
UTA Documentation

Release 0.1.7

Reece Hart

August 18, 2015

1	Contents	3
1.1	Overview	3
1.2	Examples	3
1.3	Getting Started	3
1.4	Installation	3
1.5	Database	3
1.6	Modules	3
2	Additional Resources	5
3	Indices and tables	7

UTA is a data archive and set of tools that aims to improve the precision of sequence-based descriptions of genomic variation. It will eventually contain transcripts and alignments to multiple reference genomes using multiple algorithms, as obtained from [NCBI](#), [Ensembl](#), [LRG](#), and [UCSC](#).

Contents

1.1 Overview

The UTA stores transcripts aligned to genome reference assemblies using multiple methods in order to improve the precision and accuracy by which the scientific and clinical communities describe variants.

It facilitates the following:

- enabling an interpretation of variants reported in literature against obsolete transcript records
- identifying regions where transcript and reference genome sequence assemblies disagree
- characterizing transcripts of the same gene across transcript sources
- projecting (“lifting over”) variants from one transcript to another
- identifying transcripts and genomic regions with ambiguous alignments that may affect clinical interpretation
- querying for multiple transcript sources through a single interface

1.2 Examples

1.3 Getting Started

1.4 Installation

1.5 Database

1.5.1 Schema

1.5.2 Database Loading

1.6 Modules

1.6.1 uta.db.transcriptdb

Additional Resources

- [UTA Code Repository and Issue Tracker](#)

Indices and tables

- `genindex`
- `modindex`
- `search`