Abstract

Structurally complex loci underlie many diseases. These loci can be very challenging to resolve by currently available methods such as karyotyping, clinical array, PCR-based tests, and next-generation sequencing. Next-generation mapping (NGM) by BioNano Genomics Irys® System offers a high-throughput, genome-wide method able to interrogate genome structural differences in the range of two kilobase pairs to hundreds of kilobase pairs. The Irys System uses extremely long reads to span interspersed and even long tandem repeats making it ideally suitable for interrogating genome structural differences in the range of two kilobase pairs to hundreds of kilobase pairs. The Irys System results in different lengths of the resultant Lp(a) protein, and there is a direct correlation between the size of the protein and risk of coronary heart disease, cerebrovascular disease, atherosclerosis, thrombosis, and stroke. Another important variable length tandem repeat is D4Z4, associated with facioscapulohumeral muscular dystrophy (FSHD). FSHD muscular dystrophy is strongly associated with a low copy number (< 10 units), occurring in 95% of FSHD cases. Copy number of tandem repeats is extremely hard to measure accurately with available methods, but we show that NGM on the Irys System can accurately measure the copy number of the kringle IV domain and D4Z4. The second class of complex structural variation is those that involve genes with paralogs such as amylase and UGT2B17, two genes whose copy number have been shown to be involved in human health (testosterone and estradiol metabolism, osteopathic health and graft versus host disease). We show deletions of UGT2B17 in a family trio and > 10 different structures at the Amylase region. The third class of genomic variation which is very difficult to interrogate are those flanked by segmental duplications. These are especially important because spontaneous rearrangements are common between paralogous segmental duplications causing copy number aberrations and translocation, thus resulting in developmental disorders, such as the 22q11.2 deletion syndrome mediated by segmental duplication rearrangements. We show the assembly of the region, including the normal and pathogenic alleles, using molecules that span and disambiguate the structure of the segmental duplications. We demonstrate that NGM using the Irys System is proving to be a highly accurate method for detection of clinically relevant structural variation.

Methods

1. Isolate High Molecular Weight DNA Molecules
2. Label DNA Molecules at Motif-Specific Locations
3. Linearize DNA Molecules in NanoChannels
4. Automatically Image Linearized DNA Molecules
5. Map DNA Molecules at Motif-Specific Locations

Gene Deletion in Mother and Son Genomes from the Ashkenazi Trio

Amylase Structural Variants

Conclusions

Next generation mapping can be used for the detection of clinically relevant genomic features otherwise difficult to detect including:

- Insertions/CNV (Amylase variants)
- SVs involving paralogous regions and segmental duplications (UGT2B17)
- Tandem repeat arrays (D4Z4)
- Balanced events (BCR/ABL and Amylase)

References