Xist SNP Analysis: Consolidated Executive Summary

This report provides a streamlined overview of the Nanopore amplicon sequencing analysis of the *Xist* gene across four experimental conditions. The goal was to quantify allele-specific expression (Cast vs. B16/129) and validate the stoichiometry of the single nucleotide polymorphisms (SNPs).

1. Validated SNP Architecture

The 584 bp amplicon matches **Exon 2** of the *Xist* gene (`chrX:103466603–103467186`). We target three high-confidence SNPs that differentiate the B16 (B6) and Cast alleles.

Position (GenBank)	Genomic Pos (chrX)	B6 Allele	Cast Allele
SNP 67	103466669	**T**	**C**
SNP 334	103466936	**G**	**A**
SNP 553	103467155	**G**	**A**

2. Allele Quantification Results

We used a majority-rule assignment (at least 2 matching SNPs) to assign reads to their parent allele.

Allele Ratio Summary

Condition	Total Primary Reads	B6 Reads	Cast Reads	Cast Ratio
WT (diff)	1081	521	547	**0.512**
WT (Dox 72h)	1172	1140	28	**0.024**
dTsix (Dox 72h)	837	722	111	**0.133**
dTsix dSPEN (Dox 72h)	144	87	55	**0.387**

Key Biological Insights

- **Biallelic Baseline**: WT (diff) cells show approximately equal expression from both alleles.
- **Dox Skewing**: Dox treatment in WT cells triggers a massive shift (>97%) toward the B6 allele.
- **SPEN Dependence**: Deletion of SPEN (`dTsix dSPEN`) significantly reverses this skewing, shifting the Cast ratio back toward 0.39, reflecting a loss of silencing efficiency.

3. High-Fidelity Stoichiometry

To ensure results were not artifacts of sequencing noise, we analyzed the co-occurrence of SNPs within individual reads.

SNP Hit Distribution (B6 Allele)

Condition	1 SNP Hit	2 SNP Hits	3 SNP Hits
WT (diff)	6.3%	33.2%	**60.5%**
WT (Dox)	6.2%	33.2%	**60.6%**
dTsix dSPEN	9.2%	28.7%	**62.1%**

Co-occurrence Analysis (WT-B6)

- **SNP1 & SNP2**: 70.6%
- **SNP2 & SNP3**: 79.8%
- **SNP1 & SNP2 & SNP3**: **60.5%**

Finding: The majority of reads (~60%) carry all three target SNPs simultaneously, providing extremely high confidence in the allelic assignment and confirming the technical robustness of the Nanopore long-read data.

4. Technical Summary

- **Data Quality**: Mean Q-score of **36** (>99.9% accuracy).
- **Mapping**: **>85%** mapping rate to the target amplicon.
- **Coverage**: High depth ranging from **135x to 1100x** per sample.

All intermediate files, scripts, and detailed reports are archived in the `results/` and `scripts/` directories.