU" and code value (5 digits)
- Definition specified by joining defined term elements according to the NPU syntax
- Type of value set or scale type
- Group status where relevant (list header, context dependent)
- Classification according to branch of laboratory science
- History, present status (active, retired), synonyms and notes

EXAMPLE

NPU 54329
DNA(specification)—AQP1 gene(NG_007475.2:g.63650); sequence variation
Value set: Narrative
Molecular biology and genetics (MBG)
Established 2016.01.08, never revised
Status: Active
Synonyms: Aquaporin 1 (Colton blood group)

A full data entry for a concept (term element) consists of a set of data elements specifying

- Alphanumerical code value
- Term for the concept in English
- Abbreviation of term (optional)
- Reference to international terminology, classification, ontology *etc.* defining the concept, where available, or a formal definition
- History, present status (active, retired), synonyms and note

EXAMPLE

QU100148
Term: ZFP57 gene
Reference: HGNC 18791
Established/Last revised 15-05-2009
Status: Active

When published in readable form, the format may vary with the intended use. A formal presentation in printed media should contain at least:

- The unabbreviated expression in English, ordered according to the syntax. Including line breaks between the four main syntax parts may enhance readability. If no line breaks are used, the string '=?' should be placed after the kind-of-property and procedure specifications (if any) and before the unit, indicating the place where the result value would be expected.
- The type of result value set or result scale type.
- The code system qualifier and code value, followed by a space and the abbreviated definition.

Lists are traditionally printed with the list header in bold type, and the list content, if included, below and indented, in normal type.

EXAMPLE

DNA(specification)—
Alzheimer Disease related gene;
sequence variation(list)
Scale type: Narrative
NPU28729 DNA(spec.)—Alzheimer Disease related gene; seq.var.(list)
NPU19009 DNA(spec.)—APOE gene; seq.var. = ?
NPU30497 DNA(spec.)—APP gene; seq.var. = ?
NPU38775 DNA(spec.)—PSEN1 gene; seq.var. = ?
NPU38776 DNA(spec.)—PSEN2 gene; seq.var. = ?

Abbreviated expressions have proved sufficient for the identification of result types in most communication to and from laboratories (*e.g.* laboratory orders and reports, local documentation, statistics). In this publication, the abbreviated expression is used in all examples of NPU definitions. It is recommended that the unabbreviated definition, including status, value set or scale type, and terminological references are available from *e.g.* a national resource, in order to resolve cases of doubt.

7 Terminological references

List of references relevant for terms used in NPU definitions in this field and sources of reference identifiers

HGNC symbols for genes
http://www.genenames.org/

HGVS Sequence Variant Nomenclature
http://varnomen.hgvs.org/

RefSeq reference sequences – NCBI Reference Sequence Database
https://www.ncbi.nlm.nih.gov/refseq/

CYP gene allele nomenclature
http://www.cypalleles.ki.se/

HLA gene allele nomenclature
http://www.ebi.ac.uk/ipd/imgt/hla/

LRG reference sequences
http://www.lrg-sequence.org/

Genetic syndromes (OMIM phenotypic series)
https://www.ncbi.nlm.nih.gov/omim

NCI Term Browser (cell types, variants, phenotypes)
https://ncit.nci.nih.gov/ncitbrowser/pages/multiple_search.jsf?nav_type=terminologies

ISCN Cytogentic Nomenclature [25]

KIR haplotypes
http://www.ebi.ac.uk/ipd/kir/haplotypes.html

Membership of sponsoring bodies

Membership of the IFCC Scientific Division during the period 2013–2017 was as follows:

Chair: Philippe Gillery (France) from 2017, Ian Young (United Kingdom); **Vice-Chair:** Christa M. Cobbaert (Netherlands) from 2017, Philippe Gillery (France); **Secretary:** Joseph Passarelli (United States); **Members and Consultants:** Konstantinos Makris (Greece), Tsutomu Nobori (Japan), Mario Plebani (Italy), James F. Pierson-Perry (United States), Gary Myers (United States), Heinz Schimmel (Belgium), Chris Burns (United Kingdom), Karen W. Phinney (United States), Giampaolo Merlini (Italy), David Bunk (NIST Representative)

Membership of the IUPAC Chemistry and Human Health Division Committee during the period 2013–2017 was as follows:

President: Thomas Perun (United States); **Vice President:** Rita Cornelis (Belgium); **Secretary:** Michael Schwenk (Germany); **Titular Members:** Vincenzo Abbate (United Kingdom), Edmond Differding (Belgium), A. Ganesan (United Kingdom), Vladimir Gubala (United Kingdom), Linda Johnston (Canada), Helle Møller Johannessen (Denmark)

Membership of the IFCC Committee on Nomenclature for Properties and Units during the preparation of this report (2013–2017) was as follows:

Chair: Robert Flatman (Australia); **Members:** Urban Forsum (Sweden), Antonin Jabor (Czech Republic), Florian Scherrer (France), Karin Toska (Norway), R. Dybkaer (Denmark), Young Bae Lee Hansen (Denmark) (from 2017)

Membership of the IUPAC Subcommittee on Nomenclature for Properties and Units during the preparation of this report (2013–2016) was as follows:

Chair: Helle Møller Johannessen (Denmark); **Secretary:** Ulla Magdal Petersen (Denmark); **Members:** Ivan Bruunshuus (Denmark), Pedro de Araujo (Brazil), Johan den Dunnen (Netherlands), René Dybkaer (Denmark), Georges Férard (France), Robert Flatman (Australia), Urban Forsum (Sweden), Xavier Fuentes-Arderiu (Spain), Jens Hertz (Denmark), J. Gilbert Hill (Canada), Dongchon Kang (Japan), Daniel Karlsson (Sweden), Wolf Külpmann (Germany), Clement McDonald (United States), Gunnar Nordin (Sweden), Ariadna Padró i Miguel (Spain), Karin Toska (Norway), Young Bae Lee Hansen (Denmark).

Funding: International Union of Pure and Applied Chemistry, Funder Id: 10.13039/100006987, Grant Number: 2013-009-5-700.

Acknowledgements: This publication resulted from work carried out under IUPAC Chemistry and Human Health Division Project 2013-009-5-700.

References

[1] H. Olesen. *Pure Appl. Chem.* **67**, 1563 (1995).
[2] D. Kenny, H. Olesen. *Pure Appl. Chem.* **69**, 1015 (1997).
[3] I. Bruunshuus, W. Frederiksen, H. Olesen, I. Ibsen. *Pure Appl. Chem.* **69**, 2577 (1997).
[4] H. Olesen, D. Kenny, I. Bruunshuus, I. Ibsen, K. Jørgensen, R. Dybkær, X. Fuentes-Arderiu, G. Hill, P. Soares de Araujo, C. McDonald. *Pure Appl. Chem.* **69**, 2583 (1997).
[5] M. Blombäck, R. Dybkær, K. Jørgensen, H. Olesen, S. Thorsen. *Pure Appl. Chem.* **69**, 1043 (1997).
[6] H. Olesen, D. Cowan, I. Bruunshuus, K. Klempel, G. Hill. *Pure Appl. Chem.* **69**, 1081 (1997).
[7] U. Forsum, H. Olesen, W. Frederiksen, B. Persson. *Pure Appl. Chem.* **72**, 555 (2000).
[8] R. Cornelis, X. Fuentes-Arderiu, I. Bruunshuus, D. Templeton. *Pure Appl. Chem.* **69**, 2593 (1997).
[9] H. Olesen, I. Ibsen, I. Bruunshuus, D. Kenny, R. Dybkær, X. Fuentes-Arderiu, G. Hill, P. S. De Araujo, C. McDonald. *Pure Appl. Chem.* **72**, 747 (2000).
[10] H. Olesen, D. Kenny, R. Dybkær, I. Ibsen, I. Bruunshuus, X. Fuentes-Arderiu, G. Hill, P. S. De Araujo, C. McDonald. *Pure Appl. Chem.* **69**, 2607 (1997).
[11] H. Olesen, D. Cowan, R. de la Torre, I. Bruunshuus, M. Rohde, D. Kenny. *Pure Appl. Chem.* **72**, 479 (2000).
[12] H. Olesen, A. Giwercman, D. M. De Kretser, D. Mortimer, H. Oshima, P. Troen. *Pure Appl. Chem.* **69**, 2621 (1997).
[13] I. Bruunshuus, L. K. Poulsen, H. Olesen. *Pure Appl. Chem.* **72**, 1067 (2000).
[14] P. S. de Araujo, B. Zingales, A. Blanco-Font, X. Fuentes-Arderiu, C. Mannhalter, K. Varming, S. Bojesen, I. Bruunshuus, H. Olesen. *Pure Appl. Chem.* **76**, 1799 (2004).
[15] K. Varming, U. Forsum, I. Bruunshuus, H. Olesen. *Pure Appl. Chem.* **75**, 1477 (2003).
[16] J. Duffus, I. Bruunshuus, R. Cornelis, R. Dybkær, M. Nordberg, W. Kuelpmann. *Pure Appl. Chem.* **79**, 87 (2007).
[17] U. M. Petersen, R. Dybkær, H. Olesen. *Pure Appl. Chem.* **84**, 137 (2012).
[18] Human Genome Variation Society [Internet]. [cited 2018 Feb 2]. Available from: http://www.hgvs.org/.
[19] J. T. den Dunnen, R. Dalgleish, D. R. Maglott, R. K. Hart, M. S. Greenblatt, J. McGowan-Jordan, A. F. Roux, T. Smith, S. E. Antonarakis, P. E. Taschner. *Hum. Mutat.* **37**, 564 (2016).
[20] G. Férard, R. Dybkaer, X. Fuentes-Arderiu. *Compendium of Terminology and Nomenclature of Proeperties in Clinical laboratory Sciences Recommendations 2016*, p. 182, The Royal Society of Chemistry, Cambridge (2016).
[21] S. S. Stevens. *Science* **103**, 677 (1946).
[22] G. Nordin, R. Dybkaer, U. Forsum, X. Fuentes-Arderiu, F. Pontet. *Pure Appl. Chem.* **90**, 913 (2018).
[23] Medical Subject Headings (MeSH) [Internet]. U.S. National Library of Medicine. [cited 2018 Feb 2]. Available from: http://www.nlm.nih.gov/mesh/meshhome.html.
[24] NCI Thesaurus [Internet]. [cited 2018 Feb 2]. Available from: http://ncit.nci.nih.gov/.
[25] L. G. Shaffer, J. McGowan-Jordan, M. Schmid (Eds.) ISCN 2013. *An International System for Human Cytogenetic Nomenclature*, p. 140, S. Karger, Basel (2013).
[26] National Center for Biotechnology Information. NCBI Reference Sequence Database [Internet]. International Nucleotide Sequence Database Collaboration (INSDC). [cited 2018 Feb 2]. Available from: http://www.ncbi.nlm.nih.gov/refseq/.
[27] Locus Reference Genomic (LRG) [Internet]. [cited 2018 Feb 2]. Available from: http://www.lrg-sequence.org/.
[28] Hugo Gene Nomenclature Committee (HGNC) [Internet]. [cited 2018 Feb 2]. Available from: http://www.genenames.org/.
[29] IMGT/HLA Database [Internet]. [cited 2015 Mar 13]. Available from: http://www.ebi.ac.uk/ipd/imgt/hla/.
[30] L. V. Kalman, J. A. G. Agúndez, M. L. Appell, J. L. Black, G. C. Bell, S. Boukouvala, C. Bruckner, E. Bruford, K. Caudle, S. A. Coulthard, A. K. Daly, A. Del Tredici, J. T. den Dunnen, K. Drozda, R. E. Everts, D. Flockhart, R. R. Freimuth, A. Gaedigk, H. Hachad, T. Hartshorne, M. Ingelman-Sundberg, T. E. Klein, V. M. Lauschke, D. R. Maglott, H. L. McLeod, G. A. McMillin, U. A. Meyer, D. J. Müller, D. A. Nickerson, W. S. Oetting, M. Pacanowski, V. M. Pratt, M. V. Relling, A. Roberts, W. S. Rubinstein, K. Sangkuhl, M. Schwab, S. A. Scott, S. C. Sim, R. K. Thirumaran, L. H. Toji, R. F. Tyndale, R. van Schaik, M. Whirl-Carrillo, K. Yeo, U. M. Zanger. *Clin Pharmacol. Ther.* **99**, 172 (2016).
[31] Online Mendelian Inheritance in Man® [Internet]. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine. 1987 [cited 2017 Aug 15]. Available from: http://omim.org/.
[32] D. Monk, J. Morales, J. T. den Dunnen, S. Russo, F. Court, D. Prawitt, T. Eggermann, J. Beygo, K. Buiting, Z. Tümer; Nomenclature group of the European Network for Human Congenital Imprinting Disorders. *Epigenetics* **13**, 117 (2018).