

# TYPELOADER 2: Automated full-length submission of novel alleles



Bianca Schöne<sup>1</sup>, Markus Fuhrmann<sup>1</sup>, Vineeth Surendranath<sup>1</sup>, Julia Phielers<sup>1</sup>, Alexander H Schmidt<sup>1,2</sup>, Vinzenz Lange<sup>1</sup>, Gerhard Schöfl<sup>1</sup>

<sup>1</sup>DKMS Life Science Lab, Dresden, Germany; <sup>2</sup>DKMS gGmbH, Tübingen, Germany

## Motivation

In order to use full-length sequencing for genotyping of HLA and KIR, it is essential that novel alleles are characterised and submitted in full length and that known alleles are extended to cover the complete length of the gene. However, the manual annotation and submission of full-length sequences to the European Nucleotide Archive (ENA) and the IPD-IMGT/HLA or IPD-KIR databases is a time-consuming and error-prone task. In 2016 we presented TypeLoader, a tool for automated full-length submission of novel HLA alleles. Here, we present an update, TypeLoader2, with new features and extended capabilities.

## TypeLoader (2016): semi-automated submission to ENA and IPD

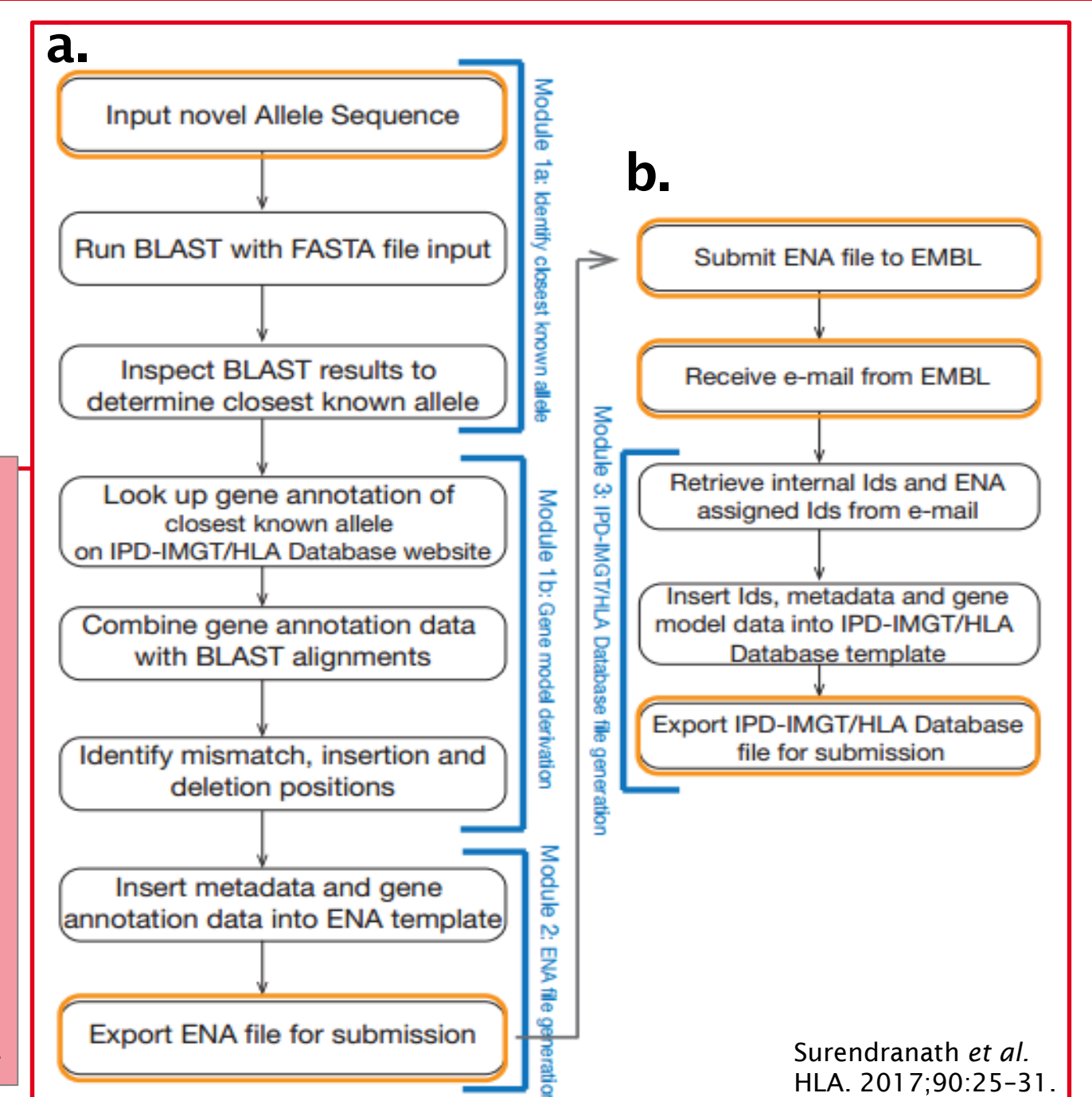
- Accepts full-length sequences of novel alleles → Automatically creates all files for submission
- Web application that needs a Linux-server backend

### Features:

- ✓ Support for all HLA genes
- ✓ Upload sequences as fasta or XML
- ✓ Generates all submission files for ENA and IPD-IMGT/HLA
- ✓ 95% less hands-on time than with manual submission
- ⇒ Submitted >1000 novel HLA alleles at DKMS Life Science Lab

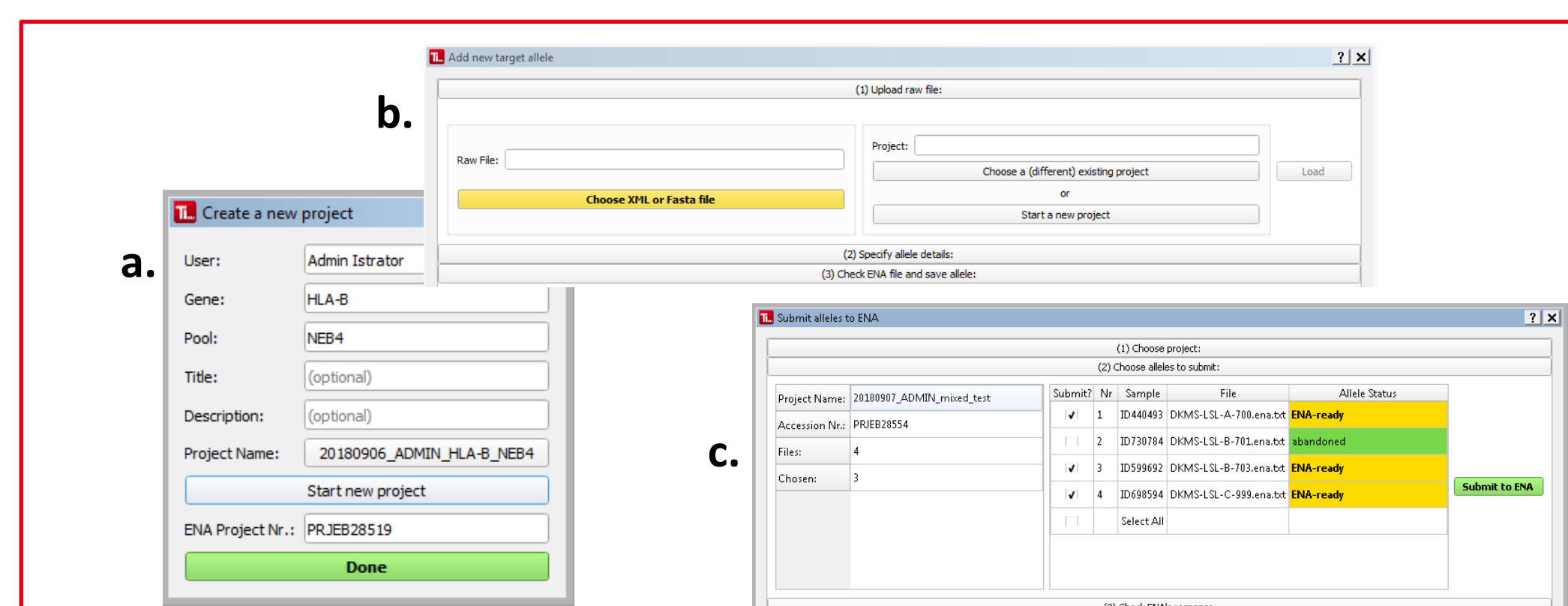
### TypeLoader workflow.

- ENA submission.** To annotate a novel allele sequence, the closest known allele is identified using BLAST against the reference HLA database (Module 1a) and the exon/intron information is transferred to the novel allele (Module 1b). A file ready for submission to ENA is created from the sequence data together with all required metadata (Module 2).
- IPD-IMGT/HLA submission.** A file ready for submission to IPD-IMGT/HLA is created from the sequence data together with all required metadata (Module 3). The sections marked in blue are steps that have been automated, while the steps marked in orange are steps that require manual intervention.



## Ti... TypeLoader2 (2018): convenient administration of all your novel allele data

- ✓ Standalone application: no server required
  - Just download, install, and use
- ✓ Integrated SQLite database: maintain all metadata of your novel alleles
  - Keep all files in a central, organised structure
  - Add, access, edit, or submit data directly

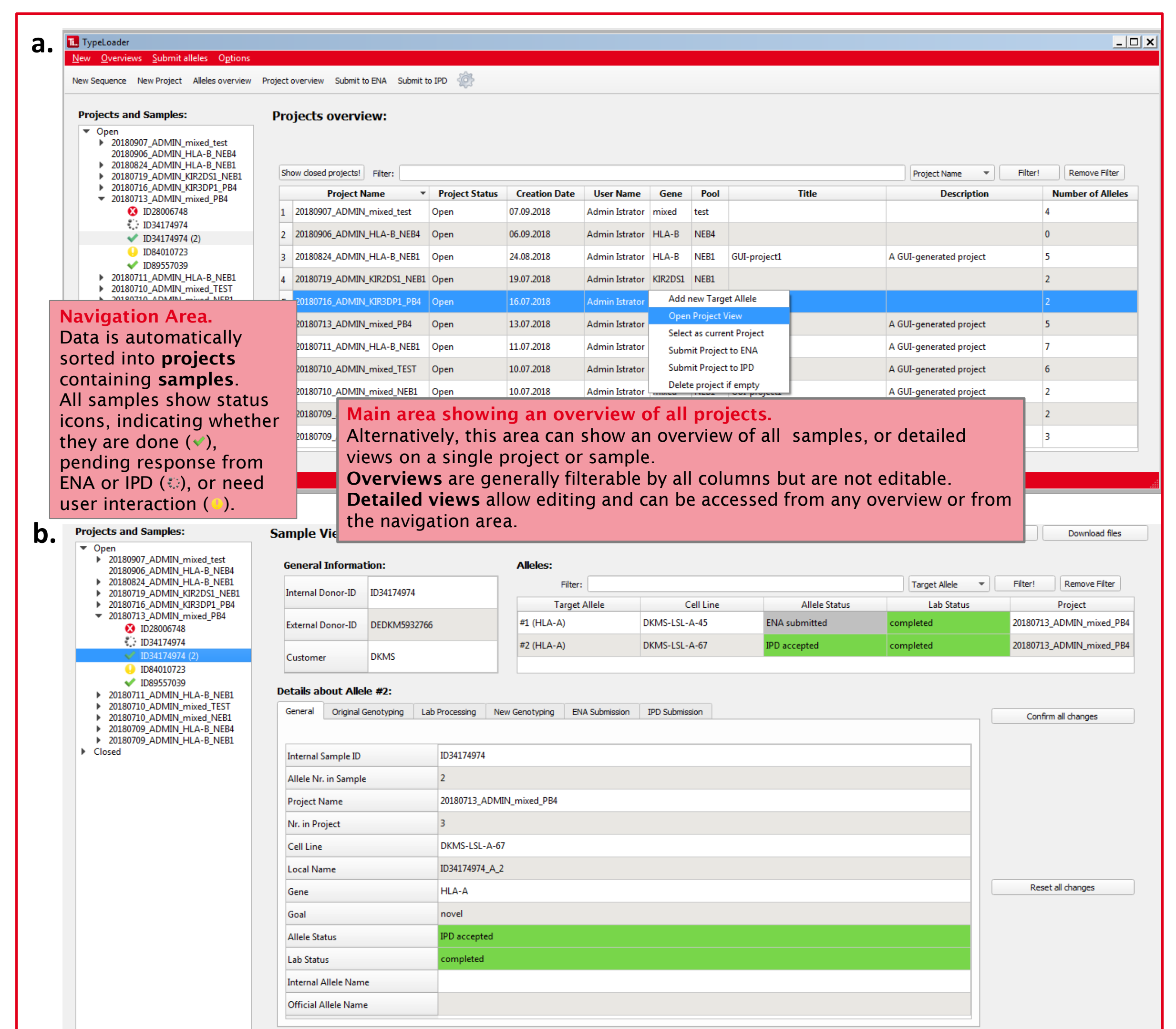


### Examples of TypeLoader user dialogs.

- New Project.** Generate a new project, including retrieval of an ENA ID.
- New Sequence.** Upload a new fasta or XML file. TypeLoader then automatically parses it, annotates the sequence, and generates an ENA submission file, which is stored.
- ENA Submission.** Choose any samples of a project to submit to ENA. Automatically updates each allele's status.

## Additional new features:

- ✓ Programmatic submission to ENA via a REST API
- ✓ Automatic detection of null alleles caused by frameshift mutations
- ✓ Experimental support for KIR alleles



### TypeLoader's Graphical User Interface (GUI).

- Complete TypeLoader-GUI displaying the Projects Overview.** In this overview, aggregated data about all projects is shown. Closed projects are hidden by default but can be shown optionally.
- Sample View.** This detailed view shows all data about the novel alleles of one sample and allows viewing, updating and editing of all metadata as well as downloading, viewing and editing of all files for this sample.

## Conclusion

TypeLoader2 can easily be used on standard PCs without dependencies on other software.

With the added internal database, **TypeLoader2 can be used as a lab's central information platform on their novel full-length sequences.**

We hope that the increased convenience and scope of TypeLoader2 will foster the submission of more full-length sequences to the IPD-IMGT/HLA and IPD-KIR databases, ultimately promoting the widespread use of full-length sequencing for genotyping of both HLA and KIR.



Download from GitHub:  
<https://github.com/DKMS-LSL/typeloader>

