Omics Technologies

Reliable genotyping of APOE-ε4 and other markers in GC-rich regions with a new Axiom Workflow

Monica Chadha, Anu Mittal, Joseph McGirr, Neeraja Vegesna, Carsten Bruckner; Thermo Fisher Scientific, Santa Clara, CA 95051, USA

Abstract

- **Purpose:** Genotyping variants in GC-rich genomic regions remain a challenge for microarrays and short-read sequencing technologies. APOE-ε4 (rs429358), a key genetic risk factor variant for late-onset Alzheimer's disease, has been particularly difficult to genotype due to its location within GC-rich sequences. We have developed a modified whole-genome amplification Axiom assay that not only improves the genotyping performance of APOE but also enhances the detection of other biologically and functionally important variants, especially those located in GC-rich promoters, CpG islands, pharmacogenes, and the MHC- these regions remain technically difficult to genotype with high confidence.

- **Methods**: The study involved amplifying and precipitating gDNA from various sample types using the Axiom™ SwiftArray™ Assay, followed by hybridization and measurement of Axiom™ PangenomePro targets using the SwiftArrayStudio™ Microarray Analyzer. As a representative benchmark, over one hundred SNPs with local GC content ≥ 70% from the 1000 Genomes high coverage call set were evaluated.

- **Results:** The modified workflow achieved a call rate above 99.5%, an average cluster resolution improvement of 1.8 units, and 100% concordance with the 1000 Genomes truth set. Notably, GC-rich loci—including APOE-ε4 (rs429358) and other functionally important variants across key genes such as CYP2A6 (rs1801272, linked to altered nicotine metabolism), KCNQ1 (rs2237897, implicated in type 2 diabetes), and HLA-DPB1 (rs1042151, associated with immune response and aspirin-induced asthma)—showed tighter genotype clusters, improving by 1–4 units compared to the original Axiom protocol. Overall, the modified assay delivered a genome-wide average cluster resolution of ~7.8 units, providing clear evidence that it consistently resolves variants in high-impact genes.

Introduction

- The new SwiftArrayStudio Microarray Solution includes the Applied Biosystems™ Axiom™ SwiftArray™ Assay, the Applied Biosystems™ Axiom™ SwiftArrayStudio™ Microarray Analyzer, and the Automated Axiom Analysis Suite as shown in Figure 1.

- This advanced solution offers a two-day workflow from sample to insight, featuring improved whole genome amplification that ensures reliable performance for genotyping variants in GC-rich genomic regions while delivering results under 30 hours.

- The new assay couples (i) a modified isothermal whole-genome amplification chemistry with enhanced strand-displacement kinetics, (ii) modified enzyme recipe that protects enzyme inhibition and depolymerization via phosphorolysis and (iii) streamlined post-amplification cleanup. The workflow eliminates several Axiom workflow and wash steps, reducing hands-on time to under 3 hours while preserving array throughput. For enhanced sustainability, the solution uses fewer pipette tips with a universal tray pour-and-go mechanism, and the hands-on time is reduced to just 3 hours.

- The new workflow exhibits high concordance with the Axiom 2.0 workflow for genotype, copy number call rate, copy number gain/loss, and PGx + other module concordance. Crucially, the new chemistry maintains good representation across the full 20–85 % GC spectrum, yielding tighter genotype clusters.

Materials and Methods

- **Microarray data:** The 96 samples (Blood, Buccal, Saliva, Cell Line) were run on Axiom™ PangenomePro array using Axiom™ SwiftArray™ Assay, Axiom™ SwiftArrayStudio™ Microarray Analyzer and workflow (Figure 1). Sample preparation was done with Axiom™ SwiftArray™ Assay workflow and the current Axiom 2.0 Workflow. All in process QCs met the acceptance criteria. The DAT and CEL files were generated, and Axiom Best Practices Workflow was run to generate genotyping calls. Array genotyping calls were directly compared to 1000 Genomes Phase 3 reference calls to calculate concordance. For matched Blood, Buccal and Saliva samples, genotype call reproducibility was assessed at high GC-rich loci.

- **Performance evaluation:** The performance of genotyping calls from microarray data was assessed for specific markers in GC-rich regions by analyzing the cluster patterns in a two-dimensional space, based on signal intensity and contrast from two channels. Probesets that perform well for a given marker typically exhibit distinct and well-separated clusters. (Figure 2) For rare markers where only a single homozygote cluster is observed, additional samples were examined to confirm that the probeset can make accurate heterozygote calls.

Figure 1. The new Applied Biosystems™ SwiftArrayStudio™ Solution includes:
(1a) Applied Biosystems™ Axiom™ SwiftArray™ Target Preparation method
(1b) Applied Biosystems™ Axiom™ SwiftArrayStudio™ Microarray Analyzer for automated array processing
(1c) Illustrates the advancements of the Axiom™ SwiftArray Assay compared with the Axiom 2.0 Assay.

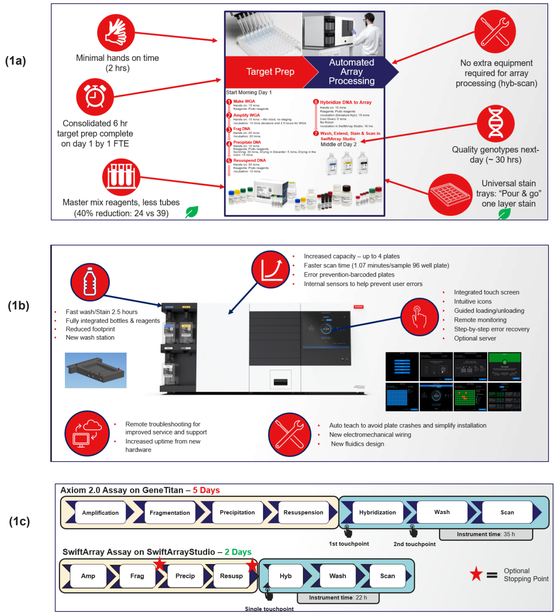
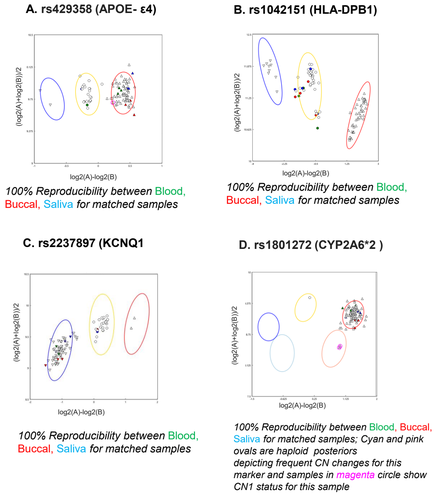


Figure 2: Genotyping cluster resolution is measured by a metric called, Fisher's Linear Discriminant (FLD). FLD metric is essentially the mean distance in signal contrast between a heterozygous cluster and the nearest homozygous cluster divided by the standard deviation across clusters. FLD is a direct measure of the cluster quality of a SNP. High-quality SNP clusters have well-separated centers (with respect to other clusters) and little variance about the center of the cluster (i.e., the clusters are narrow).

Figure 3 : Axiom™ PangenomePro provides good genotyping accuracy for markers in GC-rich genes across multiple sample types. We selected a few examples to show the genotype clustering resolution of APOE-ε4 (rs429358) and other functionally important variants across key loci, such as CYP2A6 (rs1801272), linked to altered nicotine metabolism; KCNQ1 (rs2237897), implicated in type 2 diabetes; and HLA-DPB1 (rs1042151), where next-gen sequencing often fails to deliver consistent calls.

For enhanced comparison, four matched Blood, Buccal, Saliva samples are color coded (green, red, blue) to illustrate genotyping performance across sample types. Non-colored samples are the cell line HapMap samples. Each point represents a sample, with x and y coordinates solely used for visual mapping.
For the Axiom array, genotyping calls are determined based on signal intensity and contrast derived from A and B channels, distinguishing between the two alleles. Ovals indicate prior probabilities used in genotype calling. Well-segregated homozygote and heterozygote clusters indicate that the probeset is performing effectively.



100% Reproducibility between Blood, Buccal, Saliva *for matched samples*

100% Reproducibility between Blood, Buccal, Saliva *for matched samples*

100% Reproducibility between Blood, Buccal, Saliva *for matched samples*

100% Reproducibility between Blood, Buccal, Saliva *for matched samples; Cyan and pink ovals are haploid posteriors depicting frequent CN changes for this marker and samples in magenta circle show CN1 status for this sample*

Figure 4: Comparison of the Axiom™ Swift Assay workflow on the SwiftArrayStudio™ Microarray Analyzer and the Axiom 2.0 workflow on the GeneTitan™ MC for probesets with GC content ≥ 70%. High GC content probesets within pathogenic and pharmacogenetic modules on the Axiom™ PangenomePro array are shown (n = 256). The y-axis shows the median FLD of a probeset across three 96-sample plates. GC-rich probesets tend to show higher FLD using the Axiom SwiftArray workflow compared to the Axiom 2.0 workflow.

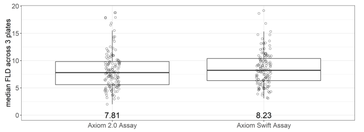
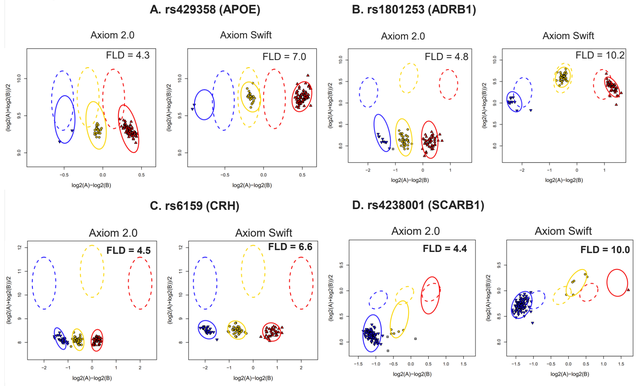


Figure 5: Cluster plots showing side-by-side comparison for probesets with GC content ≥ 70% for four pharmacogenetic markers. The four probesets show good cluster resolution with both workflows, with the Axiom SwiftArray Assay exhibiting slightly better resolution as determined by the FLD metric. Ovals indicate the prior and posterior probabilities used in genotyping calls, where the dotted ovals represents the prior probability (before observing the data) and the solid ovals represents the posterior probability (updated after incorporating the observed data). Well-segregated homozygote and heterozygote clusters indicate that the probeset is performing effectively.



Conclusions

- For important biomarkers in complex, GC-rich genomic regions, the Axiom™ PangenomePro using the new SwiftArrayStudio Microarray Solution and Axiom Swift Assay workflow deliver accurate genotyping calls and cluster resolution that is better or at par with the current Axiom 2.0 workflow.

- Axiom arrays remain highly effective tools for predictive genomics research, making them particularly valuable in fields of disease research. We have established the new Applied Biosystems™ SwiftArrayStudio™ Solution as a versatile platform that can be leveraged to design arrays targeting lower frequency variants in GC-rich regions, while supporting a wide range of sample types (Blood, Buccal, Saliva and cell lines). This further enables the discovery of disease-associated genetic variations.

References

- The 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature 526, 68–74 (2015)
- Byrska-Bishop M, Evani US, Zhao X, Basile AO, Abel HJ, Regier AA, Corvelo A, Clarke WE, Musunuri R, Nagulapalli K, Fairley S, Runnels A, Winterkorn L, Lowy E; Human Genome Structural Variation Consortium; Paul Flicek, Germer S, Brand H, Hall IM, Talkowski ME, Narzisi G, Zody MC. High-coverage whole-genome sequencing of the expanded 1000 Genomes Project cohort including 602 trios. Cell. 2022 Sep 1;185(18):3426-3440.e19. doi: 10.1016/j.cell.2022.08.004.

Science at a scan
Scan the QR code on the right with your mobile device to download this and many more scientific posters.

Learn more at thermofisher.com/us/en/home/life-science/microarray-analysis/

For Research Use Only