TruSight® HLA Assign™ 2.1 Data Analysis Software

A powerful software package that assigns and reports HLA typing results quickly and confidently.

Highlights

- **Automated Analysis**
  Sequencing data is aligned, phased, and typed automatically for quick access to results
- **Confident Answers**
  Unbiased analysis produces high-confidence results regardless of novel alleles, rare alleles, or missing reference sequence
- **Rapid Decision Making**
  Optimized and comprehensive data visualization provides users with the tools to assign consistent and reliable types quickly

Introduction

Genetic variations within the human leukocyte antigen (HLA) system have been associated with transplant rejection, autoimmune disorders, cancer, and drug sensitivity.[1-4] Sequencing and analyzing this region of the genome is difficult due to high levels of sequence homology and dense variability.[5,6]

To address this particular challenge, the TruSight HLA v2 Sequencing Panel incorporates data analysis software developed by Conexio Genomics, a pioneer in HLA sequencing and bioinformatics. TruSight HLA Assign 2.1 Data Analysis Software, creates a comprehensive, DNA-to-report, HLA sequencing solution (Figure 1).

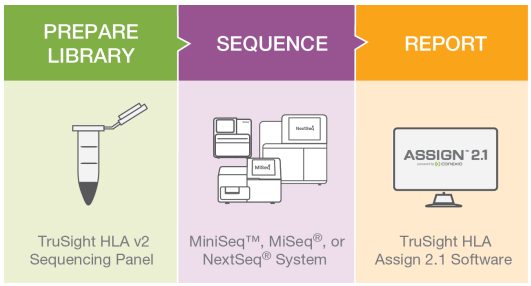

Figure 1: DNA-to-Report HLA Sequencing—This fully integrated HLA sequencing solution begins with preparing sequencing-ready libraries using the TruSight HLA v2 Sequencing Panel. Prepared libraries are then sequenced on an Illumina MiniSeq, MiSeq, or NextSeq System. Finally, TruSight HLA Assign 2.1 Software performs data analysis and reporting.

Automated Analysis

TruSight HLA Assign 2.1 Software is a sophisticated HLA bioinformatics platform designed for HLA data analysis and reporting (Figure 2). It automatically aligns millions of raw sequencing reads to a reference that includes all International ImMunoGeneTics Information System (IMGT)/HLA database alleles.[7] TruSight HLA Assign 2.1 Software then phases all heterozygote positions to one another to generate phased HLA alignments. Each allele sequence is then compared base by base to every allele in the IMGT/HLA database, which currently includes more than 15,000 unique alleles. This process produces high-confidence HLA typing results from 11 HLA loci (HLA-A, -B, -C, -DRB1/3/4/5, -DQB1, -DPB1, -DQA1, and -DPA1).

Alignment

TruSight HLA Assign 2.1 Software imports the direct output of MiniSeq™, MiSeq®, and NextSeq® 500 Systems in the form of .fastq.gz files, which contain the raw sequencing reads for each sample. These systems produce between 1 million and 400 million sequencing reads per run. The TruSight HLA v2 Sequencing Panel uses paired-end 150 base pair reads (2 × 150 bp), so each sequencing run generates 300 million to 120 billion individual base calls. These reads are aligned to consensus reference sequences for each locus.

Phasing

HLA genes are densely variable such that polymorphic positions and paired-end reads can be used to phase chromosomes to determine allele-specific variation. This is made possible by the complementary and integrated design of the TruSight HLA v2 Sequencing Panel and TruSight HLA Assign 2.1 Software. TruSight HLA sequencing selects for molecules of 500—1300 base pairs in length, and the paired reads are known to be in phase with one another. Heterozygote positions are rarely further apart than 1300 bp, allowing the algorithms to phase these positions and produce full gene sequences for each allele at each locus.

Typing

TruSight HLA Assign 2.1 Software includes the entire IMGT/HLA database of alleles and compares the allele sequences for each locus to every allele in the database. Perfect matches to alleles in the database are reported to 2, 3, or 4 fields of resolution.

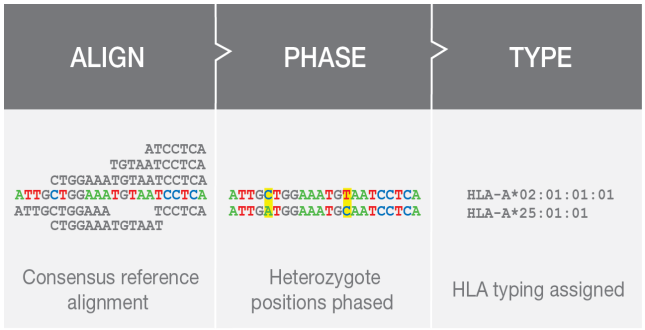

Figure 2: TruSight HLA Assign 2.1 Software Automated Workflow—TruSight HLA Assign 2.1 Software automatically performs alignment, phasing, and typing, the 3 key bioinformatics algorithms for HLA sequence interpretation.

Informative Data Visualization

TruSight HLA Assign 2.1 Software enables rapid, reliable, and traceable analysis and typing. This is made possible by the unique data visualization within the software interface. Rather than provide a list of numbers and quality control metrics, TruSight HLA Assign 2.1 Software displays all relevant information with visual cues in the Summary Screen (Figure 3). Clicking any result opens the Coverage View, which contains information about sequence quality, depth of sequencing, base-call frequency, noise, read diversity, phase, and reference (Figure 4). These cues allow for rapid decision making, and the inclusion of so many variables (Table 1) allows for high confidence in the HLA type that is called.

High-Confidence Results

TruSight HLA Assign 2.1 Software simplifies the review and reporting process. Only the data relevant to decision making is presented and only 3 confidence indicators are needed:

1. Results with low-quality sequencing, low depth of coverage, and/or allele imbalance are flagged with a red indicator (Figure 3, inset).
2. Rare alleles are easily identifiable and comparable to the base position at which the rare allele differs from the most closely matched common allele, allowing for rapid confirmation of rare alleles.
3. The lack of a typing result is an indicator of a novel allele, an ambiguity, contamination, or low-confidence base call, which can be quickly and easily determined.

Table 1: TruSight HLA Assign 2.1 Software Views and Data

Summary Screen	
Typing Results (11 loci, up to 4 fields)	
Minimum depth of coverage per locus	
Mean depth of coverage per locus	
Percent bases over Q30	
Coverage View	
Locus coverage map	Reads not contributing to consensus
Gene map	Read diversity
Phasing blocks	Read quality
Consensus reference sequence	Results Panel
Basecall confidence indicator	Closest matched allele pairs
Reference sequences	Core exons mismatches
Sample consensus sequence	Coding mismatches
Basecall frequency	Noncoding mismatches
Basecall depth of coverage	Phase mismatches
Phase alignment	IMGT/HLA sequence coverage
	Common, well-documented (CWD)
Alignment View	
Closest matched alleles pairs	
Sample alignment to reference	
Reads View	
Individual reads	
Read quality	
Read direction	
Read diversity	
Read alignment	
Reference View	
All IMGT/HLA allele sequences	
Reference filtering	
Alignment to sample consensus	

Rigorous Typing with Manual Review

Next-generation sequencing (NGS) increases coverage and resolution of the HLA region compared to conventional Sanger/capillary electrophoresis (CE) sequencing methods. [8,9] Given the highly polymorphic nature of this genomic region, novel alleles are being identified frequently with NGS in exons and in noncoding sequence.[10] Because novel alleles are frequently found in noncoding sequence, they are likely to occur in regions not covered by conventional methods, which makes proper sequence alignment of paramount importance.[9]

TruSight HLA Assign 2.1 Software requires a “perfect match” to an allele in the IMGT/HLA database to produce a typing result. This means that there must be 100% sequence identity between the sequencing result and the reference. Other HLA typing applications tolerate mismatches during alignment of reads to the reference sequence and report typing results for imperfect matches. Although this approach will boost concordance with conventional typing methods, it risks reporting results inaccurately. If the mismatch is due to a novel allele that occurs outside the reference sequence (determined by conventional methods), the HLA type that is called

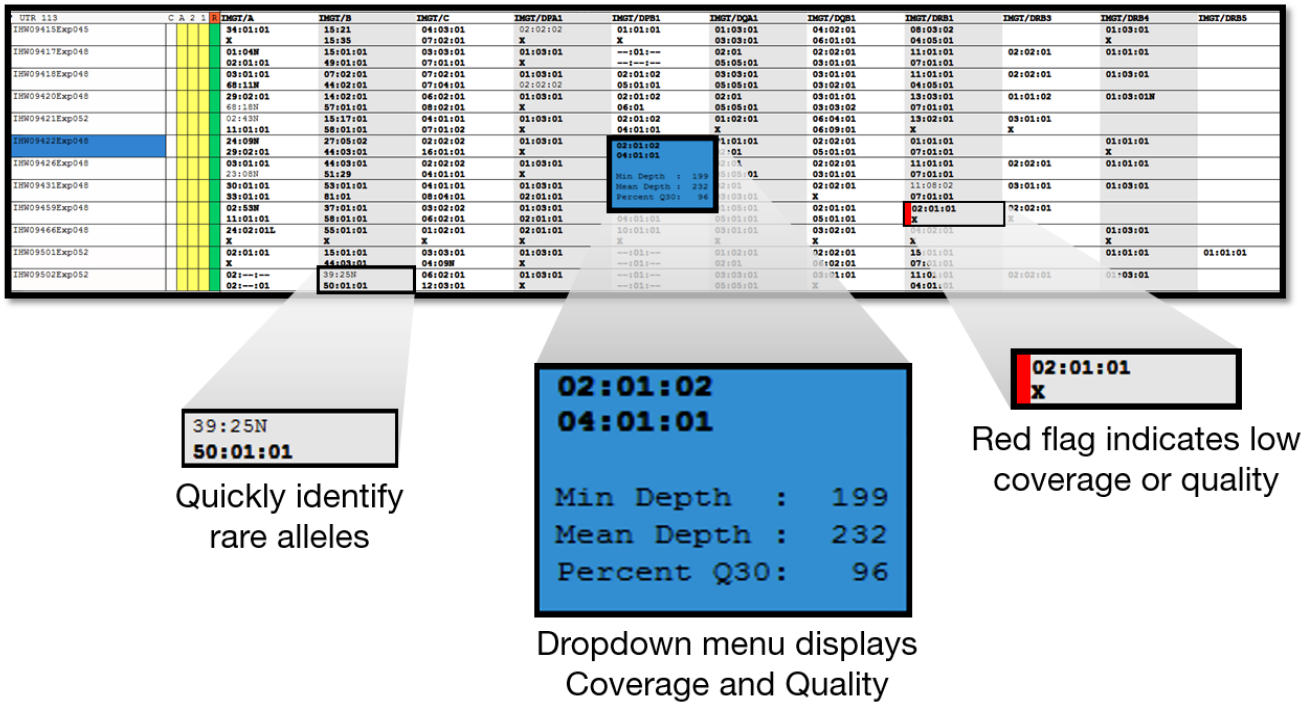

Figure 3: TruSight HLA Assign 2.1 Software Summary Screen—The TruSight HLA Assign 2.1 Software Summary Screen shows all the typing results for every sample and every allele to the number of fields of resolution specified by the user. Common, well-documented (CWD) alleles are highlighted in bold. Sequencing quality and depth are provided for every locus and flagged in the event they drop below thresholds.

may be concordant with the reference but inaccurate if that novel allele is a significant variant in terms of protein expression (eg, null alleles, expression variants).

Rapid Decision Making

The "perfect match" approach used by TruSight HLA Assign 2.1 Software requires analysts to investigate the novel variant and decide, with a wealth of data, whether they agree with the call. Analysts have 3 options:

- Leave the sequence as it is and call the mismatch a novel allele.
- Edit the sequence to match the reference.
- Make a "no call" at the position of the mismatch if the data is not clear. The "no call" option will then call an ambiguity between all alleles that share a variant at that position.

While sequence editing is simple with a single click, the decision that is made is traceable and auditable with multiple levels of review.

HLA Reporting and Workflow Management

TruSight HLA Assign 2.1 Software supports multiple users with unique login credentials. The software includes a flexible reporting engine that allows users to output text, Excel, XML, and FASTA formats. Reports can be configured to 2, 3, or 4 fields of resolution, G groups, P groups, or National Marrow Donor Program (NMDP) codes. Reports can also include the consensus sample sequence, user edits, mean depth of coverage per locus, user comments, and per locus percent of bases over Q30. All user edits, confirmations, and reports are tracked and auditable.

Summary

TruSight HLA Assign 2.1 Software is a powerful tool designed to assign and report HLA typing results quickly and confidently. It is streamlined with data visualization cues that present all relevant information in a single screen. It enables rapid, high-confidence decision making and reporting without the need for tedious calculations. TruSight HLA Assign 2.1 Software employs a "perfect match" approach to reporting, ensuring that decision making is evidence-based and traceable.

Learn More

To learn more about TruSight HLA Sequencing , visit www.illumina.com/hlaseq

Ordering Information

TruSight HLA Assign 2.1 Software is part of the TruSight HLA v2 Sequencing Panel. Licenses are sent via email after TruSight HLA v2 kits are ordered. Licenses support unlimited users, systems, and samples. Download TruSight HLA Assign 2.1 Software from the TruSight HLA v2 Support Page. The Support Page also includes the

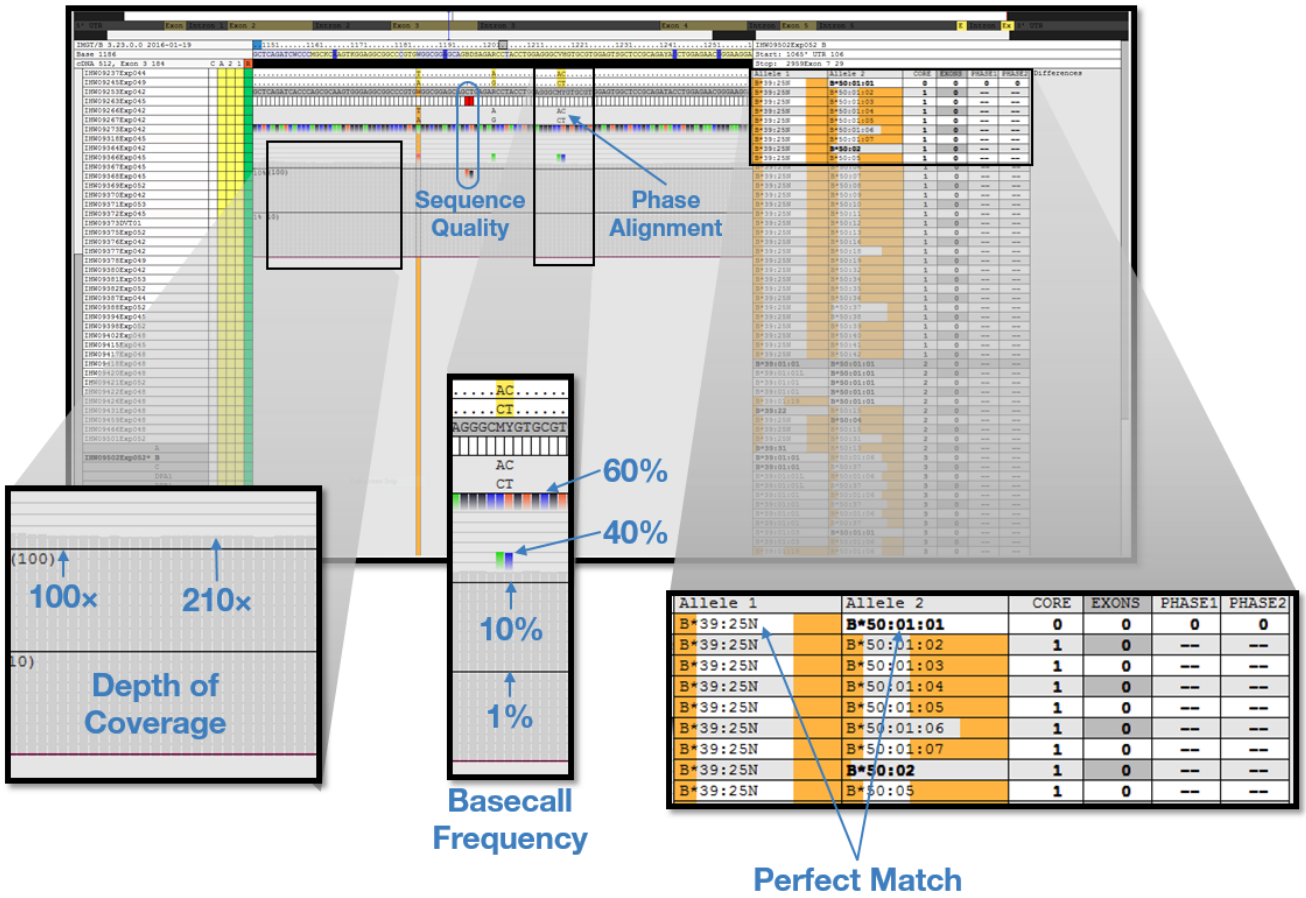

Figure 4: TruSight HLA Assign 2.1 Software Coverage View—The TruSight HLA Assign 2.1 Software Coverage View displays all the relevant details for each sequenced base. The visual cues automate calculations and allow users to make decisions rapidly and confirm results

TruSight HLA Assign 2.1 Software User Guide, sample datasets, training videos, and release notes. The software can be installed and demo datasets can be loaded without an active license.

References

1. Center for Comparative Genomics. HLA associated diseases and phenotypes (ccg.murdoch.edu.au/private/yurek_kulski/mhc/Table4.html). Accessed June 21 2016.
2. Ayala García MA, González Yebra B, López Flores AL, Guaní Guerra E The major histocompatibility complex in transplantation. *J Transplant*. 2012;2012:842141.
3. Anasetti C, Amos D, Beatty PG, et al. Effect of HLA compatibility on engraftment of bone marrow transplants inpatients with leukemia or lymphoma. *N Engl J Med*. 1989;320:197–204.
4. The International HIV Controllers Study. The major genetic determinants of HIV-1 control affect HLA Class I peptide presentation. *Science*. 2010;330: 1551-1557.
5. Lange V, Bohme I, Hofmann J, et al. Cost-efficient high-throughput HLA typing by MiSeq amplicon sequencing. *BMC Genomics*. 2014;15:63.
6. Wittiq M, Anmarkrud JA, Kassens JC, et al. Develpment of a high-resolution NGS-based HLA-typing and analysis pipeline. *Nucleic Acids Res*. 2015;23;43 (11):e70.
7. The International ImMunoGeneTics Information System (IMGT)/HLA database. www.ebi.ac.uk/ipd/imgt/hla. Accessed June 21 2016.
8. Lange V, Bohme I, Hofmann J, et al. Cost-efficient high-throughput HLA typing by MiSeq amplicon sequencing. *BMC Genomics*. 2014;15:63.
9. Ehrenberg PK, Geretz A, Baldwin KM, et al. High-throughput multiplex HLA genotyping by next-generation sequencing using multi-locus individual tagging. *BMC Genomics*. 2014;15:864.
10. HLA Sequencing of IHWG Reference Samples with TruSight HLA v2 Sequencing Panel. www.illumina.com/content/dam/illumina-marketing/documents/products/whitepapers/trusight-hla-white-paper-070-2016-012.pdf

Illumina, Inc. • 1.800.809.4566 toll-free (US) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

illumina®