New concept proposal

Transcript

Author	Jan Armida	Date last updated	10/11/2022
Project	Genomic Concepts	Contact person	Jan Armida
Status	Accepted	Consulted expert	WG

1   Rationale

During transcription, an RNA molecule is synthesized using a DNA template. The resulting transcripts are characterised by their capability to either coding or not-coding for a protein, their number of exons and total length in base pairs. Depending on their specific expression patterns, alternative transcripts can have clinically-relevant implications. Concerning alternative transcripts and tissue expression patterns in variant interpretation is of outmost importance to avoid missing diagnosis and incorrect diagnostic conclusions. Therefore, it is important to enable the possibility to refer not only the canonical transcript, but also other transcripts that might have biologically and clinically relevance for the variant interpretation process. The Transcript concept allows the use of the transcript identifier deposited in different repositories to point to a transcript of interest.

2   Comparison to other standards/data models

2.1   HL7 FHIR

In HL7 FHIR a transcript can be described using the element "*Transcript reference sequence ID*", which is described as the identifier for the transcribed reference sequence that is part of the genetic reference sequence that gets converted to messenger RNA.

A project of

SAMW ASSM
Schweizerische Akademie der Medizinischen Wissenschaften
Académie Suisse des Sciences Médicales
Accademia Svizzera delle Scienze Mediche
Swiss Academy of Medical Sciences

SIB
Swiss Institute of Bioinformatics

SIB | Swiss Institute of Bioinformatics
PHI | Personalized Health Informatics Group
www.sphn.ch | dcc@sib.swiss

3 Concept information

Contextualized concept name	Contextualized description	Type	Standard	Value set	Meaning binding
Transcript	RNA molecules that are made from a DNA template				SO: 0000673 transcript
transcript identifier	unique transcript id according to a specific nomenclature, e.g. Ensembl transcript ID	Code	Ensembl or other		
organism	organism associated to the transcript	Organism			
protein	protein product of the transcript	Protein			

4 Impact on the SPHN Dataset

The addition of *Transcript* does not require any further change in the current SPHN Dataset release.

5 Discussion

The possibility to refer to a transcript as a separate instantiation and not only when associated to a specific variant increases the flexibility and reusability of the concept in different contexts. In addition, the concept allows to specify the connection between a transcript and its protein product.