Comparative Analysis of Optical Genome Mapping and Genomic Proximity Mapping for Genomic Stability Assessment of hPSCs

Seth Taapken[1], Kim Leonhard[1], Anna Lisa Larson[1], Amber Kuhn[1], James Johnson[1], Tenneille Ludwig[1,2]

[1]*WiCell Research Institute* [2]*University of* Wisconsin-Madison — *Madison, WI*

Introduction

Genomic stability assessment is critical to human pluripotent stem cell (hPSC) characterization and quality control. G-banded karyotyping and SNP microarray are accepted methods for evaluating genomic integrity in research and translational settings and are included in the ISSCR Standards for Human Pluripotent Stem Cell Research (www.isscr.org/guidelines).

However, these methods have known limitations: restricted resolution, limited sensitivity for low-level mosaicism, and inability to detect certain classes of structural variants.

Optical genome mapping (OGM) and genomic proximity mapping (GPM) are emerging alternatives that use distinct detection methods to provide genome-wide structural variant data with improved resolution, each with its own strengths and weaknesses. We evaluated these technologies against existing methods to highlight their potential utility.

In this study, 20 hPSC lines with previously characterized chromosomal abnormalities were analyzed by OGM, SNP microarray, and karyotyping to assess concordance and detection capability across platforms. A subset of 4 lines with structural variants not detected by OGM — including variants near centromeric regions, low-level mosaic events, and isochromosome 20q10 — were further evaluated using GPM.

Genomic Variant	GPM	OGM	Karyotype	Microarray
Insertion (kb)	50	5	5000	50
Deletion (kb)	25	7	5000	50
Duplication (kb)	25	150	5000	50
Inversion (kb)	800	70	5000	N/A
Translocation (kb)	13	70	5000	N/A
Mosaic Detection Limit (%)	5 to 10	5 to 10	5 to 15	10 to 20
Resolution (kb)	5	5	5000	50

Table 1. Comparison of genomic variant resolution and sensitivity limits across assays. Some limits can be affected by read depth or metaphase counts.

Methodology

Samples for assessment were selected based on the different abnormalities present (identified by karyotype and SNP array). These represent many types of different variants at differing levels of mosaicism within the cultures. Samples were selected to challenge OGM and GPM based on their detection methodologies.

Karyotype Analysis

Visual assessment of human chromosomes. G-banded karyotyping performed per published protocol (McIntire et al. 2021); results reported per the International System for Human Cytogenomic Nomenclature (ISCN).

SNP Microarray

Use of SNPs to determine copy number variation and absence of heterozygosity. Genomic DNA extracted from cell cultures and run on the Illumina Global Diversity Array; analysis in Bionano VIA.

Optical Genome Mapping

Optical mapping of the human genome using fluorescent labeling. High-quality UHMW DNA isolated and fluorescently labeled (Bionano Prep), then run on the Bionano Saphyr system; Bionano VIA aligns and calls structural variants.

Genomic Proximity Mapping

Sequencing-based measure of chromosome interaction from which genomic variants can be inferred. Chromatin crosslinked and adjacent DNA fragments ligated (Phase Genomics platform); sequencing of junctions identifies structural rearrangements.

Results

92%
structural CNV concordance between OGM and SNP microarray, across 20 hPSC lines

59%
chromosomal variants concordant between OGM and g-banded karyotype, across the same 20 lines

OGM showed high concordance with SNP microarray for copy-number variation, but lower concordance with karyotype — missing more translocations (balanced and unbalanced), an isochromosome, and a low-percentage mosaic whole-chromosome gain. OGM data did not consistently resolve changes that originated near the chromosome centromere where repetitive sequences make the fluorescent labeling aspect of OGM technically difficult.

In the 4-sample GPM subset, GPM showed better concordance with karyotype (4/6 abnormalities) than with SNP microarray (14/26 variants). Conversely, OGM showed better concordance with SNP microarray in this subset than with karyotype, missing all karyotypic changes noted and identifying only 24/26 SNP-microarray variants. GPM showed difficulty resolving changes that were seen in karyotype at lower mosaic levels despite having a higher theoretical sensitivity.

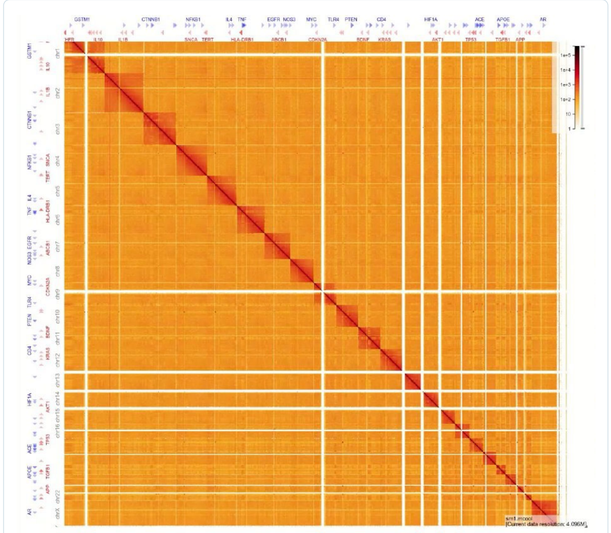

Figure 2. Example of GPM data — a heatmap of the genome-wide contact matrix where intensity indicates contact frequency. GPM is sequencing-based, measuring spatially adjacent DNA fragments; breakpoints are inferred from expected vs. observed contacts.

4 / 6
karyotype abnormalities concordant with GPM (n=4 subset)

20
hPSC lines analyzed by OGM, SNP microarray, and karyotype for this study

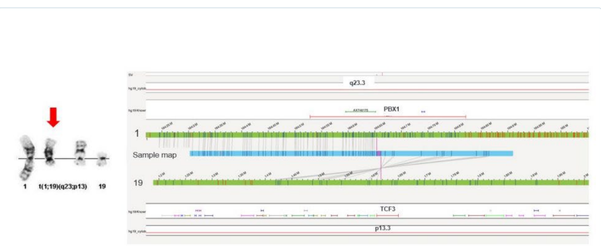

Figure 3. Example of OGM data consistent with a translocation. OGM is an imaging-based technique that linearizes and fluorescently labels ultra-high-molecular-weight DNA to detect genome-wide structural variants; breakpoints appear as alignment discontinuities.

Summary

- OGM and GPM both showed clear strengths over g-banded karyotype in resolving genomic structural variants with high resolution and more precise breakpoints.
- OGM showed high concordance with SNP microarray (81/88 CNVs) but lower concordance with karyotype (14/24 variants).
- GPM showed better concordance with karyotype data in the smaller subset of samples but still did not resolve some problematic findings including the isochromosome 20q, which is a known recurrent abnormality in hPSC.
- Both OGM and GPM identified additional changes not resolved by any other technology — notable, though these findings could not be independently validated and were excluded from this analysis. We acknowledge that these changes may be some of the real strengths of these new methods; however, this study's goal was to evaluate these technologies against existing standard methodologies.
- Some categories of change appeared specifically difficult for OGM to resolve. These were specifically changes that originated in areas of segmental duplications or repetitive sequences near the chromosome centromere. This characteristic led to several karyotypic changes not being identified by OGM.
- While theoretical detection of mosaicism is higher in OGM or GPM as compared to karyotype (*see table 1*), both assays are initiated from a pooled starting material and did not demonstrate higher sensitivity to known changes in this study. Both OGM and GPM missed structural and copy number changes detected by karyotype at a low mosaic level.
- The robustness of the OGM assay was assessed as part of this study as well. While acknowledging that WiCell was relatively new to running the OGM assay, we did find that a high number of samples needed to be rerun. Additionally, UHMW DNA takes skill to isolate, and this is a limiting aspect of the OGM assay.

Sample	Karyotype Results (All missed by OGM, Bold missed by GPM)	Microarray Results (Green missed by OGM, Blue missed by GPM)			
		ISCN	Event	Length	GRCh38 region
2	46,XX, del(6)(q14q16)[10]/46,XX[10]	1q31.3(196801046_196886770)x1	Loss	85725	chr1:196831916-196917640
		6q14.3q16.3(87301159_103046255)x1~2	Loss	16016097	chr6:86320441-102598380
		11q14.1(78481551_78939766)x1	Loss	458216	chr11:78770506-79228721
		17p11.2(172230930_17435179)x2~3	Gain	112250	chr17:17419616-17531865
		17q12 3(34523009_34804303)x1	Loss	281295	chr17:36,154,470-36,435,764
		20p12 1(14461753_14745232)x1	Loss	103480	chr20:14661107-14764586
3	**46,XY,i(20)(q10)[2]/46,XY[18]**	6p22.1(29917099_29937896)x1	Loss	20798	chr6:29949322-29970119
		12p13.31(8569612_8710101)x1	Loss	140409	chr12:8471016-8557506
		14q13.1(33635038_33676878)x1	Loss	23371	chr14:33184302-33207672
		16q24 1(86265844_86265059)x1	Loss	29216	chr16:86232238-86261453
		16q24 3(89628776_90007034)x3	Gain	378259	chr16:89562368-89940626
		18p11.31p11.23(7076562_7567500)x2~3	Gain	490939	chr18:7076563-7567502
		19q13.2(41349318_41381212)x1	Loss	31895	chr19:40843413-40875307
		Xp22.33 or Yp11.32 - p11.31(60425_26981)x1	Gain	2637746	chrX:10425-2780129
4	45,XX,der(14;15)(q10;q10)[20]	4q26(118217638_118841712)x3	Gain	624075	chr4:117296482-117920551
		20p12 1(14946105_15086816)x1	Loss	140712	chr20:14965459-15106170
		Xp21.1(31977286_32044080)x1	Loss	67233	chrX:31959169-32026491
5	46,XX,t(12;16;15)(q21;q11.2;q26.1)[18]/46,XX,t(12;16;15)(q21.2;q11.2;q26.1),**t(18;19)(q11.2;q13.3)[2]**	2q21.1(131477911_131997466)x1	Loss	519556	chr2:130720338-131239893
		3p26.1(5389728_5440226)x1	Loss	50499	chr3:5348043-5398540
		4p15.33(13394077_13571720)x3	Gain	177644	chr4:13392453-13570096
		7p22.3(1940748_2003725)x1	Loss	62978	chr7:1901112-1964090
		8q22.3(103052377_103109921)x1	Loss	57545	chr8:10240149-10207693
		12q21.31(81156788_81211782)x3	Gain	54995	chr12:80763009-80818003
		13q21 1(57789934_57858013)x1	Loss	68808	chr13:57215800-57283879
		19p12 20099816_20716993)x1~2	Loss	117124	chr19:20471010-20534133
		Xq23 (115886817_116045594)x3	Gain	158778	chrX:116752849-116911626

Table 2. Results comparison across karyotype, SNP microarray, OGM, and GPM for representative sample variants (ISCN nomenclature, event type, length, and GRCh38 region).

References

McIntire E, Leonhard K, Taapken S, Larson AL. G-banded karyotyping of human pluripotent stem cell cultures. In: Hu K, ed. Nuclear Reprogramming. Methods Mol Biol. 2021;2239. Humana.

International Society for Stem Cell Research. Guidelines for Stem Cell Research and Clinical Translation. ISSCR, 2021. www.isscr.org/guidelines.

Acknowledgements

We thank Bionano and Phase Genomics for working with and training WiCell on sample preparation and data analysis. We thank WiCell staff for their efforts in generating this data, and ISSCR for the opportunity to share this work.