Next generation sequencing of *JK* (*SLC14A1*) gene reveals higher frequency of variant alleles, novel allele-defining SNPs (allele reference fingerprints) and reassignment of a purported *JK Null* allele

Malik Altayar
PhD Student, Plymouth University

Topics

- Introduction
- Workflow
- *JK* NGS sequencing results

Blood Group Genotyping Applications

- **Foetal genotyping (HDFN)**
- **Testing multiply transfused patients (sickle cell disease)**
- **Solving difficult serological results**
- **Autoimmune patient**
- **Limited serological reagents (Dombrock)**

(High demand on genotyped blood units)

Genotyping needs to be **Accurate, high-throughput and relatively cheap** application (**if to be routinely used**)

High throughput Genotyping methods

- **The BLOODchip**

bloodgen
blood grouping & genotyping

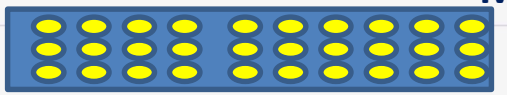
- **The human erythrocyte antigen (HEA BeadChip ™)** platform
- Based on a **DNA-array analysis**

Insufficient to define and **discover unknown**, **emerging mutations, alleles** in blood group systems.

Multiplex PCR Amplicons + Dye

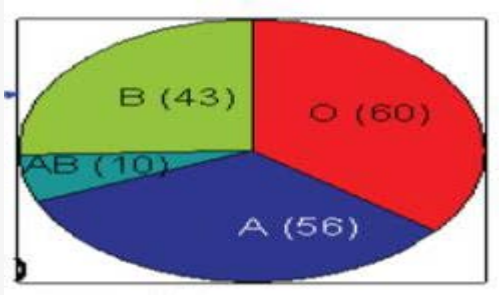
Hybridisation

Microarray of oligonucleotides



Increase in the number of blood groups alleles

- No. of alleles ABO
- 2001 (72)
- 2007 (215)
- 2013 (320)



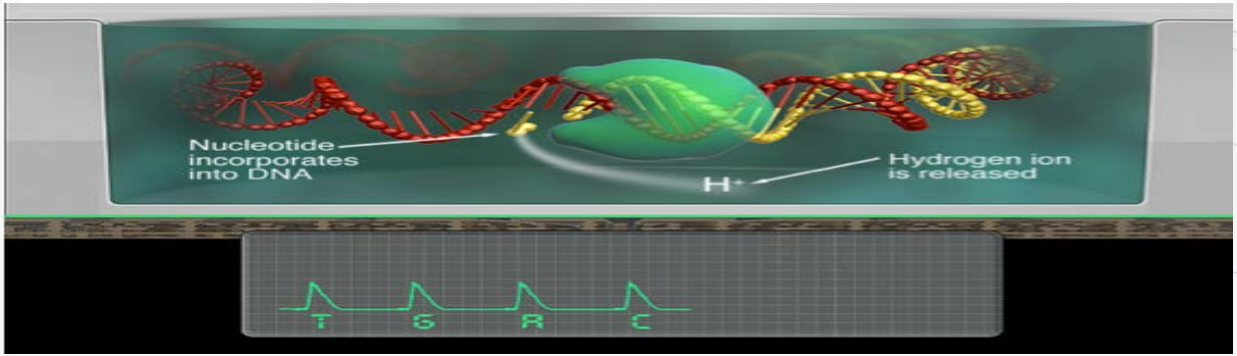
Next Generation Sequencing (NGS)

- Enables **high throughput** sequencing **(Massively Parallel DNA sequencing**).
- i.e. capable of sequencing various regions of interest in a significant number of samples in one run.
- **Fast**
- **Cost effective**

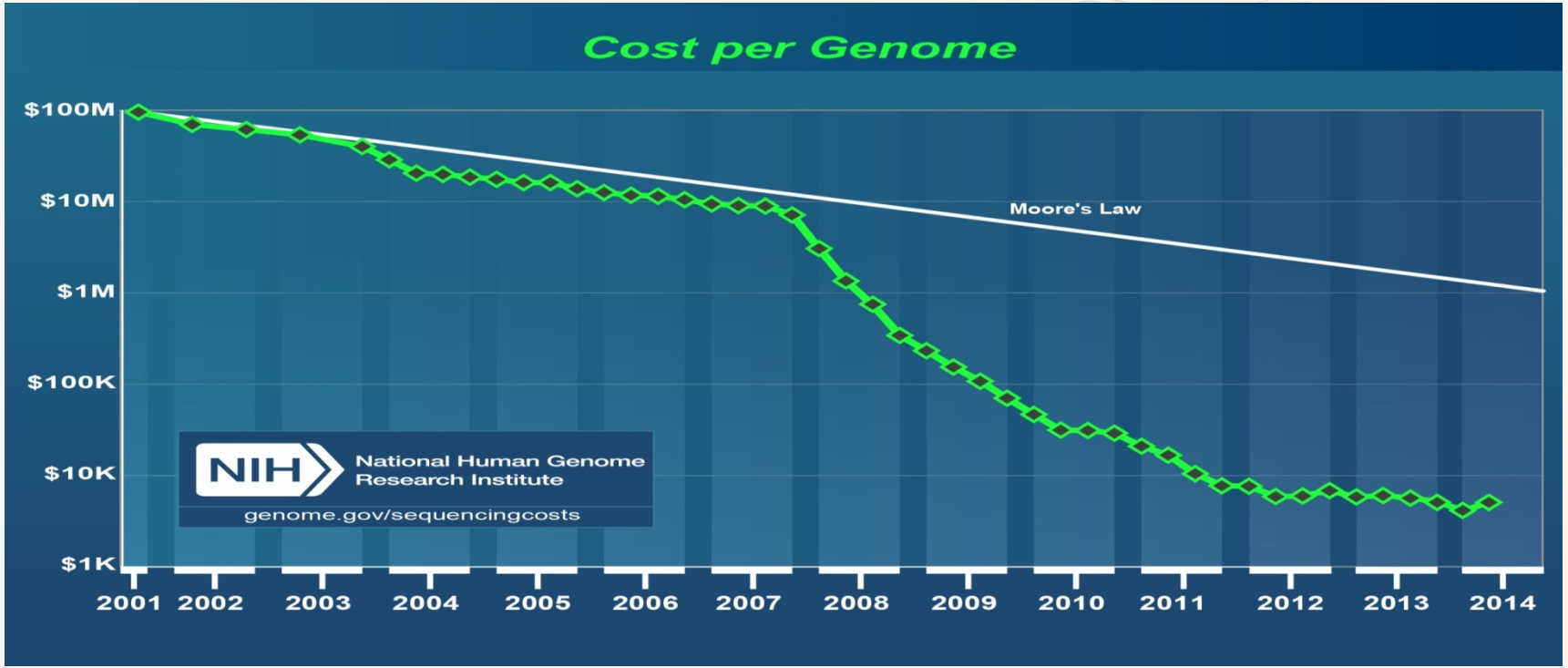
Ion Torrent Personal Genome Machine (PGM™)

Principle:

- **Sequencing by synthesis**, **Nucleotide incorporation** into DNA releasing Proton.
- Detect change in **pH**



Costs


Cost per Genome
$100M
$10M
$1M
$100K
$10K
$1K
Moore's Law
NIH
National Human Genome Research Institute
genome.gov/sequencingcosts
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

Project

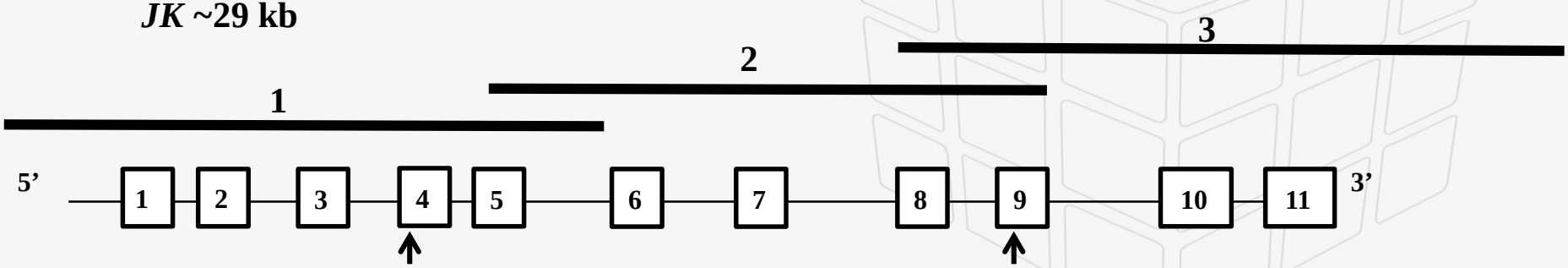
Extensive NGS-based genotyping of blood group genes:

- *ABO*
- Duffy (*FY, ACKR1*)
- **Kidd (*JK, SLC14A1*). 36 *JK* alleles**
- **Sequence the entire gene (exons and introns)** plus flanking regions **(regulatory regions)**

Long-Range PCR

3 Long Amplicons (~11-14kb)

JK ~29 kb



The workflow

Construct Library

↓

Prepare Template

↓

Run Sequence

↓

Data analysis

Data Analysis (Variant analysis)

Several **bioinformatics software packages** and websites are used such as:

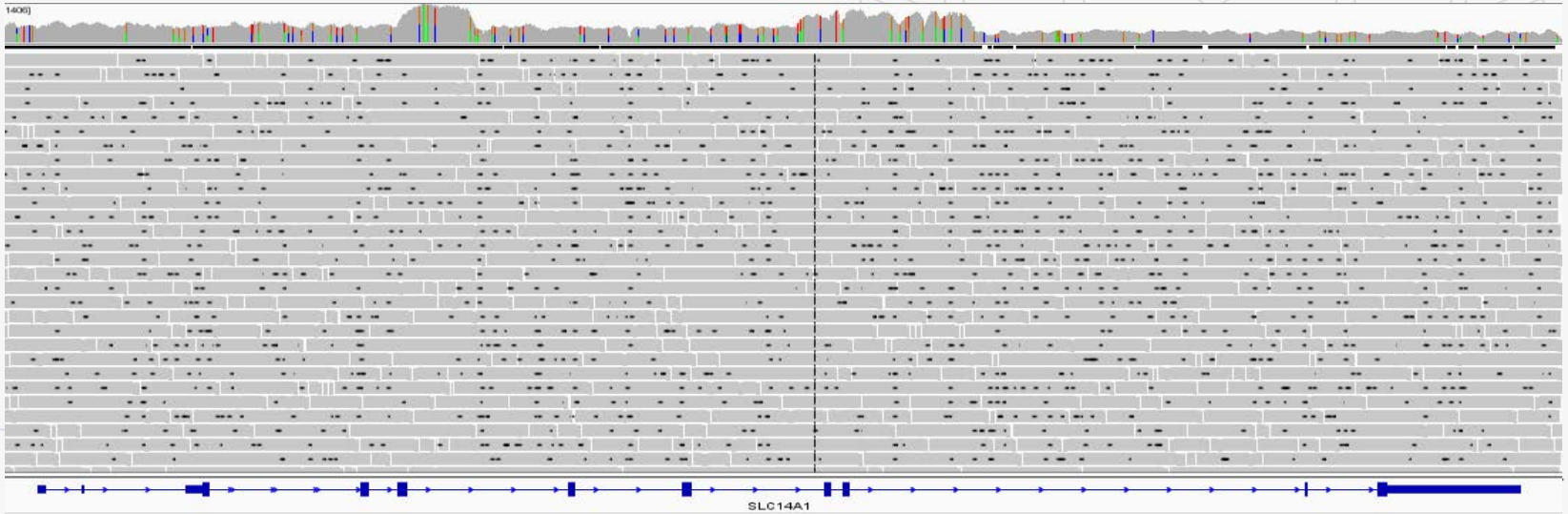
- **Ion Torrent Suite ™** plugins (such as coverage analysis, VariantCaller, Alignment and FastQC)
- Ion Reporter ™
- Integrative Genomics Viewer (**IGV**)
- Seattle annotation web site.
- Database, such as **NCBI**.
- Some cases of SNPs were confirmed by further analysis, e.g. Cloning and **Sanger sequencing.**

JK

- **67** samples
- **3.5 million** reads
- Coverage depth of more than **700x.**
- Data quality more than 99% accurate
- Significant number of **intronic** SNPs (80) were found.

NGS sequence

- Complete coverage across the gene
- All existing polymorphisms in various parts are visible



Findings

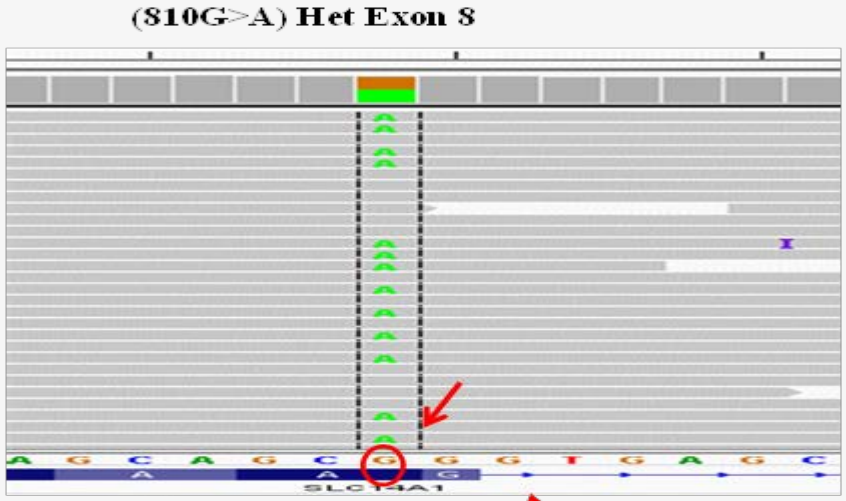
- **10/67** samples carry the **SNP 130G>A (assigned to allele *JK*01W.01* encoding weak Jka), one homozygous**, all with normal antigenicity.
- **Reassignment of purported *JK*B Null* allele (SNP 810G>A)**, was found in **10/67** samples, all with **normal serological Jkb** expression.
- **Intronic polymorphisms analysis:**

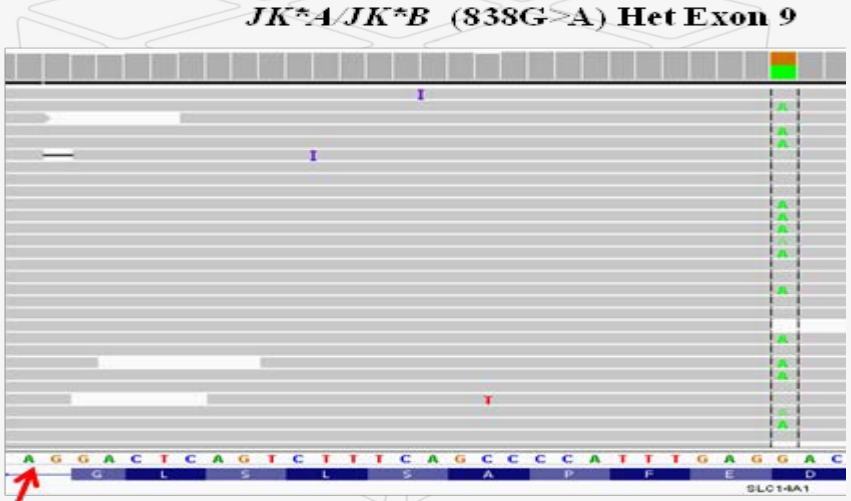
1. **Unique *JK* alleles' fingerprints (suggested allele reference sequence).**
2. ***JK*01W.01* allele sequence resembles *JK*A/JK*B* hybrid**
3. **SNP close to *JK*A/JK*B* critical polymorphism 838G>A, may lead to allelic dropout during BGG**

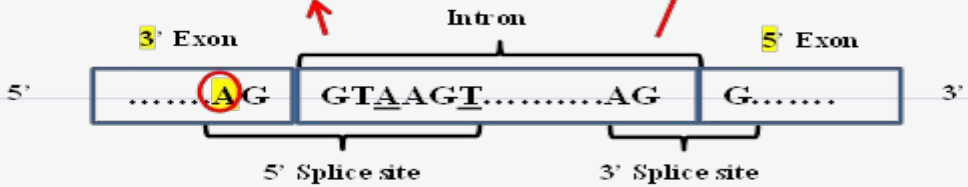
*Novel JK*B 810G>A Null* allele?

- **810G>A SNP** encodes for synonymous amino acid substitution Ala270Ala.
- located at the exon 8/intron 8 boundary (**the second last nucleotide of exon 8**)
- Due to its location, it was suggested to alter the expression of the Jk^b antigen in 2 Jk(a+**b-**) samples > disturbing the splice site > novel *JK*B* allele. (Henny et al., 2014)

*Novel JK*B 810G>A Null* allele?







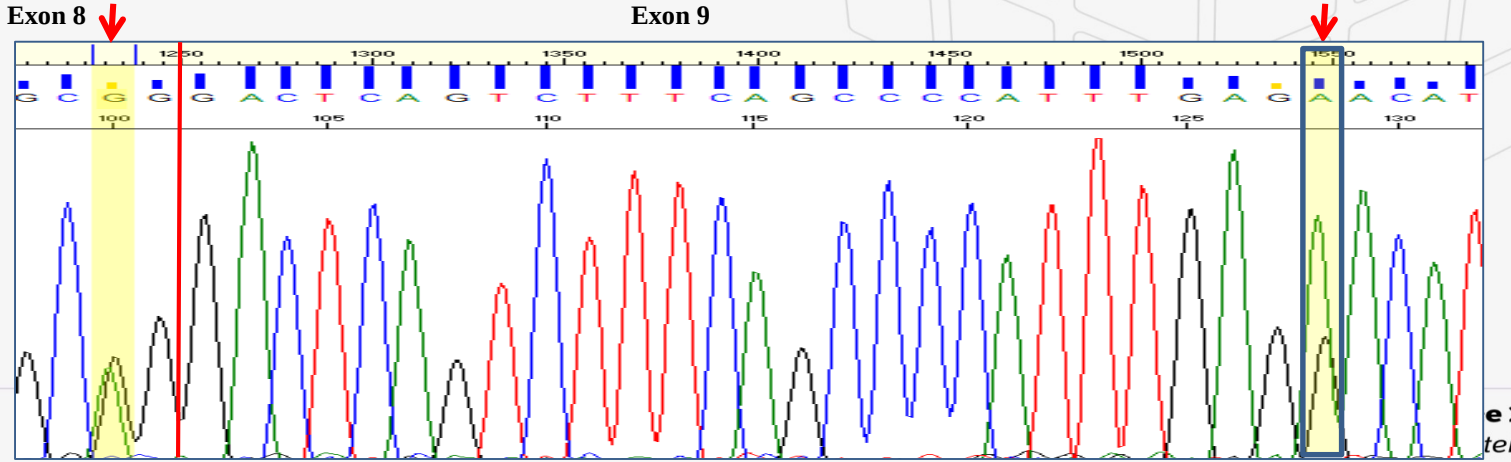
No effect on (Phenotype nor Genotype)?

- **cDNA analysis**
- **Phenotype**: All (10 samples) with this 810G>A SNP were (Jka-**b+**) or (Jka+**b+**)
- **Genotype:** no silencing effect **(no exon skipping)**

Sample JK009.55 (Jk a+b+, *JK*A/JK*B*)

(810G>A) Het Exon 8

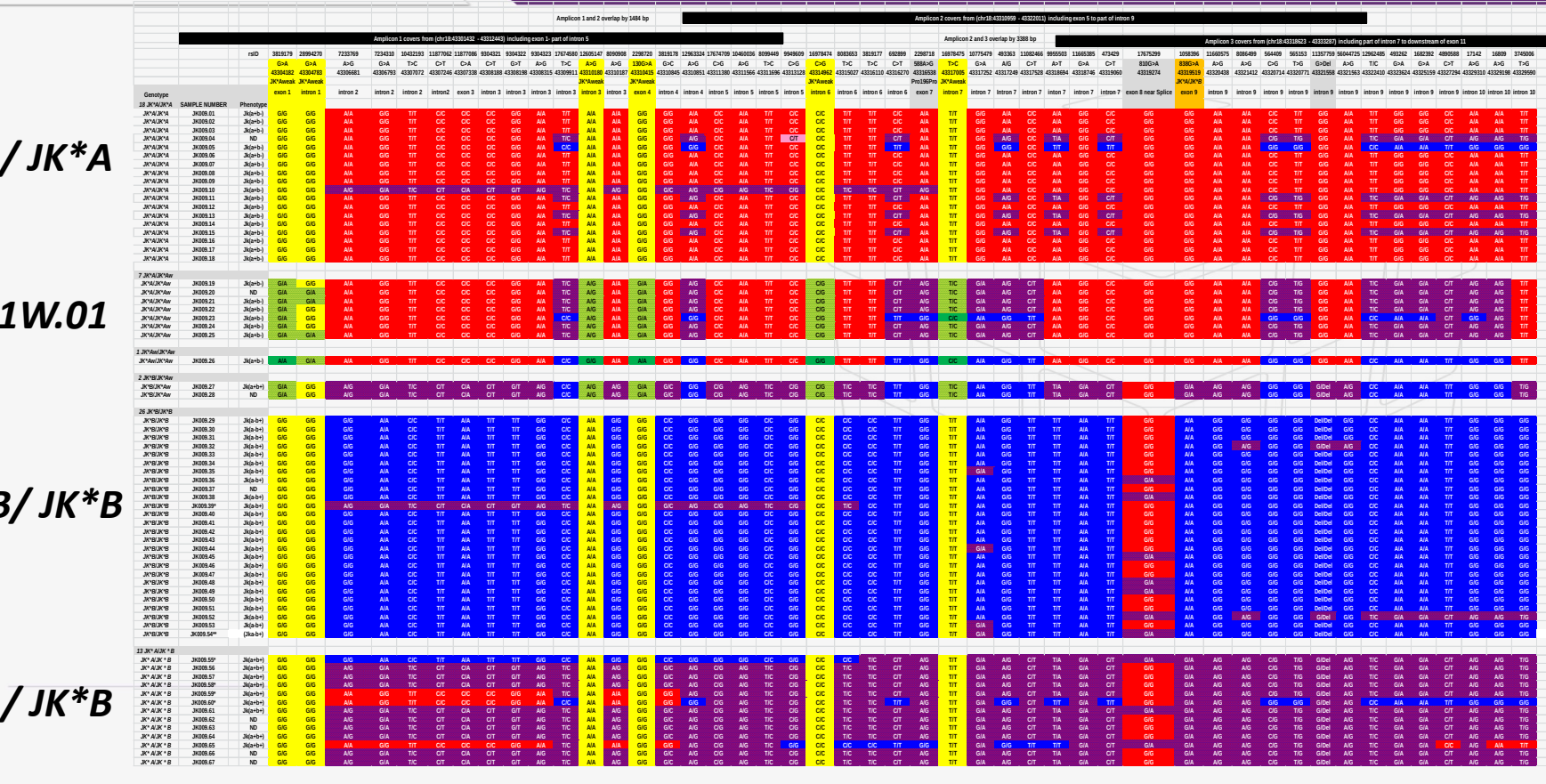
JK*A/JK*B **(838G>A) Het Exon 9**

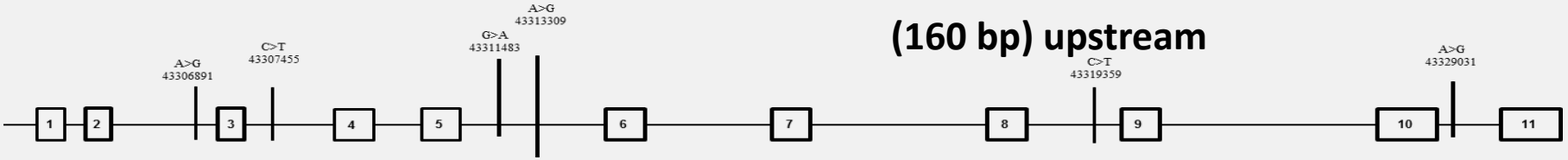


Allele-specific intronic polymorphisms ('fingerprints')

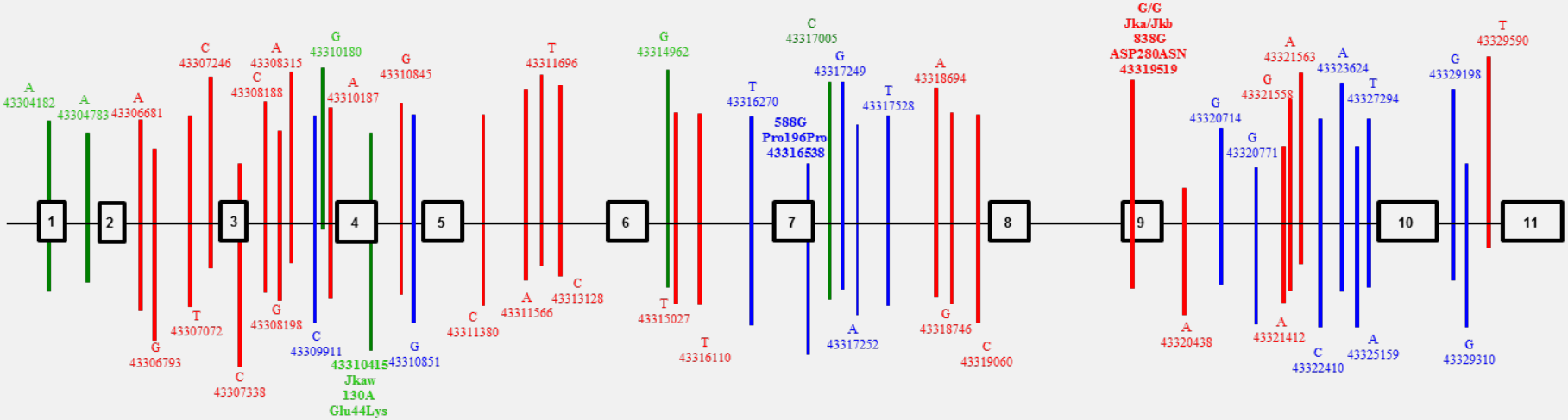
- The **correlation** of intronic polymorphisms with *JK* alleles (*JK*A*, *JK*B* and *JK*01W.01)* was assessed in **homozygous allele samples**.
- *JK* allele ***(JK*A, JK*B* and *JK*01W.01)*** specific patterns were found (**suggested reference sequences**), those samples **differing from these patterns represent new alleles**.
- *JK*01W.01* allele sequence *(*hybrid *JK*A/JK*B).*

BBTS Annual Conference 2016





D) *JK*Aw* (*JK*01W*)



- **130G>A** might not be the only factor for weakening the Jka expression. (Cumulative effect?)

Other findings

- SNP 588A>G and -46 G from intron 9 SNPs described to be associated with *JK*B* and *JK*01W.01,* however, have also been found in ***JK*A* allele samples.**

Key points

- **NGS** allows simultaneous **comprehensive sequencing of a large number of samples** for various **blood group genes**.
- **Discovery** mode, **novel and rare alleles**.
- **Accurate phenotype predictions**. Unlike microarray platforms.
- Exploring polymorphisms across the gene (exons, introns and flanking regions) provides comprehensive genotyping.

Key points

- Also, helps in studies of the **evolution of alleles**
- **Polymorphisms frequency. (high throughput)**
- Allows the **discovery of the causative factors** in **Discrepant and unusual** phenotype **samples due to novel or rare weak or null allele** (for example, explore **splice sites and regulatory regions**).
- **Cataloguing** polymorphisms **close to critical SNPs** to be taken into consideration while **designing genotyping primers.**
- These information helpful to develop genotyping software (allele-specific patterns).

- This NGS approach has been also applied to other blood group systems such as ABO, RH, KEL and FY.
- Near future, **NGS** will become the **potential methodology of choice for genotyping patients and donors**

Acknowledgements

- Professor Neil Avent
- Dr. Tracey Madgett
- Plymouth University research team.
- Tabuk University, Saudi Arabia

Thank You