# CRISPR/Cas9 Customization of Electronic Genome Mapping Repeat Expansion Assays

Kasra Tabatabaei[1], Shuk Shukor[2], Dong Zhang[1], Xu Tan[1], Lindsay Schneider Opsahl[1], Syndi Koltz[1], Michael Kaiser[1], and John Thompson[1]

[1]Nabsys 2.0 LLC, Applications, Providence, RI, [2]Hitachi High-Tech America, Inc., San Diego, CA

## Introduction

Repeat expansions underlie numerous neurological and neuromuscular disorders, with pathogenicity often determined by repeat length. Accurate sizing can be difficult because expanded regions may be long, repetitive, and GC-rich, limiting conventional PCR and sequencing approaches.

Electronic genome mapping (EGM) analyzes ultra-long, single DNA molecules by measuring distances between sequence-specific tags. An expansion changes the interval spanning the repeat locus, providing direct molecule-level evidence of allele size. The RepX™ Repeat Expansion Analysis pipeline uses these intervals to detect, phase, and quantify expanded alleles. Previously, RepX performance was demonstrated for *FMR1* and *FXN*. Herein, we demonstrate RepX pipeline performance in *DMPK* and *C9orf72*.

**Analytical Challenge**
Long and sequence-complex repeat regions can be difficult to characterize using amplification- or sequencing-based methods.

**Single-Molecule Analysis**
EGM measures molecules independently, allowing distinct alleles and heterogeneous repeat populations to be observed.

**Study Objective**
Evaluate standard EGM and RepX for detecting and sizing repeat expansions in *DMPK* and *C9orf72* research samples.

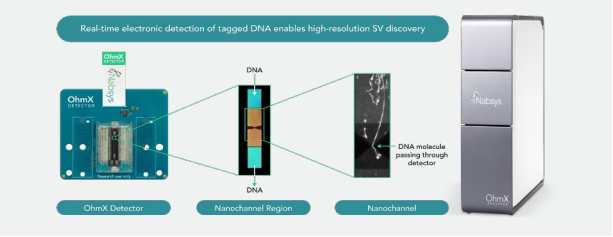


**Figure 1. EGM detector and analyzer.** EGM is enabled by the novel, solid-state nanodetectors developed by Nabsys. The OhmX™ Detector houses 256 parallel nanochannels, each with its own electronic sensor. Nabsys has developed DNA tagging chemistry that provides a high signal-to-noise. The high signal and spatial resolution allows tags to be closely spaced on the high molecular weight DNA molecules. This yields optimal tag density such that structural variations as small as 300 bp can be routinely identified and mapped genome-wide.

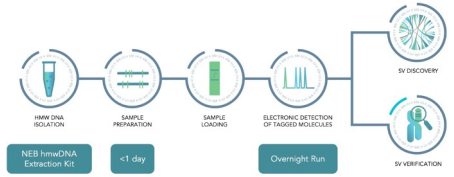


**Figure 2. Overview of EGM and OhmX Genome Preparation workflow plus CRISPR.** The EGM workflow consists of the following steps: (1) high molecular weight (HMW) DNA isolation, (2) genome preparation using the OhmX Velocity Optimized Labeling and Tagging (VOLT™) Kit, (3) sample loading, (4) electronic detection of tagged molecules using the OhmX Analyzer, and (5) data analysis.

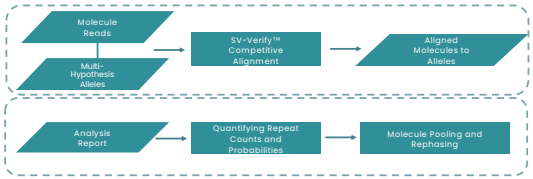


**Figure 3. Methodology for the RepX pipeline.** In the first stage, DNA molecule reads are competitively aligned to multiple alleles with various repeat expansion lengths (multi-hypothesis alleles). In the second stage, aligned molecules are pooled, phased, and quantified for repeat expansion metrics. Analysis report output provides phased repeat unit counts, 95% confidence intervals, as well as probabilities of observed repeat expansions. This pipeline is customizable for other target genes and regions suspected to contain repeat expansions.

## SV-Discover™ *DMPK* and *C9orf72* Normal Repeat Locus Assembly

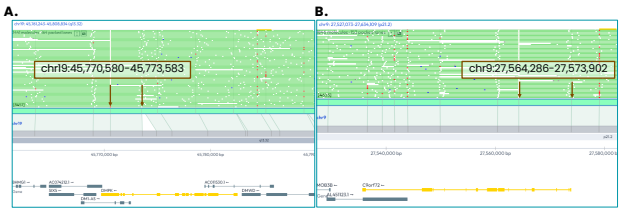


**Figure 4. SV-Discover pipeline assembled high-coverage contigs at repeat expansion sites in HG002 (300×).** Repeat expansion interval start and end labels highlighted on GRCh38 reference with red arrows. **A.** Assembled contig at *DMPK* gene with molecule support visualized. **B.** Assembled contig at *C9orf72* gene with molecule support visualized. No mapping issues are expected at either locus.

## RepX Pipeline on Known Positive *DMPK* and *C9orf72* Coriell Cell Lines

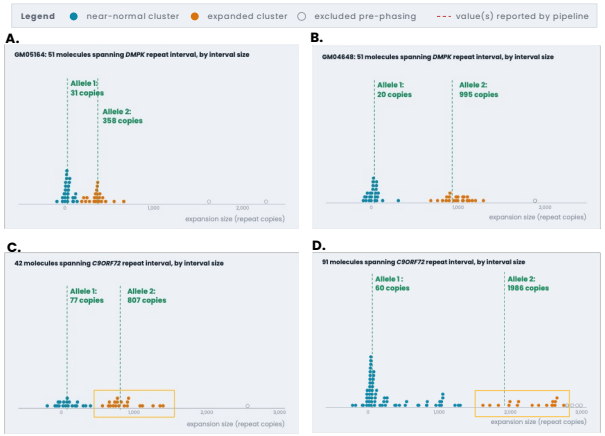


**Figure 5. Representative data of Coriell cell lines with expected positive *DMPK* and *C9orf72* repeat expansions.** Plotting observed molecule interval sizes aligned repeat expansion interval of interest. X-axis expansion size (repeat copies) is a function of observed minus GRCh38 repeat expansion interval size. Y-axis denotes frequency of molecules at specific repeat copies. Blue, orange, and grey dots denote normal, expanded, and outlier repeat copies, respectively. Orange boxes highlight sporadic expanded repeat copies and slippage. **A.** GM05164 RepX *DMPK* pipeline results. B. GM04648 RepX *DMPK* pipeline results. **C.** ND10221 RepX *C9orf72* pipeline results. **D.** ND10773 RepX *C9orf72* pipeline results.

## Result Highlights

| Gene | Cell Line ID | Expected Repeat Counts from Coriell (* or literature) | Observed Repeat Counts |
|---|---|---|---|
| DMPK | GM04648 | 5–27 / ~1000 | 20 / 995 |
| | GM05164Rep1 | 21 / ~340 | 31 / 358 |
| | GM05164Rep2 | 21 / ~340 | 12 / 335 |
| | GM23265 | 12 / 60 | 62 / 62 |
| C9orf72 | ND10221 | 6 / 249 / 738 * (3 bands; 1 normal + 2 expanded) | 77 / 807 |
| | ND10773 | 2 / Unresolved * (2 bands; 1 normal + 1 unresolved) | 60 / 1986 |
| | ND12099 | 2 / 90 / 171 * (3 bands; 1 normal + 2 expanded) | 50 / 413 |

**Table 1. Tested cell lines and results summary.** Expected Repeat Counts with "*" were only reported as positive by Coriell. Bram et al. 2018 used PCR/gel sizing to assess repeat counts. Normal allele lengths vary within that range.

## Targeted Nicking of Custom Sites in Human DNA

Multiple classes of nicking sgRNA have been developed and tested for human sample applications:

- Repeat sgRNA that target many sites across the genome where the nicking density is not optimal for molecule alignment, including targeted nicking of the centromere for Robertsonian translocation analysis
- Unique sgRNA targeted to the following:
  - Repeat Expansions—such as Fragile X
  - Genes of interest—such as genes involved in evaluating FSHD

The RepX pipeline maps ultra-long DNA molecules to quantify the number of repeats associated with a repeat expansion. When a repeat expansion of interest falls within an interval that is not optimal for the RepX pipeline, using CRISPR/Cas9 allows for optimization of the target interval size and surrounding nick sites. Fragile X was studied as a proof of concept.

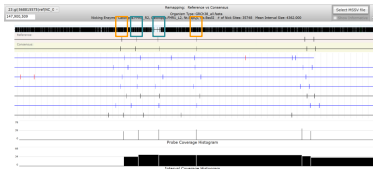

**Figure 6. sgRNA nicking changes the interval size for repeat expansion sizing of the *FMR1* locus.** The interval size is decreased from 6,099 bp (orange boxes) to 1,728 bp (blue boxes).

**Table 2. Repeat sizing of Fragile X samples with and without CRISPR nicking.** The CRISPR assay is efficient at introducing nicks into the sample, allowing accurate repeat expansion calling.

| Sample | CRISPR Nicking | Expanded Allele | | | | |
|---|---|---|---|---|---|---|
| | | Repeat Annotation | P (Elevated) | P (Full Mutation) | Repeat Count | 95% Confidence Interval |
| GM20239 | No | 183–193 | 100% | 43.5% | 192 | 117–271 |
| GM20239 | Yes | 183–193 | 100% | 48.8% | 198 | 151–247 |
| GM20239 | Yes | 183–193 | 100% | 21.9% | 184 | 148–221 |
| GM20239 | Yes | 183–193 | 100% | 65.7% | 207 | 170–247 |

## Using EGM for sgRNA/Cas9 Characterization

Further analysis of samples using Cas9 D10A nickase and dCas9 shows that EGM can be used for base-pair-level analysis of sgRNA and for assessing Cas9 proteins.

In addition to targeted genetic analysis, proof of concept testing shows that EGM can be used for the following:

- Test the distance an sgRNA can be from a nick site and still block using dCas9 (Table 3)
- Characterize the relative efficiency of Cas9 enzymes to nick and block target sites

| Distance from sgRNA | Number of examples | Median Percent Tagged (Assembly 1) | Median Percent Tagged (Assembly 2) |
|---|---|---|---|
| 5 | 18 | 53.4% | 45.7% |
| 9 | 40 | 96.9% | 97.1% |
| 10 | 4 | 94.9% | 91.8% |
| 12 | 1 | 94.4% | 97.1% |
| 14 | 4 | 92.7% | 96.6% |
| 15 | 1 | 98.7% | 97.4% |
| 24 | 1 | 96.4% | 95.0% |
| 9–24 | 51 | 96.4% | 96.9% |

**Table 3.** Distance of sgRNA from nick site for blocking evaluated across 2 assemblies using Human Chromosome Explorer®.

## Conclusions

- **Concordance:** RepX pipeline accurately phased and estimated repeat counts concordant to literature. Wide range of expanded intervals, suggesting genomic instability, also observed in interval size plots.
- **RepX Panel Improvements:** Successfully demonstrated RepX pipeline performance in quantifying *DMPK* and *C9orf72* repeat expansion counts, expanding upon published *FMR1* and *FXN* performance.
- **Pipeline Performance Updates:** Analytics have been improved for more robust RepX pipeline performance

**Learn more at Nabsys.com**

© 2026 Nabsys. For Research Use Only. Not for use in diagnostic procedures.