orphanet

| Version 01 | April 2017

Procedural document
Orphanet inventory of genes related to rare disorders

Instituts thématiques Inserm
Institut national de la santé et de la recherche médicale

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE
MINISTÈRE DES AFFAIRES SOCIALES ET DE LA SANTÉ

20102206

Table of contents

I. Introduction

1. Purpose/objectives

This procedural document aims to describe how the Orphanet inventory of genes related to rare disorders is maintained and updated.

In order to better define rare disorders of genetic origin, Orphanet provides information on every gene related to a rare disorder. This information includes the genetic international nomenclature, the gene typology, the chromosomal location, the cross-mappings with other international genetic databases. Orphanet also defines the relationship between genes and their related rare disorders and provides the evidence for establishing these gene-disorder relationships.

2. Disclaimer

- This publication is part of Joint Action 677024 RD-ACTION which has received funding from the European Union's Health Programme (2014-2020).

- The content of this publication represents the views of the author only and is his/her sole responsibility; it can not be considered to reflect the views of the European Commission and/or the Consumers, Health, Agriculture and Food Executive Agency or any other body of the European Union. The European Commission and the Agency do not accept any responsibility for use that may be made of the information it contains.

- Candidate genes and biomarkers are excluded from the inventory unless a clinical test is made available in one of the Orphanet consortium's countries.

3. Range of application

The present procedure applies to all genes listed in the Orphanet database.
The Orphanet inventory of genes is managed by the gene data manager, under responsibility of the rare disorder database manager. The Orphanet Advisory Board on Genetics and external experts are regularly consulted to adjust decisions. The Orphanet Advisory Board on Genetics was also consulted for the establishment of this procedure.

4. References

Some international genetic databases are referred to the Orphanet inventory of genes, they are:

- HUGO Gene Nomenclature Committee - HGNC.
- Online Mendelian Inheritance in Man - OMIM
- Universal Protein Resource - UniProt
- Ensembl - Ensembl
- International Union of Basic and Clinical Pharmacology - Database- IUPHAR-DB
- Reactome database - Reactome
- The GenAtlas database - GenAtlas

- Regulation (EC) N°141/2000 of the European Parliament and of the Council of 16 December 1999 on orphan medicinal products)

Procedural documents:

- Orphanet inventory of rare diseases.
- The membership, mandate and role of the Orphanet Advisory Board on Genetics is available here.

5. Definitions

A **typology of genes** is used to differentiate between the following:

- **Gene with protein product**: Basic unit of heredity, consisting of a segment of DNA arranged in a linear manner along a chromosome that is transcribed in RNA and translated in a protein.
- **Disorder-associated locus:** Chromosomal region associated with a single heritable disorder. The heritable disorder may be mapped to a chromosome but generally has not been associated to a specific gene.
- **Non-coding RNA:** RNA encoded by a gene but not translated into a protein (ie: Transfer RNA).

Relationships between genes and rares disorders are made explicit:

- **Disease-causing germline mutation(s) in**: A gene mutation in a germ cell that is sufficient to produce the disorder and that can be passed on to offspring.
- **Disease-causing germline mutation(s) (gain of function) in:** A gene mutation in a germ cell that provides a new function of the corresponding protein that is sufficient to produce the disorder and that can be passed on to offspring.
- **Disease-causing germline mutation(s) (loss of function) in:** A gene mutation in a germ cell that impairs the function of the corresponding protein and that is sufficient to produce the disorder and that can be passed on to offspring.
- **Disease-causing somatic mutation(s) in:** A gene mutation in a somatic cell that is sufficient to produce the disorder but that cannot be passed on to offspring.
- **Major susceptibility factor in:** A gene mutation in a germ cell that predisposes to the development of a disorder, and that is necessary but not sufficient to develop the disorder.
- **Modifying germline mutation in:** A gene mutation in a germ cell that modifies the clinical presentation of the disorder and that can be passed on to offspring.
- **Part of a fusion gene in:** A coding or regulatory DNA sequence from a gene that has fused with another coding and/or regulatory DNA sequence from a different gene.
- **Role in the phenotype of:** A gene included in a chromosomal rearrangement, and proved to have a major influence in the phenotype of the chromosomal rearrangement.
- **Biomarker tested in:** A gene in which a variation is used to monitor disease activity and/or patient outcome.
- **Candidate gene tested in:** A gene in which a mutation is suspected, but not yet proven, to be responsible for a disorder, and that is tested for in a clinical setting.

Chromosomal location: Cytogenetic location on the chromosome of a gene with protein product, non-coding RNA or disorder-associated locus.

Rare disorder: A rare disorder is defined according to the European legislation defining a prevalence threshold of not more than 5 affected persons per 10’000 (Regulation (EC) N°141/2000 of the European Parliament and of the Council of 16 December 1999 on orphan medicinal products)

Group of phenomes: A group of phenomes is a collection of clinical entities, sharing a given characteristic and that are therefore classified together.

Subtypes of a disease: Subtypes are a set of sub-forms of a disorder that could be clinical, aetiological or histopathological.

Genetic functional study: Genetic functional studies aim to explain the physiological and pathophysiological roles of a gene. The functional studies used to validate the association of a gene with a disorder can be carried out in animal models such as geneknock-out, cell experiments such as restoration of a gene/protein activity and/or *in vitro* experiments such as testing the gene expression level and/or protein activity.

Experts: Experts mentioned in this procedural document are the health professionals identified by Orphanet as leaders in the medical field for a rare disorder or a group of rare disorders.

Orphanet Advisory Board on Genetics: This Board is composed either by geneticists who are members of the Orphanet Management Board and who join the Advisory Board on a volunteer basis, or by geneticists not belonging to the Orphanet Management Board and invited by the Orphanet Management Board to join the Advisory Board. Rules and procedures of the Orphanet Advisory Board on Genetics are here.

Biomarker: A molecule used to monitor disease activity and/or patient outcome.

Somatic mosaic mutation: a postzygotic mutation that results in co-occurrence of two genetically distinct populations of cells within an individual.

Penetrance: the proportion of individuals that express an associated phenotype within a population carrying a particular variant of a gene.

Oligogenic: Inheritance of a disorder in which the combination of mutated alleles of two or more genes are sufficient to express the phenotype.

6. Filing and updates

This procedural document is updated at least annually and as often as necessary by the gene data manager. The most up-to-date version is available on the Orphanet website:
http://www.orpha.net/orphacom/cahiers/docs/GB/eproc_Orphanet_inventory_genes_PR_R1_Gen_02.pdf

II. METHODOLOGY

1. Flowchart

Figure 1: Flowchart of the procedure

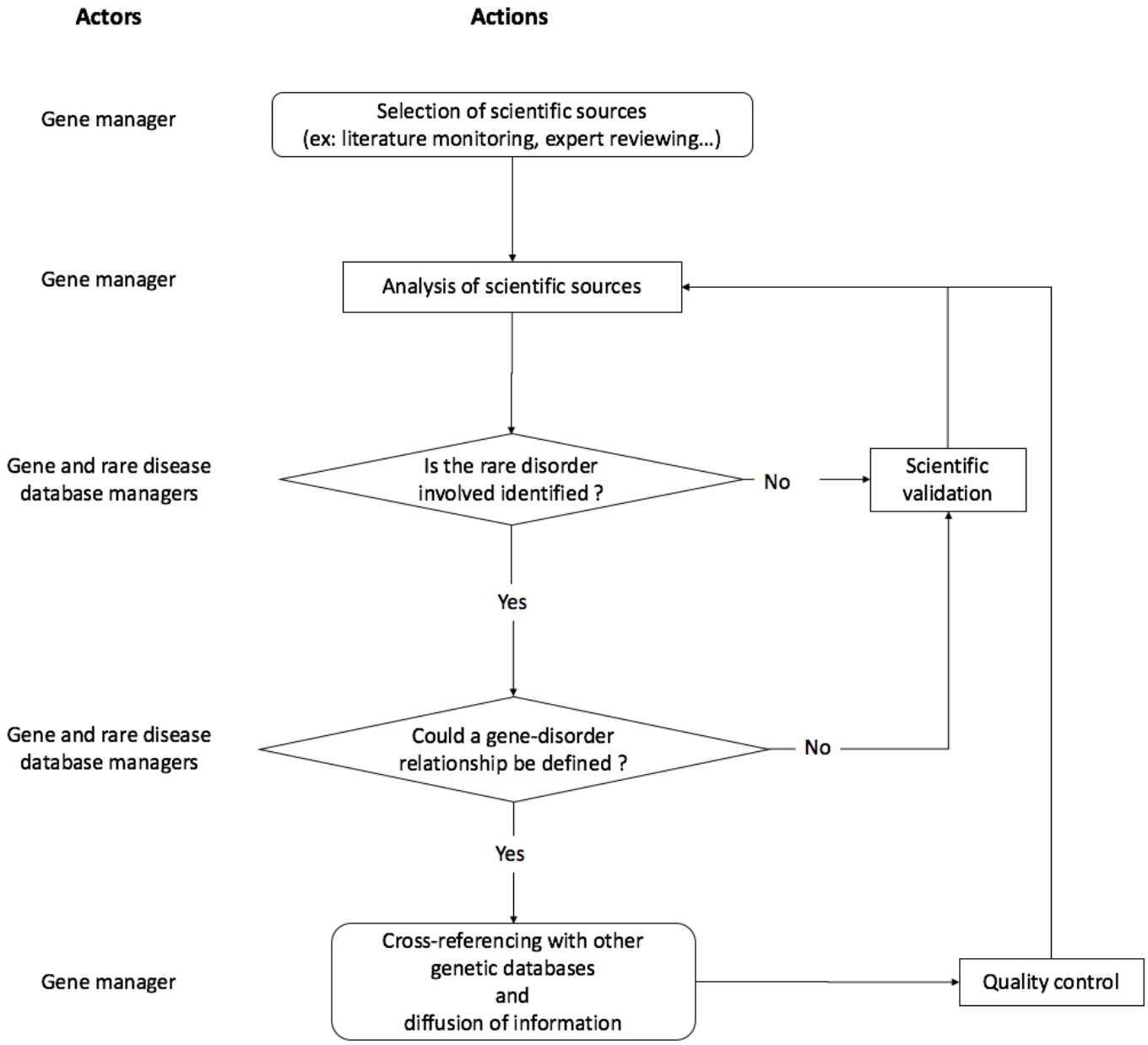


Figure 2: Decisional diagram: gene-disorder relationship for a germline point mutation in a gene

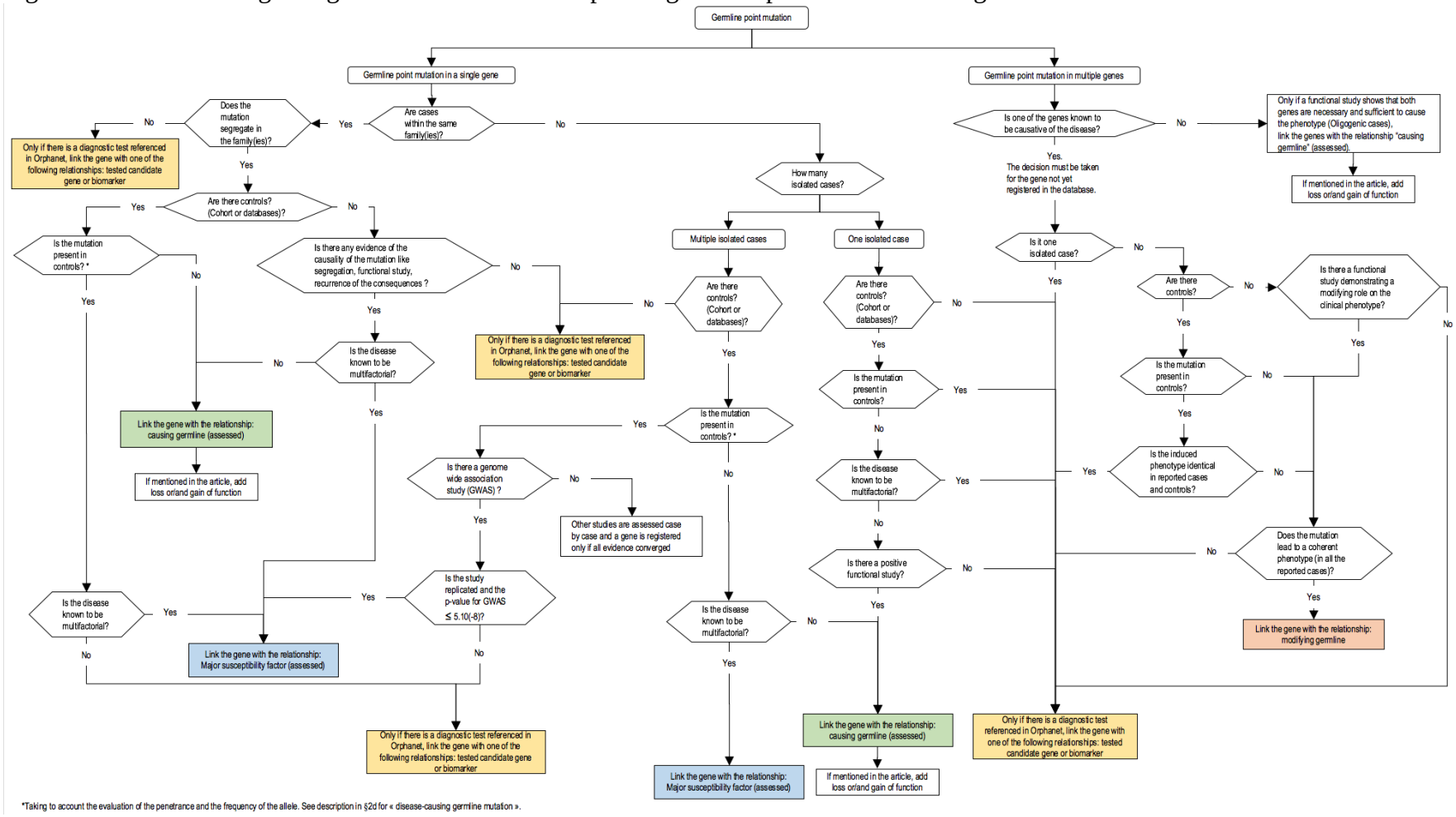


Figure 3: Decisional diagram: gene-disorder relationship for a somatic point mutation in a gene

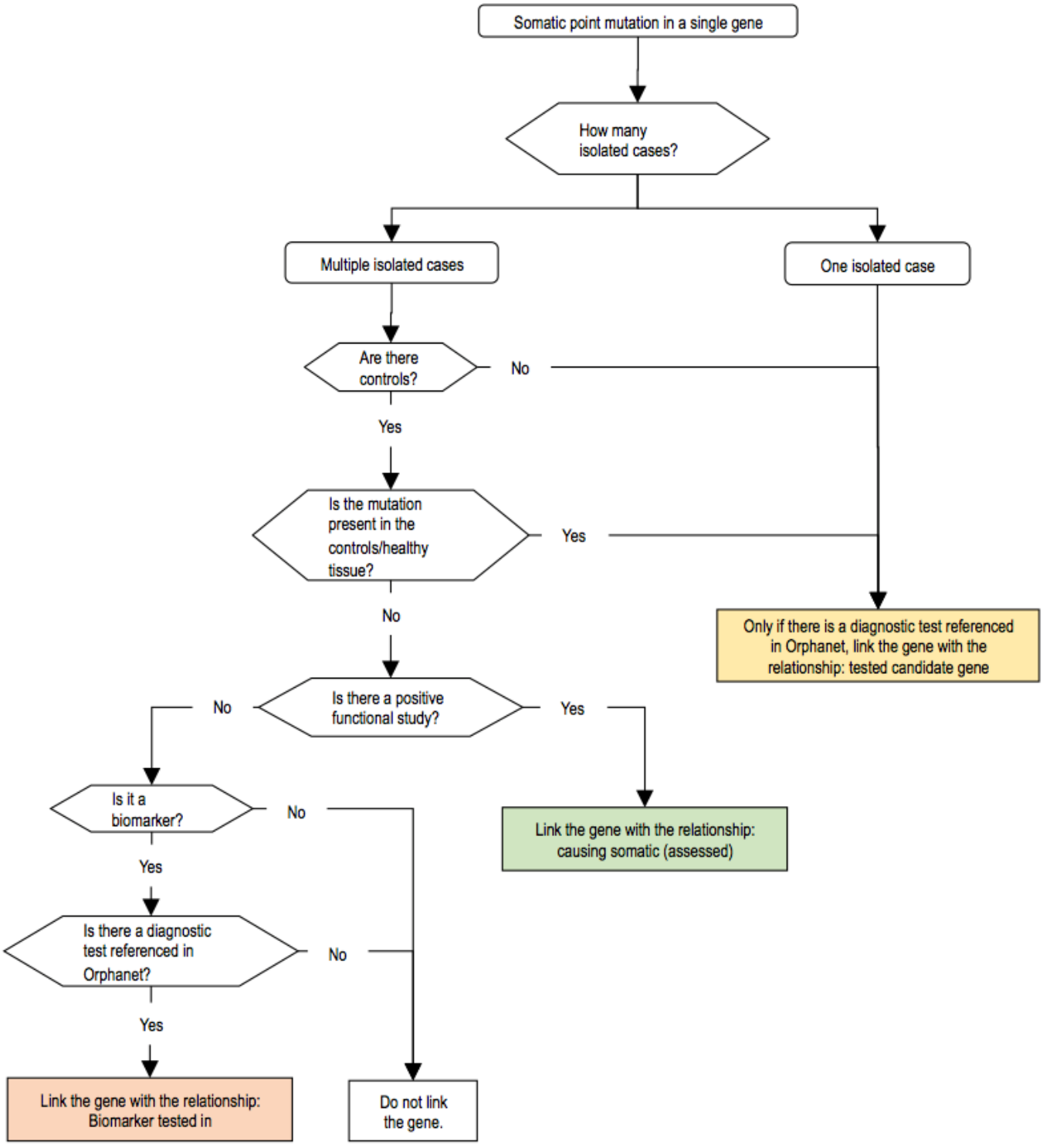
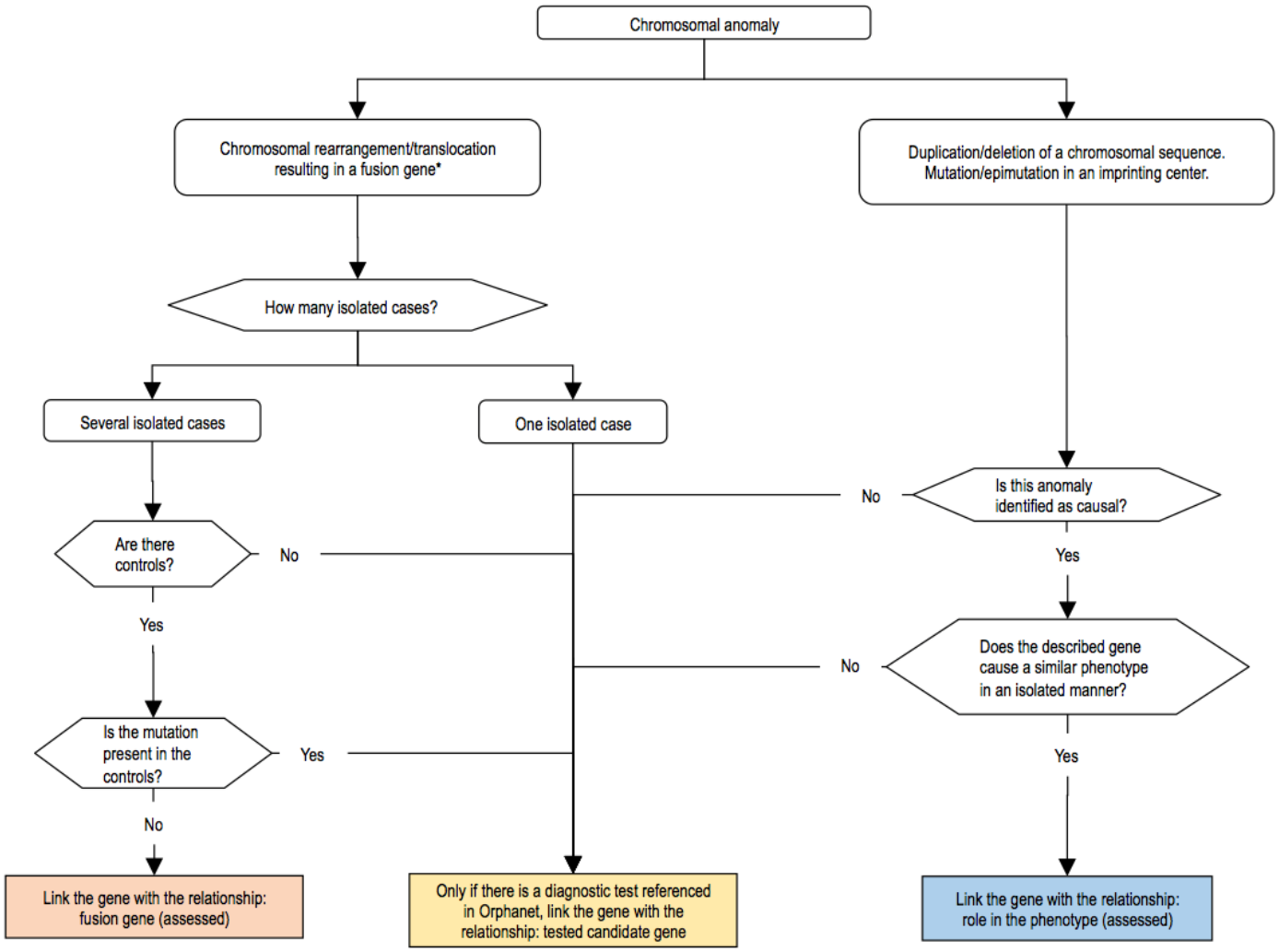


Figure 4: Decisional diagram: gene-disorder relationship for a chromosomal anomaly



*see definition section

2. Description

The whole flowchart, from sources to diffusion of information is described in figure 1 and detailed below.

a. Scientific sources

Different sources highlight the need for creation or update of a gene entry:

- Literature monitoring.
- Surveys of expert websites.
- Expert advice (spontaneously or upon request by Orphanet).
- Requests from national information scientists of the Orphanet Consortium.
- Cross-referencement with medical and/or genetic databases (such as OMIM).

b. Documentation

Only peer-reviewed publications are consulted. Decisions are based on publications establishing the gene-disorder relationship only. As an exception, candidate genes and biomarkers can be recorded without peer-reviewed publication following a request from an expert regarding a clinical test in use.

c. Identification of the related rare disorder

Based on the publication(s), the rare disorder is identified in the Orphanet rare disorder inventory. If needed, the addition of a new disorder is studied according to the inventory of rare disorders operating procedures.

d. Identification of the gene-disorder relationship

The gene-disorder relationship is qualified according to the pathogenicity of the described mutations. Thus, genes are annotated as causative (from germline or somatic mutations), modifiers, major susceptibility factors or playing a role in the phenotype (for chromosomal anomalies). When causative mutations are of germline origin, loss or gain of the protein function are documented if available.
Two relationships are specifically dedicated to candidate genes and biomarkers, that are registered only if a clinical diagnostic genetic test is made available in one of the Orphanet consortium's countries.

Criteria to link a gene to a disorder differ from a relation to another. Decisional diagrams (figure 2, 3 and 4), validated by the Orphanet Advisory Board on Genetics, allow the data manager to standardise this step. However, the scientific knowledge of the data manager cannot be easily represented in a simple diagram and a number of the decisional steps taken depend on the data manager's expertise.

The main decisional criteria are described below.

Disease-causing germline mutation(s) in

Every gene identified with a causative germline mutation is linked to a rare disorder. The study demonstrating the causality has to fulfil all of the following criteria:

- The mutation segregates in one or more family(ies) or is identified in several isolated cases. Segregation of the putative pathogenic mutation is assessed case by case. The structure of the family, the mode of inheritance and the penetrance of the disease are carefully examined.
- The mutation is not found in the general population (absent from control cohort and/or genetic variants databases). In case of low penetrance, the mutation's occurrence in a control population is permitted. Then, causality is assessed according to the degree of penetrance, the frequencies of the mutation in the patient and control populations. In this specific case, functional studies may help to confirm the relationship.

- If there is only one isolated case, functional studies are needed to confirm the association.
- If the inheritance is digenic (or oligogenic more generally), functional studies are needed to confirm the association.

Functional studies are assessed case by case.
If available, loss or gain of the protein function are documented, and both relations *Disease-causing germline mutation(s) in (gain-of-function)* and *Disease-causing germline mutation(s) in (loss-of-function)* may coexist for a single couple gene-disorder.

Disease-causing somatic mutation(s) in
Every gene identified with a causative somatic mutation is linked to a rare disorder. The study demonstrating the causality has to fulfil the following criteria:

- The mosaic mutation is found in several patients.
- The mosaic mutation is absent from controls (persons or healthy tissues).
- Functional studies confirm the association.

Functional studies are assessed case by case.
If there is only one isolated case, the gene is not registered in the database.
Concerning cancers, the causality of the mutation is considered if the mutation is an early event in oncogenesis and if an *ex vivo* or *in vivo* study confirms its oncogenicity.

Major susceptibility factor in
Only germline mutations are considered. The study has to fulfil the following criteria:

- The variant segregates in a family or is present in several isolated cases.
- The variant is usually found in the general population and/or the disease is known to be multifactorial.
- If the variant has been identified by a Genome wide association study (Gwas), only meta-analysis and a p-value less than 5.10(-8) are considered.

Currently, there is no accepted consensus for other types of association studies (ie: whole exome sequencing). Consequently, these studies are assessed case by case and gene is registered only if all evidence converges (replicated studies, functional studies …). Isolated cases are not considered.

Modifying germline mutation in
A modifying germline mutation is considered only if the causative gene is already identified in the patients described in the study. The study has to fulfil the following criteria:

- The variant is present in several cases.
- The variant is absent from the general population or does not lead alone to a similar phenotype.
- The consequences of the mutation are identical from a patient to another.

Isolated cases are not considered.

Part of a fusion gene in
Genes fused by chromosal rearrangement are registered only if the molecular study shows evidence of the causality of the fusion.The study has also to fulfil the following criteria:

- The fusion is observed in several cases.
- The fusion is absent from controls.

Role in the phenotype of
A gene with a major role in the phenotype of a chromosomal rearrangement is linked to a disorder only if amutation hitting only this gene gives rise to similar phenotypic consequences when it occurs in isolation.

Mutations or epimutations in an imprinting centre are also considered as having a major role in a phenotype as the phenotypic consequences are similar to those of a chromosomal rearrangement.

Candidate gene tested in or Biomarker tested in
Candidate genes or biomarkers are registered only if a clinical genetic test is available in at least one of the Orphanet consortium's countries.Depending on the purpose of the test, either for diagnosis or to monitor disease activity, the candidate gene or biomarker relationship is attributed.

e. Scientific validation

Scientific validation is sometimes necessary to confirm the identification of the related disorder and/or its gene relationship. These particular cases are first submitted for scientific validation to the rare disorder database manager. If necessary, experts on a specific disease or members of the Orphanet Advisory Board on Genetics are also consulted.

f. Update of the Orphanet gene database

When a request fulfils all the inclusion criteria, the gene is introduced in the database. Computational cross-referencement with international genetic databases is completed and allows to inform on:

- Gene nomenclature, including main name, symbol, synonyms and previous names and symbols according to HGNC.
- Typology of the genetic entity.
- Chromosomal location.
- Mappings with HGNC, OMIM, UniProtKB, Genatlas, Ensembl, Reactome and IUPHAR-DB.

Every gene-disorder relationship that has been curated and fulfils our inclusion criteria is mentioned as « assessed ». As exceptions, some gene-disorder relationships waiting for expert validation are not « assessed ». Candidate genes are never « assessed » as there is no scientific evidence to attest the gene-disorder relationship.

g. Quality control

Quality control of the gene-disorder relationship is carried out according to new peer-reviewed publications or expert advices. Quality control of the cross-referencement is carried out every month by an automatic mapping with HGNC database in order to maintain an uptodate the nomenclature and referencing of the inventory of genes.

The consistency of the database is monitored bi-annually based on systematic queries. The inconsistencies observed are either corrected or removed.

III. Availability of data

The Orphanet inventory of genes is released via various different dissemination channels at different frequencies (daily for the website - www.orpha.net -, monthly for the Orphanet download platform - www.orphadata.org - and the Orphanet Rare Disease Ontology – ORDO).

For any questions or comments, please contact us: contact.orphanet@inserm.fr

Editors of this procedural document: Catherine Gonthier, Annie Olry- This procedural document has been approved by : Ana Rath - Quality control : Charlotte Gueydan

The correct form when quoting this document is :

« Procedural document on the Orphanet inventory of genes related to rare disorders , Orphanet, April 2017, Number 01 : http://www.orpha.net/orphacom/cahiers/docs/GB/eproc_Orphanet_inventory_genes_PR_R1_Gen_02.pdf