

From Sample To Submission

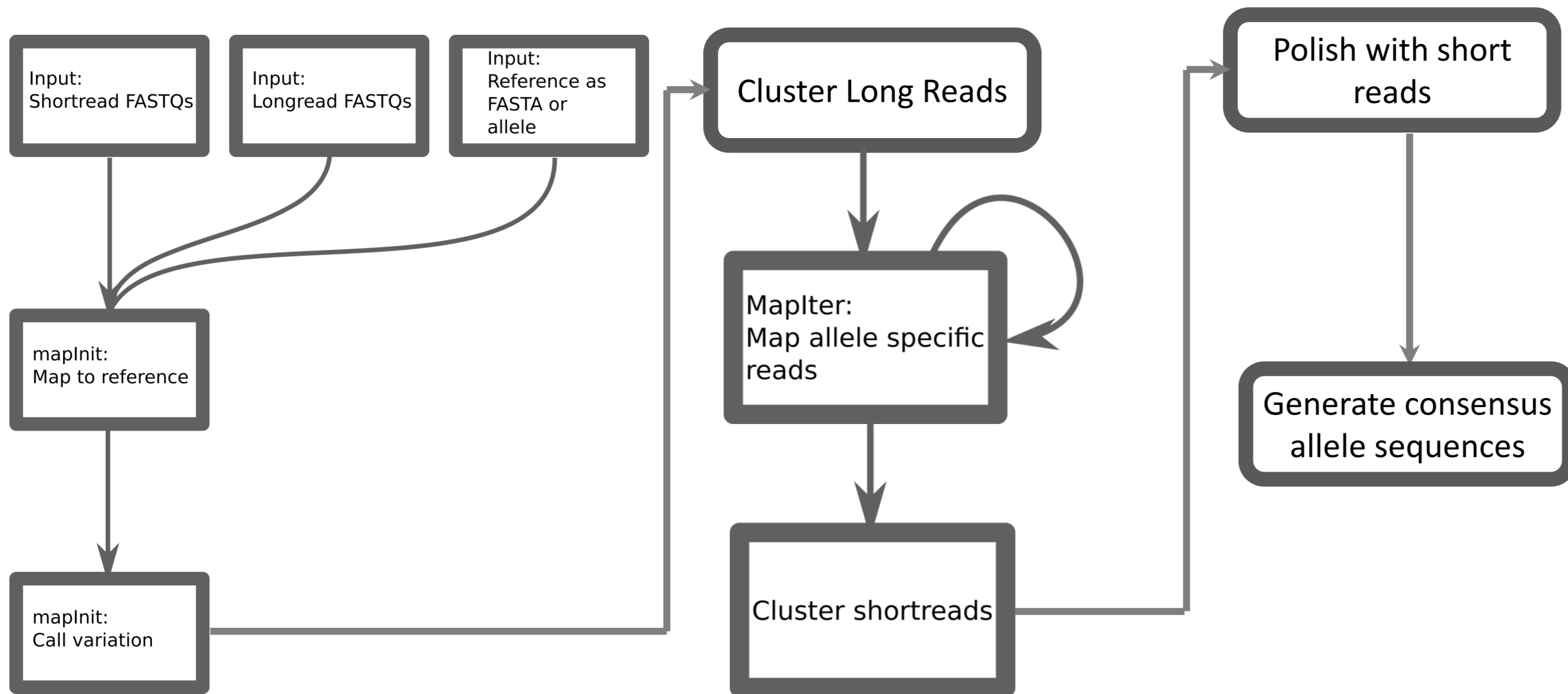
An Integrated Workflow For Characterising Novel HLA And KIR Alleles In Full Length

Vineeth Surendranath¹, Kathrin Putke¹, Marie Günther¹, Viviane Albrecht¹, Steffen Klasberg¹, Markus Fuhrmann¹, Bianca Schöne¹, Alexander H Schmidt^{1,2}, Anja Klussmeier¹, Gerhard Schöfl¹, Vinzenz Lange¹
¹DKMS Life Science Lab, St. Petersburger Str. 2, 01069 Dresden, Germany;
²DKMS, Kressbach 1, 72072 Tübingen, Germany

Introduction

The high throughput of genotyping samples at the DKMS Life Science Lab results in regular discovery of novel HLA and KIR alleles. We have established a dual redundant sequencing strategy that couples sequence data from complementary sequencing technology platforms. The phase delineation issue that plagues short reads is addressed by the corresponding long reads, and the sequencing errors issue that plagues long reads is addressed by the corresponding highly accurate short reads. Subsequent to the characterisation of a novel allele sequence using an algorithmic workflow developed at the DKMS Life Science Lab, the data is submitted to the allele databanks using an almost automatic software tool that has been developed in-house, ameliorating error-prone laborious manual curation efforts.

Dual Redundant Sequencing Strategy Algorithm

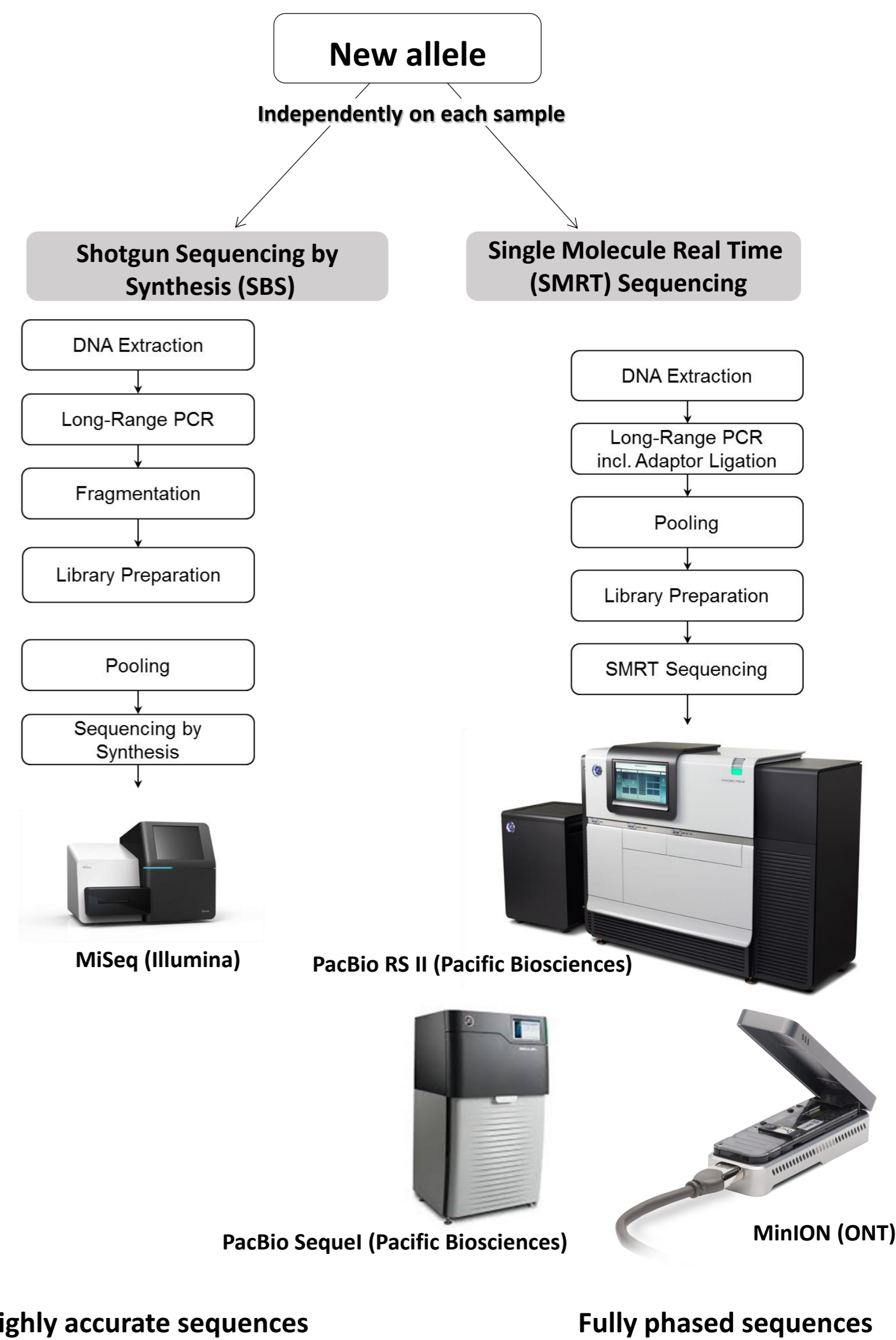


implemented in the R package DR2S available on GitHub



Download from GitHub:
<https://github.com/DKMS-LSL/dr2s>

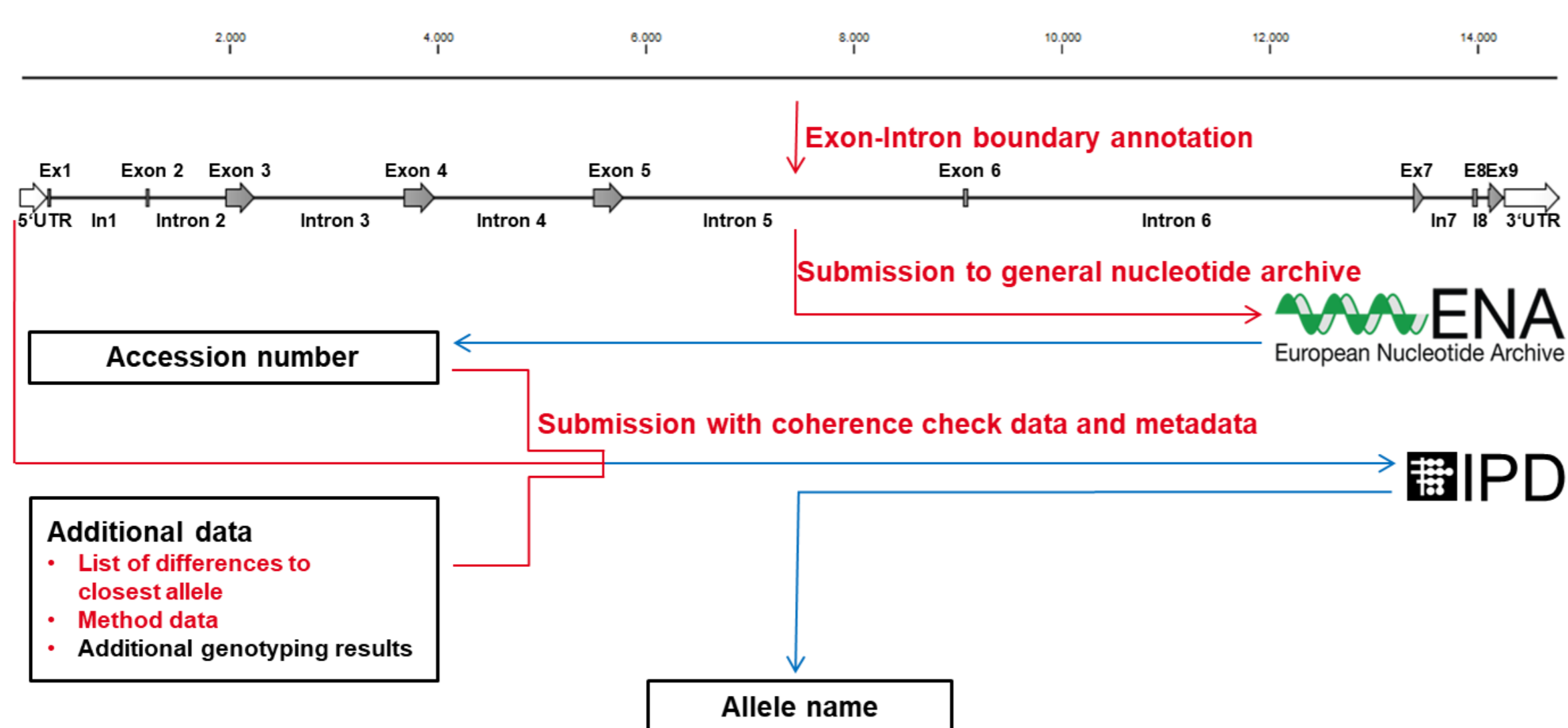
Dual Redundant Sequencing Strategy



Highly accurate sequences

Fully phased sequences

Novel Allele database submission workflow

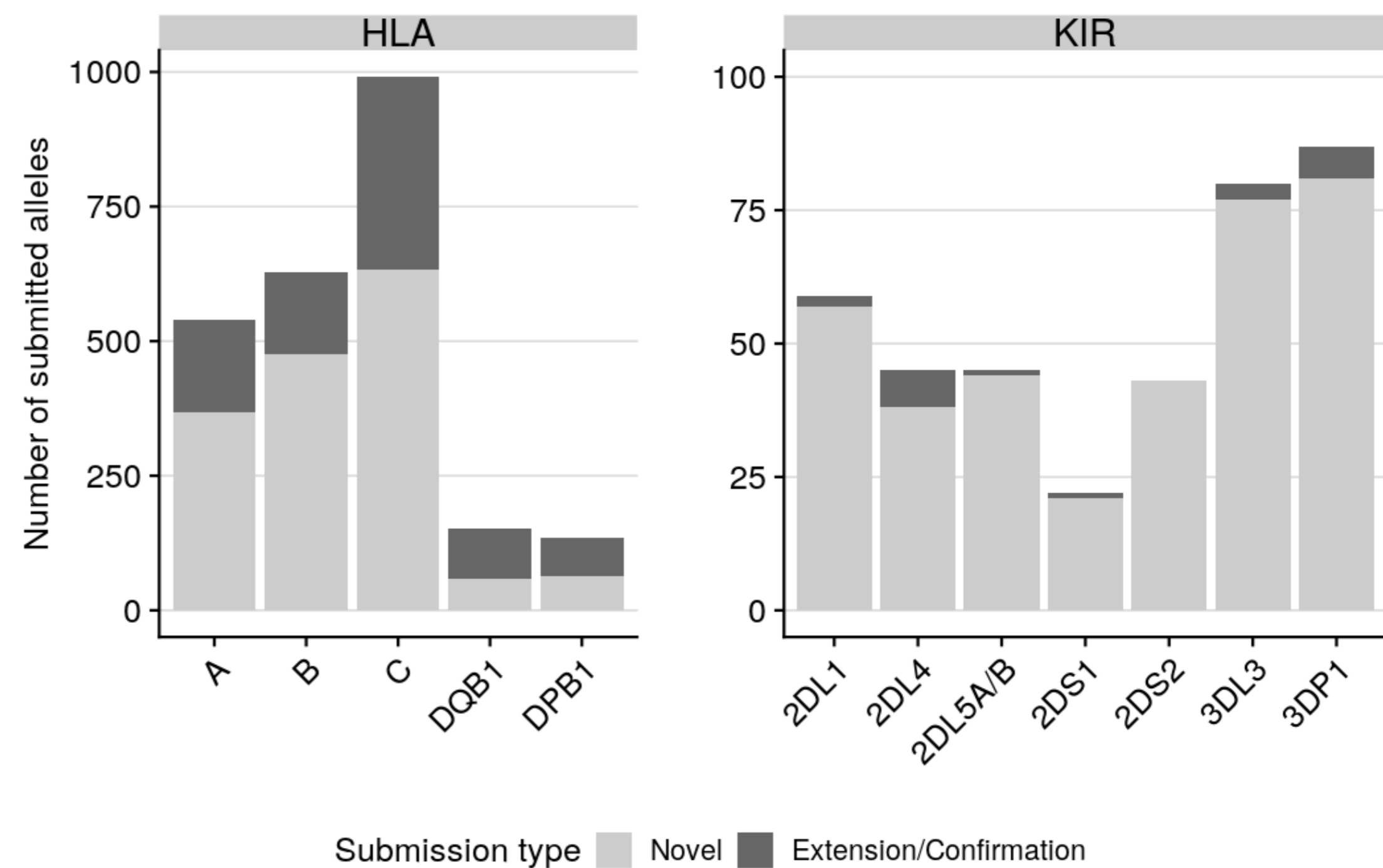


implemented as a standalone application TypeLoader available on GitHub



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

Allele submissions



Conclusion

Using the Dual Redundant Sequencing Strategy (DR2S) workflow at the DKMS Life Science Lab for both, partially known and novel alleles, we have characterised and submitted to the IPD-IMGT/HLA and IPD-KIR databases more than 1000 HLA and 500 KIR alleles in 2018 using the TypeLoader annotation and submission tool. This seamlessly integrated workflow represents a robust and efficient system to characterise alleles in full length, thus contributing significantly to improving data quality of the reference databases.

