Hi-C detects novel structural variants in HL-60 and HL-60/S4 cell lines

Elsie C. Jacobson[1], Ralph S. Grand[1,2], Jo K. Perry[1], Mark H. Vickers[1], Ada L. Olins[3], Donald E. Olins[3], Justin M. O'Sullivan[1]*

[1] Liggins Institute, University of Auckland, Auckland, New Zealand
[2]Friedrich Miescher Institute for Biomedical Research
[3]University of New England, Portland, ME, USA

Abstract

Cancer cell lines often have large structural variants (SVs) that evolve over time. There are many reported differences in large scale SVs between HL-60 and HL-60/S4, cell lines derived from the same acute promyelocytic leukemia sample. However, the stability and variability of inter- and intra-chromosomal SVs between different sources is unknown. Here, we used Hi-C and RNA-seq to identify and compare large SVs in HL-60 and HL-60/S4 cell lines.

Comparisons with previously reported karyotypes identified two non-canonical SVs in HL-60. Ten previously unreported SVs were identified in HL-60/S4. The unreported SVs were generally small enough to be plausibly undetected with traditional karyotyping methods. An expansion centered on MYC was found in a novel genomic location in HL-60. The HL-60 cell line has more within-line structural variation than the HL-60/S4 derivative cell line.

HL-60 and HL-60/S4 karyotypes are generally consistent with the current literature, with some exceptions. Hic_breakfinder is an effective tool for identifying all SVs, but intra-chromosomal SVs are less reliably detected across different samples. The orientation of SV patterns, and strandedness of gene fusions, allowed us to differentiate inversions from other forms of intra-chromosomal SV. Visual inspection of Hi-C heatmap patterns allow further characterization of SVs without additional methods, although orthogonal information such as gene fusions contribute to the characterization.

Background

Karyotyping with fluorescence in situ hybridization (FISH) is an effective way of identifying large scale structural variation, but it has relatively low resolution [1]. Genome wide chromatin conformation capture (Hi-C), which was developed to identify spatial genome organization [2,3], is emerging as a useful method for de-novo genome assembly [4] and identifying structural variants [5,6].

The HL-60 cell line was derived from a female acute myeloid leukemia patient in 1976 [7]. Since that time, the HL-60 cell line has been instrumental to breakthroughs in cancer, immunology, and cell biology. For example, HL-60 cells were used to identify retinoic acid as a therapeutic for acute promyelocytic leukemia [8,9], investigate myelocytic differentiation [10,11], study neutrophilic behaviors such as migration and respiratory bursting [12–15], and understand the mechanical properties of nuclear envelope components [16,17].

Several cell lines have been derived from the original HL-60 cells, including HL-T [18], HL-60/MX2 [19], and HL-60/S4 [20], however, the original HL-60 line [7] and the HL-60/S4 variant are currently the most frequently used in empirical studies [20,21]. The HL-60/S4 subline was generated using random mutagenesis and selection for cells supersensitive to retinoic acid-induced

differentiation [20]. A recent karyotype identified novel structural variants in the HL-60/S4 genome when compared to the parental HL-60 cell line [22].

A recent survey of MCF7 sublines showed rapid development of genetic diversity and subsequent functional changes in this breast cancer cell line, demonstrating that they cannot be used interchangeably for drug discovery screens or other purposes [23]. While HL-60 and HL-60/S4 are recognized as different cell lines by ATCC, they are both used for similar studies [22,24]. It is notable that despite the HL-60/S4 variant being developed as a super-responder to retinoic acid-induced differentiation; modern studies typically treat both HL-60 and HL-60/S4 with 1μM retinoic acid for four days to induce differentiation into the granulocytic form [22,24].

Here, we analyzed Hi-C and RNA-seq data from two HL-60 sublines (HL-60 and HL-60/S4) to identify large structural variants (SVs) and gene fusions. We confirm that previously reported SVs can be detected in our HL-60 and HL-60/S4 samples. We also find that the known MYC gene expansion is highly stable, although it has inserted into different genomic locations in the HL-60 and HL-60/S4 genomes. We identify a number of novel intra-chromosomal SVs, many of which were only computationally detected in one dataset using hic_breakfinder but were visually identified in other dataset(s). Karyotyping of HL-60/S4 has previously identified a single inversion on one homolog of chromosome 2. Here we identified multiple additional intra-chromosomal SVs in HL-60/S4.

Results

Structural variants detected with Hi-C and RNA-seq

We used Hi-C data from two sublines of the HL-60 cell line to detect SVs. Hi-C detects a high frequency of contacts between regions of DNA in close linear proximity, a feature which has been exploited to improve genome assemblies [4] and detect structural variation [5]. It has been previously noted that very high contact frequencies between distal regions of the genome are likely to represent a translocation or inversion compared to the reference genome [4–6].

We detected SVs from Hi-C data in three independent experiments across two versions of the HL-60 cell line. Two independently generated Hi-C experiments were performed using HindIII and MboI and HL-60/S4 cells grown in RPMI 1640 with 1% pen/strep and 10% serum, in the same laboratory, two years apart [25]. An additional Hi-C data set, generated using HaeIII and HL-60 cells grown in RPMI 1640 with 1% penicillin/streptomycin and 10% serum [24] was retrieved from GEO GSE93997. The Hi-C data set from Jacobson et al [25] was generated using HL-60/S4 cells that had been differentiated into granulocytes using all-trans retinoic acid (ATRA) treatment over four days. Therefore, to avoid any treatment-specific effects, we only investigated the genomic structure in ATRA-treated HL-60/S4 cells.

We used hic_breakfinder v1 [6] to identify structural variants from our Hi-C data. Regions of unusually high contact frequency, indicating a possible SV ('breaks'), were identified in genome structures captured in the HL-60-HaeIII [24], HL-60/S4-HindIII (this study) and in the HL-60/S4-MboI [25] datasets. To further investigate each of these breaks, we visually inspected the putative SVs in Hi-C contact matrices visualized in HiGlass [26]. We found that hic_breakfinder identified five short range intra-chromosomal breaks in the HL-60/S4-MboI data set that were not supported by visual inspection of the Hi-C heatmap. Therefore, we filtered the hic_breakfinder results to remove breakpoints that were less than 10Mb apart (Supp table 1). Of the 67 remaining breaks (Supp table 1), all except two (1—1, 3—8; both present in the HL-60/S4-MboI dataset [25]) were supported by visual heatmap inspection. In some instances, multiple break locations

represent a single SV, and in total we identified 27 SVs. Seven SVs were present in both HL-60 and HL-60/S4, one was unique to HL-60, and 19 were unique to HL-60/S4 (Table 1).

Table 1. Locations of SVs identified within Hi-C datasets for HL-60-HaeIII [24]**, HL-60/S4-HindIII (this study), and HL-60/S4-MboI** [25] **cell lines.**

Location 1	Location 2	HL60-HaeIII	HL60/S4-HindIII	HL60/S4-MboI	Fusion Gene	Strand	Distance	Predicted Type
chr1 (p36.32)	chr3 (q28)	-	breakfinder	breakfinder			Inter	Translocation
chr2 (p24.2)	chr2 (p13.1)	-	STAR-Fusion	breakfinder	SMC6--SLC4A5	-/-	Intra	Translocation
chr2 (p22.1)	chr2 (q31.2)	-	breakfinder	breakfinder			Intra	Inversion
chr2 (p15)	chr2 (p13.1)	-	breakfinder	visual			Intra	Inversion
chr2 (q14.1)	chr2 (q22.1)	-	visual	breakfinder			Intra	Deletion
chr2 (q22.3)	chr2 (q24.3)	-	visual	breakfinder			Intra	Inversion
chr3 (q28)	chr14 (q12)	-	breakfinder	breakfinder	IL1RAP--AL136018.1	+/+	Inter	Translocation
chr4 (p15.31)	chr4 (p15.1)	-	visual	breakfinder			Intra	Inversion
chr4 (p15.2)	chr8 (q24.13-q24.21)	breakfinder	-	-			Inter	Insertion
chr4 (q35.2)	chr18 (q21.1)	-	breakfinder	breakfinder			Inter	Translocation
chr5 (q23.3)	chr5 (q31.3)	breakfinder	visual	breakfinder			Intra	Deletion
chr5 (q31.2)	chr7 (q32.3)	breakfinder	breakfinder	breakfinder			Inter	Translocation
chr5 (q33.3)	chr16 (q23.2-q23.3)	breakfinder	breakfinder	breakfinder	CYFIP2--PLCG2	+/+	Inter	Translocation
chr5 (q11.2)	chr17 (p11.2)	breakfinder	breakfinder	breakfinder			Inter	Translocation
chr6 (p22.3)	chr6 (p12.1)	-	visual	breakfinder			Intra	Inversion
chr6 (q22.2)	chr8 (q24.13-q24.21)	-	breakfinder	breakfinder			Inter	Insertion
chr7 (q32.3)	chr16 (q24.1)	breakfinder	breakfinder	breakfinder			Inter	Translocation
chr8 (q24.13-q24.21)	chr11 (p14.1)	-	breakfinder	breakfinder			Inter	Translocation

Continued

Location 1	Location 2	HL60-HaeIII	HL60/S4-HindIII	HL60/S4-MboI	Fusion Gene	Strand	Distance	Predicted Type
chr9 (q31.1)	chr14 (q23.2)	breakfinder	breakfinder	breakfinder			Inter	Translocation
chr10 (p12.1)	chr13 (q12.12)	breakfinder	-	-			Inter	Translocation
chr14 (q12)	chr21 (q22.11-q22.12)	-	breakfinder	breakfinder			Inter	Translocation
chr15 (q14-q21.1)	chr16 (p12.1-p11.2)	-	breakfinder	visual			Inter	Translocation
chr16 (p13.3-p13.11)	chr17 (q11.1-q25.2)	-	breakfinder	visual			Inter	Translocation
chr16 (p12.1)	chr21 (q22.11)	-	breakfinder	breakfinder			Inter	Translocation
chr17 (q21.31)	chr17 (q25.3)	-	breakfinder	visual			Intra	Inversion
chr18 (q21.1)	chr18 (q21.32)	-	visual	breakfinder			Intra	Unclassified
chr20 (p11.23)	chr20 (q11.23)	-	visual	breakfinder	RALGAPB--RBBP9	+/-	Intra	Inversion
chr15 (q11.2)	chr21 (q22.11)	-	STAR-Fusion	STAR-Fusion	TUBGCP5--TMEM50B	-/-	Inter	Translocation

Location 1 and 2 indicate the positions of the likely breakpoints. The next three columns indicate the highest level of evidence provided for the SV occurring in that sample, in the order: Hic_breakfinder (breakfinder), STAR-Fusion, visual inspection in HiGlass (visual), or none (-). The Fusion Gene column indicates whether a fusion gene was identified at the SV. The Type column indicates the likely type of SV, based on Hi-C patterns and fusion transcripts where possible.

We used STAR-fusion v1.4.0 [27] to identify gene fusion products that have been generated by SVs within the HL-60 and HL-60/S4 cell transcription profiles. Gene fusions were identified in two RNA-seq experiments from two different laboratories in HL-60/S4 cells [22,28]. These RNA-seq datasets each contained 12 samples from three [23] and four [28] conditions respectively. We detected gene fusions in an RNA-seq experiment [24] in HL-60 cells, the same source that was used to acquire Hi-C data from, containing four samples from two conditions [24]. All experimental datasets [22,24,28] included both undifferentiated and granulocytic (i.e. all-trans retinoic acid differentiated) cells. STAR-fusion identified a total of 93 putative, unique gene fusions. To reduce the number of spurious fusion calls, we only considered gene fusions that were detected in all samples of a single condition, leaving 17 putative fusion candidates (Supp table 2). We then visually inspected the corresponding Hi-C heatmaps to identify SVs that could result in the predicted gene fusion. We identified one previously reported gene fusion in both HL-60 and HL-60/S4 [29], and four previously unreported gene fusions in HL-60/S4 (Table 1).

Translocations are reliably detected with hic_breakfinder

Both HL-60 and HL-60/S4 cell lines have previously reported cytogenetic information [22,30], so we compared these to our Hi-C breakfinder results. The DSMZ catalogue describes the HL-60 karyotype (DSMZ no.: ACC 3), and a recent study has reported FISH karyotyping of HL-60/S4 [22]. All previously reported translocations and inversions were identified in HL-60 and HL-60/S4 cells (Table 1).

As described in early HL-60 cell karyotypes [31], all HL-60 cells contained a set of SVs between chromosomes 5, 7, and 16 and a t(9;14) translocation. A translocation between chromosomes 5 and 17 was not reported in the spectral karyotype, but has been described elsewhere in an early HL-60 study [32] and is described in the DSMZ cytogenetic information. The SVs t(5;7)(q31.2;q32.3), t(5;16)(q33.3;q23.2-q23.3), t(7;16)(q32.3;q24.1), t(9;14)(q31.1;q23.2), and t(5;17)(q11.2;p11.2) were detected in all samples (Table 1).

A recent FISH karyotype reported additional SVs (i.e. inv(2)(p22.1;q31.2), t(3;14)(q28;q12), t(4;18)(q35.2;q21.1), t(16;21)(p12.1;q22.11)) specific to HL-60/S4 cells [22] that we identified in both Hi-C HL-60/S4-HindIII and HL-60/S4-MboI datasets. The previously reported t(15;16)(q14-q21.1;p12.1-11.2) and t(16;17)(p13.3-p13.11;q11.1-q25.2) translocations [22] were detected with hic_breakfinder in the HL-60/S4-HindIII dataset (this study) but not the HL-60/S4-MboI dataset [25]. This apparent discrepancy may be due to restriction enzyme choice or other library preparation factors, especially given that both missing regions involve nearby regions of chromosome 16. For instance, MboI is methylation sensitive, while HindIII cuts at any site with the correct motif, regardless of methylation status. Thus, the sequencing coverage across breakpoints can be affected by the choice of restriction enzyme [33].

MYC expansion integrates into different locations within the genome

The MYC expansion that is present in HL-60 cells involves a non-contiguous region of chromosome 8 that is centered on the MYC gene (Fig 1) and amplified many times [32,34,35]. The non-contiguous region of chromosome 8q24 that comprises the MYC expansion has been reported to exist as: a) an extrachromosomal region; b) in the form of double minute chromosomes; and c) integrated into the genome [32,34–37]. The composition of the MYC expansion is highly conserved between HL-60 and HL-60/S4 cells (Fig 1a), consistent with previous reports. The massive copy number increase means that in an ICE-normalized heatmap, amplified regions appear as densely populated 'windows' compared to the surrounding sparser matrix (Fig 1a). Due to the high density of these windows, the location of the six amplified regions

that make up the MYC expansion can be identified to a high resolution (Table 2). The presence of the six amplified regions is consistent with the amplicon described in [35], based on the size and genes contained in each amplified region (Table 2). Similarly, we observed high sequence coverage of the six amplified regions in normalized DNase hypersensitivity data from ENCODE (Fig 1c, [38]).

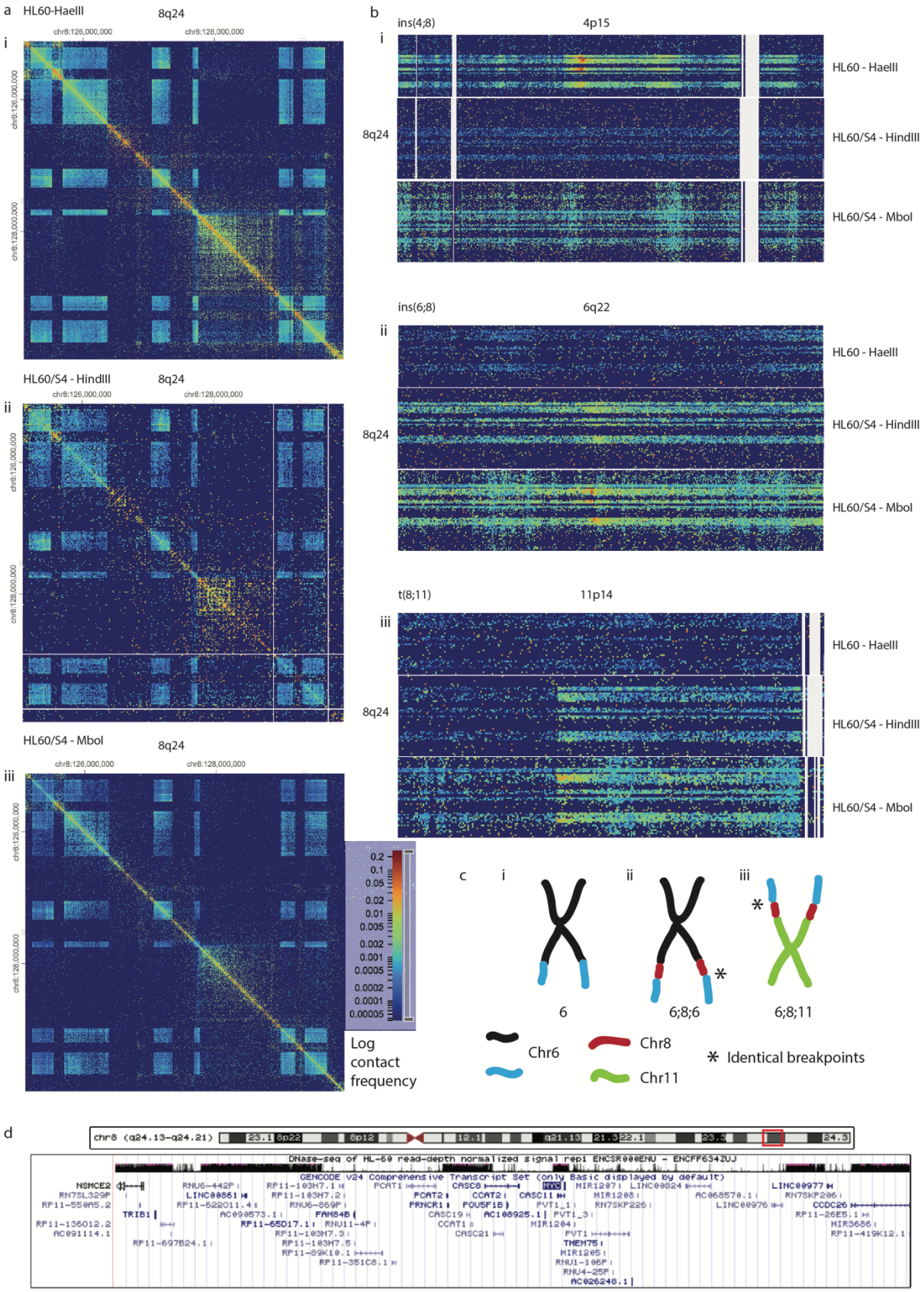
a
HL60-HaeIII
8q24
i
chr8:126,000,000
chr8:128,000,000
HL60/S4 - HindIII
8q24
ii
HL60/S4 - MboI
8q24
iii
0.2
0.1
0.05
0.02
0.01
0.005
0.002
0.001
0.0005
0.0002
0.0001
0.00005
Log contact frequency
b
ins(4;8)
4p15
i
8q24
HL60 - HaeIII
HL60/S4 - HindIII
HL60/S4 - MboI
ins(6;8)
6q22
ii
8q24
HL60 - HaeIII
HL60/S4 - HindIII
HL60/S4 - MboI
t(8;11)
11p14
iii
8q24
HL60 - HaeIII
HL60/S4 - HindIII
HL60/S4 - MboI
c
i
ii
iii
6
6;8;6
6;8;11
Chr6
Chr8
Chr11
* Identical breakpoints
d
chr8 (q24.13-q24.21)
DNase-seq of HL-60 read-depth normalized signal rep1 ENCSR000ENU - ENCFF634ZUJ
GENCODE V24 Comprehensive Transcript Set (only Basic displayed by default)

Figure 1. MYC expansion and insertions were present in both HL-60 and HL-60/S4. Hi-C heatmaps show the spatial contacts between two genomic locations, with the color of each pixel indicating the contact frequency. a) Hi-C heatmaps have been ICE-normalized to adjust for variable coverage across different regions, but the massive copy number increase of the region centered on MYC has resulted in densely-population regions of the heatmap. The amplified regions are highly conserved across i) HL-60-HaeIII, ii) HL-60/S4-HindIII, and iii) HL-60/S4-MboI. b) The amplified MYC expansion is i) inserted into chromosome 4 in HL-60-HaeIII ii) inserted into chromosome 6 in HL-60/S4-HindIII and HL-60/S4-MboI, and iii) translocated onto chromosome 11 in HL-60/S4-HindIII and HL-60/S4-MboI. c) FISH karyotyping has shown that the t(8;11) translocation is part of a derivative chromosome t(6;8;11) [22]. Given that there was only one breakpoint location identified between chromosomes 6 and 11, we can surmise the order of MYC movement throughout the HL-60/S4 genome. i) A karyotypically normal chromosome 6, gained an 8q24 insertion in the q arm ii). The resultant 8q24;6q22-6qter end of the modified chromosome was then copied and translocated onto the end of chromosome 11 iii). d) Genomic features of 8q24.13-q24.21. DNase hypersensitivity sequencing of HL-60 from ENCODE identifies the six amplified regions characterizing the MYC expansion. The MYC expansion is gene-dense, particularly with non-coding genes (Gencode v24 annotations).

Table 2. Amplified regions within 8q24

	Start (kb)	End (kb)	Coding genes	Non-coding genes
Amp 1	125,212	125,532	NSMCE2, TRIB1	RN7SL329P; AC084083.1; AC091114.1
Amp 2	125,700	126,372		AC016074.1; AC024681.2; RNU6-442P; KNOP1P5; AC024681.1; RFPL4AP5; AC087667.1; SOD1P3; LINC00861
Amp 3	127,056	127,331		CASC19; AC020688.1; PRNCR1; AC018714.1; CASC8
Amp 4	127,678	127,759	MYC	CASC11
Amp 5	128,989	129,202		
Amp 6	129,356	129,685		MIR3686; CCDC26; AC103718.1; AC011257.1

Location (hg38) and gene content of the highly amplified regions surrounding MYC. These are consistent between Hi-C libraries, and with previous reports [35].

As reported in [22], chromosome 8 is involved in two SVs, ins(6;8)(q22.2;q24.13-q24.21) and der(6;8;11)(q22.2;q24.13-q24.21;p14.1). Hi-C data for both HL-60/S4 cell lines ([25], this study) confirms that the MYC expansion is involved in translocations with chromosomes 6 and 11 (Fig 1b). However, in the HL-60 cell line, the MYC expansion appears to be inserted into chromosome

4 (ins(4;8)(q35.2;q24.13-q24.21)) (Fig 1b). One previous report has identified the MYC expansion integrating into t(5;17) [32]. By contrast, the DSMZ karyotype reports an insertion of 8q24 (containing the MYC expansion) into chromosome 1p31.

The difference between the insertions of 8q24 into chromosomes 4 and 6 and the translocation of 8q24 onto chromosome 11 is apparent in the Hi-C heatmaps ([24,25], Fig 1b). Regions of increased interactions with 8q24 are observed upstream and downstream of the chromosomes 4 and 6 breakpoints. By contrast, there are no interactions upstream of the chromosome 11 breakpoint. Notably, there is only one breakpoint identified between chromosome 6 and chromosome 8. These results indicate that the derivative chromosome der(6;8;11) resulted from a translocation between chromosome 11 and the derivative chromosome ins(6;8), and not a translocation between chromosome 6 and 11 that was followed by an insertion of the MYC expansion. Thus, the precise breakpoints identified with Hi-C can identify the likely timeline of genomic rearrangements in this instance.

Collectively, these indicate that the MYC expansion is highly conserved in its composition but inserted into different regions of the genome in HL-60 derivative strains. There could be many reasons for this, but the simplest is that the extrachromosomal MYC expansion integrated into different regions of the genome in different clones of the original HL-60 sample and has remained stably inserted during clonal expansion. However, there is some evidence to the contrary. Namely, an inversion identified in HL-60/S4 but not HL-60 inv(4)(p15.31;15.1) encompasses the location of the MYC expansion insertion (4p15.2), leaving open the possibility that the inversion is the result of MYC excision. If this is the case, the MYC expansion may continue to excise and insert into different genomic locations and could even be responsible for other SVs identified in HL-60 genomes.

Hi-C identifies non-canonical translocations in HL-60 and HL-60/S4 cell lines

We detected a translocation between chromosomes 10 and 13 (t(10;13)(p12.1;q12.12)) in the HL-60 cells that was not identified in either of the HL-60/S4 samples (Fig 2a). This translocation has been reported previously in HL-60 cells [39] but is absent from the HL-60 derived cell line HL-T [18], and is not described in the DSMZ cytogenetic information. Higher resolution inspection in HiGlass [26] confirms that this translocation is not present in HL-60/S4 datasets (Fig 2a). Karyotype analysis shows that HL-60/S4 cells have two complete versions of chromosome 10 [22], further indicating that this translocation was not present in the original leukemia sample. Moreover, the Hi-C pattern indicates the derivative chromosome consists of chromosome 10 (~25-133Mb) and chromosome 13 (~23-114Mb) (Fig 2a), a large ~200Mb chromosome that should be detectable with any cytogenetic analysis method.

We identified previously unreported translocations between chromosomes 1 and 3 (t(1;3)(p36.32;q28)), and 14 and 21 (t(14;21)(q12;q22.11-q22.12)) in both HL-60/S4 Hi-C datasets ([25], this study). Visual inspection of these putative SVs at multiple resolutions in HiGlass [26] confirmed that these translocations are not present in the HL-60 Hi-C dataset (Fig 2b,c). While the HL-60 t(10;13)(p12.1;q12.12) translocation results in a large derivate chromosome, the t(1;3)(p36.32;q28) and t(14;21)(q12;q22.11-q22.12) variants that are observed HL-60/S4 are much more subtle and would require confirmation with breakpoint PCR. In particular, the chromosome 21q22.11 translocations with chromosomes 14q12 and 16p12.1 may have generated a derivative chromosome or had two independent breaks on each homolog of chromosome 21. A further putative translocation was identified between chromosome 21q22.11 and chromosome 15q11.2. Despite hic_breakfinder not detecting a translocation between chromosomes 21 and 15 in either HL-60/S4-HindIII (this study) or HL-60/S4-MboI [25] datasets, there is a gene fusion

TUBGCP5--TMEM50B identified across all HL-60/S4 samples and conditions that supports an SV between chromosomes 21q22.11 and 15q11.2. Heatmap inspection revealed poor sequence coverage across these regions, which may have obscured the breakpoint from detection.

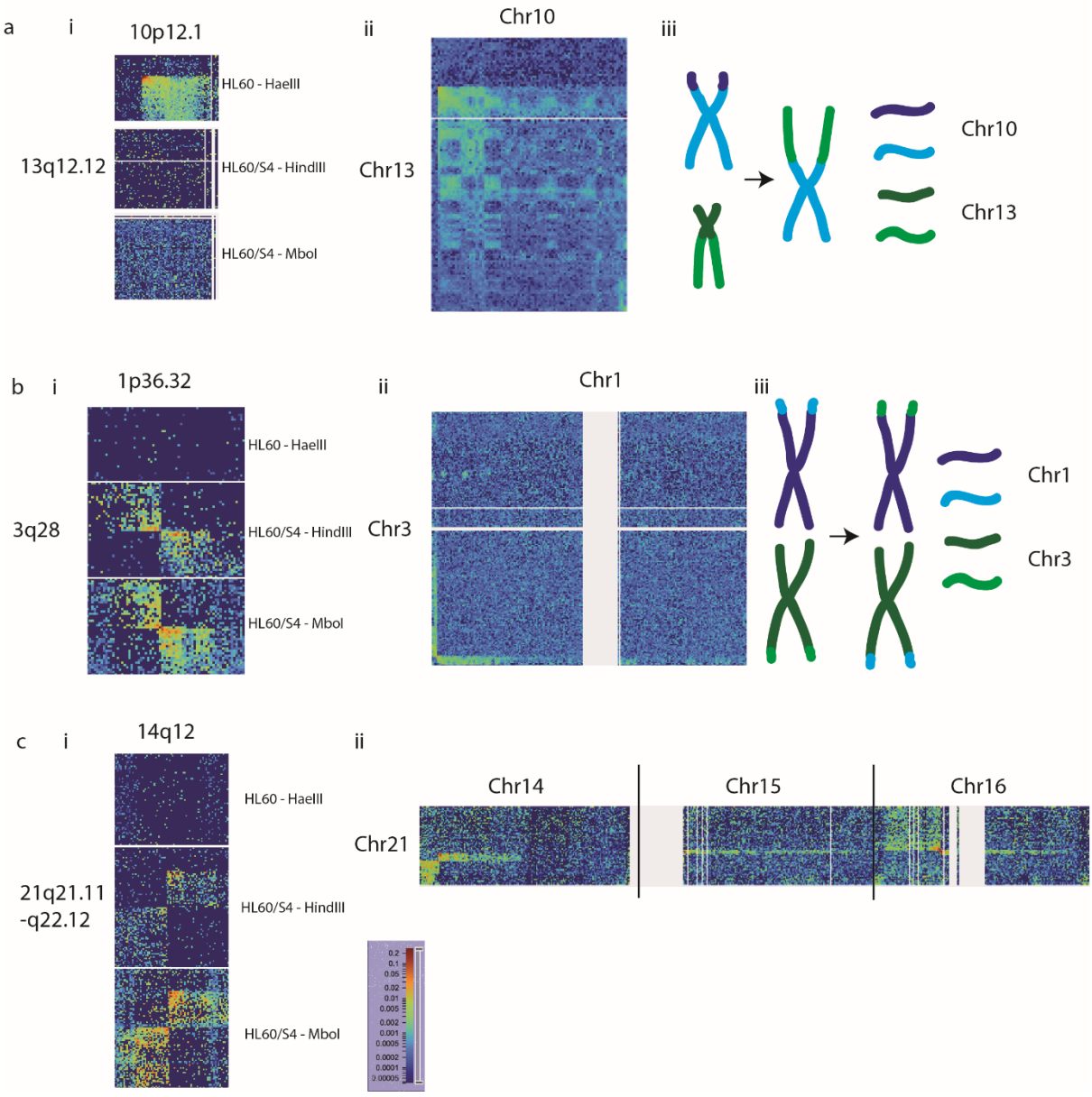

Figure 2. Non-canonical inter-chromosomal SVs in HL-60 and HL-60/S4.
a) Characterization of t(10;13) in HL-60. i) Heatmap of the breakpoint confirms that this translocation is absent from HL-60/S4. ii) Heatmap of the whole of chromosomes 10 and 13 show increased interaction frequency across the majority of both chromosomes, indicating a fusion of 13q12.2-13qter onto 10p12.1-10qter (iii). b) Characterization of t(1;3) in HL-60/S4. i) Heatmap of the breakpoint confirms that this translocation is absent from HL-60. ii) Heatmap of the whole of chromosomes 1 and 3 show increased interaction frequency between the start of chromosome 1 and the end of chromosome 3, indicating a reciprocal translocation between 1p36.2-1pter and 3q28-3qter (iii). Characterization of t(14;21) in HL-60/S4. i) Heatmap of the breakpoint confirms that this translocation is absent from HL-60. iii) Heatmap of the whole of chromosome 21 and chromosomes 14, 15, and 16, indicate a series of linked SVs.
The HL-60 t(10;13)(p12.1;q12.12) SV was present in the HL-60 Hi-C dataset, provided by the China Infrastructure of Cell Line Resources, but not the DSMZ cytogenetic information. Due to the large (~200Mb) derivative chromosome this translocation creates, its absence from cytogenetic information is unlikely to be due to usage of a low-resolution karyotyping method [40]. By contrast, the omission of the previously unreported inter-chromosomal SVs in HL-60/S4, which are small and/or complex, is likely to be due to the use of karyotyping methods that do not involve sequencing [6].

Intrachromosomal SVs are sensitive to Hi-C detection

Inter-chromosomal SVs were generally identified across all samples using both the karyotyping and Hi-C datasets. By contrast, intra-chromosomal SVs (e.g. inversions) were less likely to be previously reported, and more likely to be missed by hic_breakfinder in one of the HL-60/S4 Hi-C datasets.

A deletion on chromosome 5(q23.3-31.1) was detected with hic_breakfinder in the HL-60-HaeIII and HL-60/S4-MboI Hi-C datasets and has weak visual evidence in the HL-60/S4-HindIII Hi-C dataset (Fig 3a). This deletion has been previously detected [31], but is not reported in the DSMZ karyotype. Notably, the chromosome 5 SV contains a cancer-associated locus (5q31.1) that is frequently deleted in non-small cell lung carcinoma [41].

Inversions on chromosomes 4(p15.31;15.1) and 20(p11.23;q11.23) were identified with hic_breakfinder within the HL-60/S4-MboI Hi-C dataset, while an inversion on chromosome 17(q21.1;q21.32) was detected in HL-60/S4-HindIII Hi-C dataset (Fig 3a). An SV on chromosome 18(q21.1;q21.32) was identified with hic_breakfinder within the HL-60/S4-MboI dataset. An inversion of 18p on one homolog was reported previously [22], but this SV shows evidence of a large deletion and another unclassified SV that results in an unusual chromosome structure (Supp fig 1). Critically, all of these intra-chromosomal SVs were visible in both samples upon visual inspection of the Hi-C heatmap (Supp fig 1).

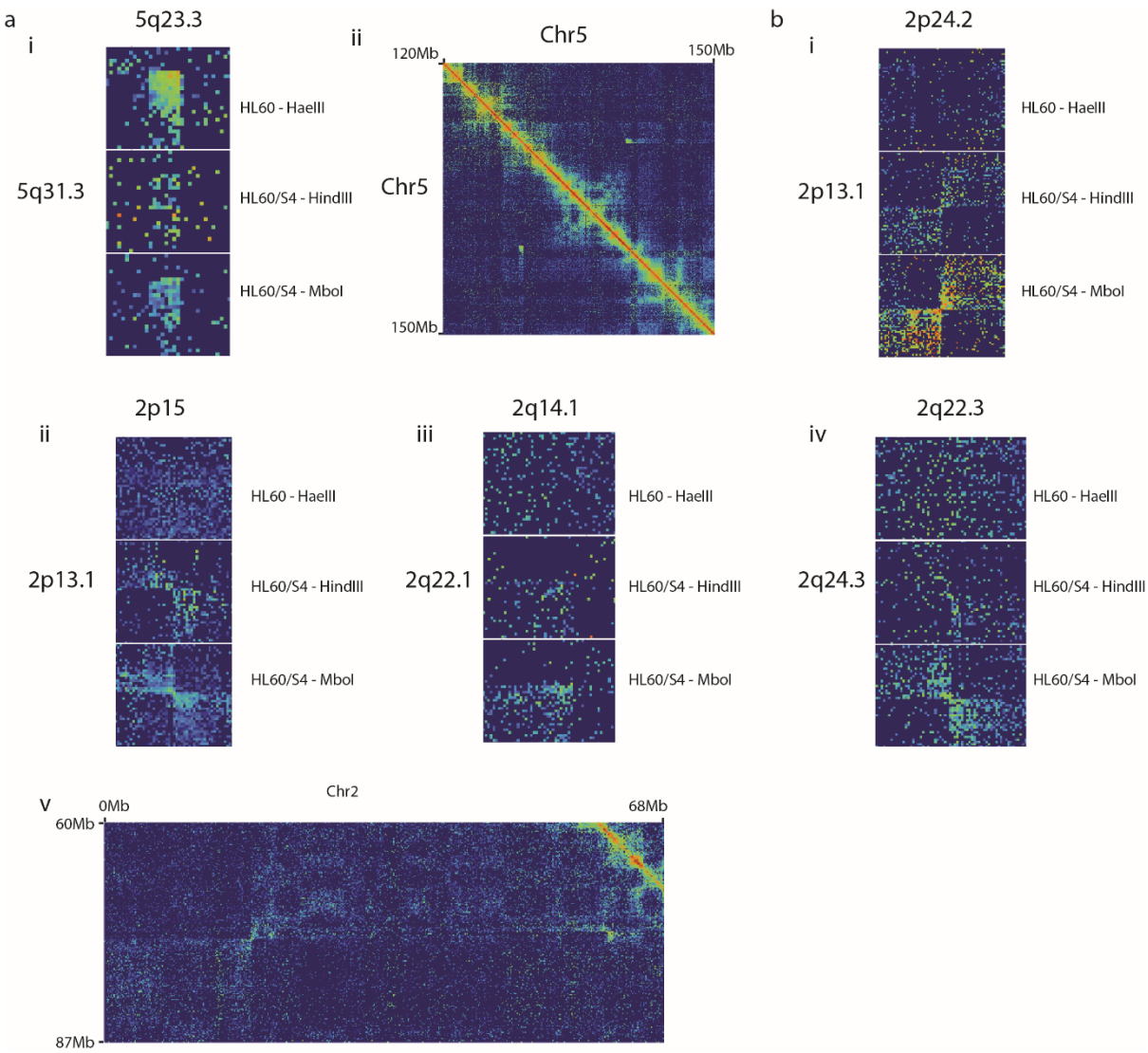

Figure 3. Heatmaps of a variety of intrachromosomal SVs.
a)i) An SV on chromosome 5 was automatically detected in HL-60-HaeIII and HL-60/S4-MboI, and there is some visual evidence in HL-60/S4-HindIII. ii) The SV pattern indicates a deletion. b) SVs on chromosome 2. i) 2p24.2—2p13.1 is present in HL-60/S4 but not HL-60 and is a likely reciprocal translocation between the two homologues of chromosome 2. ii) 2p15—2p13.1 is present in HL-60/S4 but not HL-60 and is a reciprocal inversion. iii) 2q14.1—2q22.1 is present in HL-60/S4 but not HL-60 and is likely a deletion. iv) 2q22.3—2q24.3 is present in HL-60/S4 but not HL-60 and is a reciprocal inversion. v) The SVs shown in i) and ii) both have breakpoints in 2p13.1, and so may have resulted from a single SV event or have a causal relationship.

A fusion gene arising from the t(2;2)(p24.2;p13.1) SV further indicates that the SV is the result of a translocation between homologous chromosomes, not an inversion. Specifically, gene fusions arising from inversions (e.g. RALGAPB--RBBP9) have different strandedness of the left and right parts of the chimeric transcript, as the region has been 'flipped'. Both halves of the SMC6--SLC4A5 fusion are negatively stranded, as is seen in fusions resulting from inter-chromosomal translocations such as CYFIP2--PLCG2 (Table 1).

Triaging putative gene fusions with Hi-C

Gene fusions are typically generated by genomic rearrangements [27], therefore gene fusions detected from RNA-seq data should occur at SV breakpoints identified with Hi-C data. Of the 17 gene fusions identified across all replicates of at least one condition (Table 3, Supp table 3), five unique predicted gene fusions occurred at breakpoints detected with hic_breakfinder, including two fusions located at the same breakpoint (SMC6--SLC4A5 and SMC6--AC006030.1), and a sixth (TUBGCP5--TMEM50B) had some visual support from Hi-C, despite poor sequence coverage across the region (Table 1, Fig 2cii). A further three predicted fusions had breakpoints less than 10Mb apart, and therefore their causal SV would not have been detected by our analysis. Two predicted fusions were between genes on the same chromosome, and so may have resulted from undetected SVs, given the reduced sensitivity of Hi-C to detect intrachromosomal SVs. Six predicted fusions occurred between genes on different chromosomes, and had no corresponding SV detected with Hi-C data. Although it is possible that these gene fusions were generated by an alternative mechanism, such as trans-splicing [42], we predict these are false positives and would not prioritize these candidates for follow-up (Table 3).

Table 3. Summary of predicted gene fusions.

Fusion Gene	Left Chromosome	Right Chromosome	Chromosomes	Distance	SV
AC098590.1--AC099789.1	chr4	chr1	Different	>=10Mb	No
ACTG2--ACTG1	chr2	chr17	Different	>=10Mb	No
AP004242.1--FOXP2	chr11	chr7	Different	>=10Mb	No
CYFIP2--PLCG2	chr5	chr16	Different	>=10Mb	Yes
IL1RAP--AL136018.1	chr3	chr14	Different	>=10Mb	Yes
MTRNR2L12--AC025627.3	chr3	chr17	Different	>=10Mb	No
PIBF1--KLF5	chr13	chr13	Same	<10Mb	Unknown
RALGAPB--RBBP9	chr20	chr20	Same	>=10Mb	Yes
RN7SKP71--RN7SKP48	chr12	chr4	Different	>=10Mb	No
SLC7A5--SMG1P1	chr16	chr16	Same	>=10Mb	Unlikely
SMC6--AC006030.1	chr2	chr2	Same	>=10Mb	Yes
SMC6--SLC4A5	chr2	chr2	Same	>=10Mb	Yes
SUZ12P1--TP53I13	chr17	chr17	Same	<10Mb	Unknown
TMEM183A--PLEKHA6	chr1	chr1	Same	<10Mb	Unknown
TUBGCP5--TMEM50B	chr15	chr21	Different	>=10Mb	Likely
USP22--RN7SL2	chr17	chr14	Different	>=10Mb	No
USP34--ALMS1	chr2	chr2	Same	>=10Mb	Unlikely

The 'SV' column indicates whether there is evidence of an SV that could generate the gene fusion, based on Hi-C data.

Discussion

Here we identified shared and unique SVs in the related HL-60 and HL-60/S4 cell lines. In HL-60, we identified two SVs that have been reported previously, but not consistently across HL-60 karyotypes [31,39], and one novel insertion. We identified 10 previously unreported SVs in HL-60/S4, that were likely too small to be detected by a recent FISH karyotype [22]. We found that intra-chromosomal SVs were more sensitive to Hi-C method than inter-chromosomal SVs. We identified four novel gene fusions in HL-60/S4, including one that identified a likely translocation that was not identified with Hi-C. We confirmed the conserved composition of the known MYC expansion, and found it inserted into a novel location (4p15.2) in HL-60. We also identified an inversion encompassing this region (4p15.31-p15.1) in HL-60/S4, providing possible evidence of MYC insertion and excision contributing to HL-60 genome evolution.

Hi-C data is predominantly generated with the purpose of investigating the spatial organization of genomes [2,3], but the same data can be repurposed to investigate genome structure without any additional wet-lab work [4–6,43,44]. The complexity of the genomic data that Hi-C generates has many uses and it represents a resource for SV detection.

Previous reports have identified the difference between reciprocal and non-reciprocal translocations based on Hi-C heatmap patterns [5]. In our analysis, we identified that four out of the eleven intrachromosomal SVs in HL-60/S4 cells were not inversions (Table 4), based on the orientation of the butterfly block. In one instance, this was supported by the strandedness of a gene fusion that resulted from the SV (Table 1). In some instances, the complexity of the SVs, such as those seen with chromosome 21q22.11 and 18q21 in HL-60/S4, were difficult to interpret using Hi-C datasets alone. However, these SVs were also in some instances undetected by FISH karyotyping, presumably due to the small size of the rearranged regions. Analyzing Hi-C datasets was therefore able to highlight these for further, targeted analysis.

Table 4. Summary of structural variants in HL-60 and HL-60/S4 cells.

	Inter-chromosomal translocation	Inter-chromosomal other	Intra-chromosomal inversion (>10Mb)	Intra-chromosomal other (>10Mb)
HL-60	6	1	0	1
HL-60/S4	14	1	7	4

A recent study demonstrated that inter-chromosomal SVs could be detected precisely with Hi-C in a mouse erythroid cell line [45] and primary human brain tumors (five glioblastomas and one anaplastic astrocytoma), however, they did not report any intra-chromosomal SVs [5]. An SV detection tool developed by Dixon et al [6] was able to identify both intra- and inter-chromosomal structural variations. However, they did not use the Hi-C data to distinguish different forms of structural variation. Instead, they utilized optical mapping to characterize SVs. While this is preferable, patterns in Hi-C heatmaps can also be used to differentiate inversions from other forms of intra-chromosomal SVs without additional costs.

The automated SV detection tool [6] is highly effective at identifying SVs. However, substantial additional characterization requires visual inspection of the putative SVs. New tools need to be developed to include the orientation and symmetry of SVs and thus determine whether they resemble a reciprocal or non-reciprocal translocation, inversion, or other form of SV.

In theory, copy number variation can be detected with Hi-C [5]. However, the different restriction enzymes (i.e. HindIII, MboI, HaeIII) used in each Hi-C sample analyzed here complicated the analysis of genomic coverage. Despite this, the dramatic copy number increase of a region centered on the MYC gene (8q24) created a Hi-C heatmap pattern, even after matrix balancing to adjust for varied coverage (Fig 1a). The composition of the region (Table 2) was consistent across all three datasets, and with the literature [35]. The MYC expansion was inserted into chromosome 4p15.2 in HL-60 but was involved in SVs with chromosomes 6q22.2 and 11p14.1 in HL-60/S4. The precise breakpoints that we determined with Hi-C allowed us to deduce the probable order of these two SVs (ins(6;8)(q22.2;q24.13-q24.21) and der(6;8;11)(q22.2;q24.13-q24.21;p14.1), Table 1, Fig 1c). While the location of the MYC expansion was stable across two HL-60/S4 samples, it does not appear to be stable across different HL-60 sources, with our analysis showing an insertion into chromosome 4p15.2, while the DSMZ cytogenetic information describes an 8q24 insertion into chromosome 1, a 1990 study found it inserted into a t(5;17) translocation, and a spectral karyotype reported an insertion of chromosome 8 into the p arm of chromosome 11, but no chromosome 6 insertion [31].

We propose that the MYC expansion either; a) was stably inserted into different genomic locations during seed-lot generation at different repositories, or b) is a mobile element that may be inserted and excised, contributing to HL-60 genome evolution. If the former is true, different sources of HL-60 should confirm and report the current location of the MYC expansion, and derivative strains named if their MYC expansion location is different from the canonical strain (i.e. ATCC HL-60). Further experimentation should be performed to ascertain whether the MYC expansion remains able to move around the genome. For instance, HL-60 could be grown for many passages, possibly with the addition of a mutagenic agent to encourage homologous recombination. Samples would be collected at many time points, and then PCR performed with two forward primers aligned to the end of the MYC expansion, and random hexamers as the reverse primers. This would amplify sequences that occur at a junction with the MYC expansion, and the amplified products sequenced. The resulting library would confirm whether the MYC expansion is stably inserted, or whether it is mobile and warrants further investigation.

The MYC extra-chromosomal element is particularly well described in HL-60 cells, but MYC (and MYC homolog) extra-chromosomal amplifications have been identified in a number of different cancers, most commonly leukemias and glioblastomas [46–53]. MYC deregulation in cancer generally increases the expression levels of the MYC gene, maintaining cancer cells in a highly proliferative state [54–57]. Intriguingly, overexpression of MYC is sufficient to drive genomic instability and induce the formation of extra-chromosomal elements such as double minutes [54–58], meaning that amplification of MYC can drive further amplification of MYC in a positive feedback loop. If these amplifications can also contribute to genome rearrangements via insertion and excision of extrachromosomal copies, this could provide a second mechanism by which MYC amplifications can contribute to genomic instability, and subsequently tumor development and progression.

The t(10;13) translocation was identified in the HL-60 cell line by Hi-C (Fig 2a) but not the DSMZ cytogenetic information [30]. Unlike the non-canonical deletion on chromosome 5, the t(10;13) translocation involves large regions, and it should be detectable with any form of karyotyping. The t(10;13) translocation has been previously reported [39], but its absence from HL-60/S4 datasets ([25], this study) indicates it was not in the original leukemia sample. Therefore, it is likely that the t(10;13) translocation is absent from other HL-60 stocks. In contrast, the deletion on chromosome 5 is present in all samples, and likely is unreported in DSMZ due to its small size.

Hi-C alone is not sufficient to perform a comprehensive analysis of all structural variations within a cell line. Rather, comprehensive characterization of SVs requires validation and investigation from orthogonal techniques [6]. Despite this, Hi-C can provide substantial information regarding the location and type of SVs. Moreover, Hi-C and RNA-seq libraries are often generated together in gene regulation studies (e.g. [25,59,60]). As such, integrating automated detection tools, e.g. hic_breakfinder [6] and STAR-Fusion [27], with visual heatmap inspection can provide substantial information regarding genomic SVs without any additional cost.

Conclusion

HL-60/S4 cells have a stable karyotype, with a number of small SVs that are not detected with FISH-karyotyping. HL-60 cells from the China Infrastructure of Cell Line Resources have two large SVs that are distinct from the DSMZ karyotype (t(10;13)(p12.1;q12.12), ins(4;8)(q35.2;q24.13-q24.21)), and a small deletion del(q23.3-q31.3) that is unlikely to be detected with traditional karyotype approaches. The MYC expansion has a stable composition but is found in a different genomic location in each HL-60 cytogenetic analysis, indicating it may act as a mobile element. Hic_breakfinder was effective at identifying inter-chromosomal translocations and intra-chromosomal SVs above 10Mb apart. Visual inspection of SV patterns in the Hi-C heatmap distinguished inversions from other forms of intra-chromosomal SV, which was supported by the gene fusions identified with STAR-Fusion. Hi-C and RNA-seq are effective tools to characterize the location and type of genomic SV.

Methods

Cell culture – HL-60/S4-HindIII

HL-60/S4 cells were maintained in RPMI 1640 medium, plus 10% fetal calf serum and 1% penicillin, streptomycin, and l-glutamine at 37°C in a humid incubator purged with 5% CO2/95% air. Differentiation of HL-60/S4 cells into granulocytes was induced by all-trans retinoic acid (ATRA) treatment for four days [25]. Briefly, cells were seeded at a density of 1x105 cells/ml in the RPMI medium (above) and ATRA (1μM final concentration).

Hi-C – HL-60/S4-HindIII

Dilution Hi-C library preparation was adapted from Lieberman-Aiden et al. 2009 [3]. Briefly, 15x107 cells were crosslinked with Formaldehyde (2% FC, 10 min, RT), quenched with Glycine (125mM, 5 min, RT) and scraped from the plate. Cells were lysed, digested (HindIII) and ligated according to Lieberman-Aiden et al. 2009 [3]. A single replicate was sequenced (Illumina Hi-Seq 2000, DKFZ, Heidelberg, Germany).

Hi-C data processing

Reads were aligned to hg38 and filtered using HiCUP v0.5.9 [61]. SVs were detected with hic_breakfinder [6]. SVs were filtered and summarized in R [62] with dplyr [63]. Iterative correction and eigenvector decomposition (ICE) was performed with cooler v0.7.1 [64]. Hi-C heatmaps were generated with HiGlass v1.1.5 [26].

RNA-seq analysis

Reads were aligned and fusions detected with STAR-Fusion [27,65] with hg38 and gencode annotations v27. Fusions were filtered and summarized in R [62] with dplyr [63].

Availability of data and materials

The HL-60/S4-HindIII Hi-C raw and processed data is available at GEO, GSE120815. The HL-60/S4-MboI Hi-C from [25] is publicly available on GEO, GSE115634. The HL-60/S4-mRNA from [22] is publicly available in the NCBI Short Read Archive at http://www.ncbi.nlm.nih.gov/bioproject/303179. The HL-60-HaeIII Hi-C and HL-60 RNA from [24] is publicly available on GEO, GSE93997. The HL-60/S4-total RNA is publicly available on GEO, GSE120579.

Additional files

Supplementary Figure 1. Heatmaps of previously unreported intra-chromosomal SVs in HL-60/S4 cells. a)i) An inversion on chromosome 4 is present in HL-60/S4, but not HL-60 cells. ii) An inversion on chromosome 17 is present in HL-60/S4, but not HL-60 cells. iii) An inversion on chromosome 20 is present in HL-60/S4, but not HL-60 cells. b)i) An SV on chromosome 18 is present in HL-60/S4, but not HL-60 cells. However it lacks the characteristic inversion pattern, and appears to be related to a larger scale SV and likely deletion.

Supplementary Table 1. Hic-breakfinder filtered results for all samples. This table shows all breakpoints over 10Mb apart. The log-odds score indicates the strength of the call. The bias value predicts which coordinate is closest to the breakpoint. A "+" bias indicates the "end" coordinate is closest to the breakpoint, while a "-" bias indicates the "start" coordinate is closest to the breakpoint. The resolution indicates the maximum resolution the breakpoint was identified at, as hic_breakfinder identifies breakpoints at multiple resolutions. The sample column indicates which dataset the breakpoint was identified in.

Supplementary Table 2. STAR-Fusion filtered results for all samples. This table has the gene fusion predictions for all samples of the 17 fusions identified in all replicates of at least one condition. Each row provides the STAR-Fusion results for a single RNA-seq library, thus each gene fusion may fill up to 29 rows. The first three columns indicate the dataset, condition, and replicate of the library. The remaining columns include detailed information about the predicted fusion as outputted by STAR-Fusion, including specific breakpoint locations and the number of junction and spanning reads.

Supplementary Table 3. HiCUP QC results for HL-60/S4-HindIII.

Supplementary methods. Further details of the HL-60/S4-HindIII Hi-C library preparation.

Declarations

Ethics approval and consent to participate

Not Applicable

Consent for publication

Not Applicable

Competing interests

The authors declare that they have no competing interests

Funding

This research was supported by a Health Research Council Explorer grant (HRC 15/604) to JMO and a University of Auckland FRDF grant (3702119) to JMO. The funding bodies had no role in the study design, collection, analysis, and interpretation of the data, or preparing the manuscript.

Authors' contributions

ALO and DEO grew HL-60/S4 cells and RSG prepared the HindIII Hi-C library, under the guidance of JMO. ECJ analyzed and interpreted the data, and wrote the manuscript with support from JMO, JKP, and MHV. All authors contributed to revisions and approved the final manuscript.

References

1 Hasty P, Montagna C. Chromosomal rearrangements in cancer: Detection and potential causal mechanisms. Mol Cell Oncol 2014;1.
2 Rao SSP, Huntley MH, Durand NC, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 2014;159:1665–1680.
3 Lieberman-Aiden E, van Berkum NL, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science (80-) 2009;326:289–293.
4 Dudchenko O, Batra SS, Omer AD, et al. De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. Science (80-) 2017;356:92–95.
5 Harewood L, Kishore K, Eldridge MD, et al. Hi-C as a tool for precise detection and characterisation of chromosomal rearrangements and copy number variation in human tumours. Genome Biol 2017;18:125.
6 Dixon JR, Xu J, Dileep V, et al. Integrative detection and analysis of structural variation in cancer genomes. Nat Genet 2018:1.
7 Gallagher R, Collins S, Trujillo J, et al. Characterization of the continuous, differentiating myeloid cell line (HL-60) from a patient with acute promyelocytic leukemia. Blood 1979;54:713–733.
8 Breitman TR, Selonick SE, Collins SJ. Induction of differentiation of the human promyelocytic leukemia cell line (HL-60) by retinoic acid. Proc Natl Acad Sci 1980;77:2936–2940.
9 Schenk T, Stengel S, Zelent A. Unlocking the potential of retinoic acid in anticancer therapy. Br J Cancer 2014;111:2039–2045.
10 Birnie GD. The HL60 cell line: a model system for studying human myeloid cell differentiation. Br J Cancer Suppl 1988;9:41–45.
11 Collins SJ, Robertson KA, Mueller L. Retinoic acid-induced granulocytic differentiation of HL-60 myeloid leukemia cells is mediated directly through the retinoic acid receptor (RAR-alpha). Mol Cell Biol 1990;10:2154–2163.
12 Millius A, Weiner OD. Manipulation of neutrophil-like HL-60 cells for the study of directed cell migration. Methods Mol Biol 2010;591:229–242.
13 Hauert AB, Martinelli S, Marone C, et al. Differentiated HL-60 cells are a valid model system for the analysis of human neutrophil migration and chemotaxis. Int J Biochem Cell Biol 2002;34:838–854.
14 Bréchard S, Bueb J-L, Tschirhart EJ. Interleukin-8 primes oxidative burst in neutrophil-like HL-60 through changes in cytosolic calcium. Cell Calcium 2005;37:531–540.

15 Levy R, Rotrosen D, Nagauker O, et al. Induction of the respiratory burst in HL-60 cells. Correlation of function and protein expression. J Immunol 1990;145:2595–2601.
16 Rowat AC, Jaalouk DE, Zwerger M, et al. Nuclear envelope composition determines the ability of neutrophil-type cells to passage through micron-scale constrictions. J Biol Chem 2013;288:8610–8618.
17 Olins AL, Herrmann H, Lichter P, et al. Nuclear envelope and chromatin compositional differences comparing undifferentiated and retinoic acid- and phorbol ester-treated HL-60 cells. Exp Cell Res 2001;268:115–127.
18 Paietta E, Stockert RJ, Calvelli T, et al. HL-T, a new cell line derived from HL-60 promyelocytic leukemia cell cultures expressing terminal transferase and secreting suppressor activity. Blood 1987;70:1151–1160.
19 Harker WG, Slade DL, Dalton WS, et al. Multidrug resistance in mitoxantrone-selected HL-60 leukemia cells in the absence of P-glycoprotein overexpression. Cancer Res 1989;49:4542–4549.
20 Leung MF, Sokoloski JA, Sartorelli AC. Changes in microtubules, microtubule-associated proteins, and intermediate filaments during the differentiation of HL-60 leukemia cells. Cancer Res 1992;52:949–954.
21 Campbell MS, Lovell MA, Gorbsky GJ. Stability of nuclear segments in human neutrophils and evidence against a role for microfilaments or microtubules in their genesis during differentiation of HL60 myelocytes. J Leukoc Biol 1995;58:659–666.
22 Mark Welch DB, Jauch A, Langowski J, et al. Transcriptomes reflect the phenotypes of undifferentiated, granulocyte and macrophage forms of HL-60/S4 cells. Nucleus 2017;8:222–237.
23 Ben-David U, Siranosian B, Ha G, et al. Genetic and transcriptional evolution alters cancer cell line drug response. Nature 2018;560:325–330.
24 Li Y, He Y, Liang Z, et al. Alterations of specific chromatin conformation affect ATRA-induced leukemia cell differentiation. Cell Death Dis 2018;9:200.
25 Jacobson EC, Perry JK, Long DS, et al. Migration through a small pore disrupts inactive chromatin organization in neutrophil-like cells. BMC Biol 2018;16:142.
26 Kerpedjiev P, Abdennur N, Lekschas F, et al. HiGlass: web-based visual exploration and analysis of genome interaction maps. Genome Biol 2018;19:125.
27 Haas B, Dobin A, Stransky N, et al. STAR-Fusion: Fast and Accurate Fusion Transcript Detection from RNA-Seq. BioRxiv 2017:120295.
28 Jacobson EC, Perry JK, Vickers MH, et al. TNF-α differentially regulates cell cycle genes in promyelocytic and granulocytic HL-60/S4 cells. BioRxiv 2018:471789.
29 Klijn C, Durinck S, Stawiski EW, et al. A comprehensive transcriptional portrait of human cancer cell lines. Nat Biotechnol 2015;33:306–312.
30 Deutsche Sammlung von Mikroorganismen und Zellkulturen. DSMZ catalogue: ACC 3.
31 Liang JC, Ning Y, Wang R yu, et al. Spectral karyotypic study of the HL-60 cell line: Detection of complex rearrangements involving chromosomes 5, 7, and 16 and delineation of critical region of deletion on 5q31.1. Cancer Genet Cytogenet 1999;113:105–109.
32 Von Hoff DD, Forseth B, Clare CN, et al. Double minutes arise from circular extrachromosomal DNA intermediates which integrate into chromosomal sites in human HL-60 leukemia cells. J Clin Invest 1990;85:1887–1895.
33 Lajoie BR, Dekker J, Kaplan N. The Hitchhiker's guide to Hi-C analysis: Practical guidelines. Methods 2015;72:65–75.
34 Von Hoff DD, Needham-VanDevanter DR, Yucel J, et al. Amplified human MYC oncogenes localized to replicating submicroscopic circular DNA molecules. Proc Natl Acad Sci 1988;85:4804–4808.
35 Hirano T, Ike F, Murata T, et al. Genes encoded within 8q24 on the amplicon of a large extrachromosomal element are selectively repressed during the terminal differentiation of

HL-60 cells. Mutat Res - Fundam Mol Mech Mutagen 2008;640:97–106.
36 Misawa S, Staal SP, Testa JR. Amplification of the c-myc oncogene is associated with an abnormally banded region on chromosome 8 or double minute chromosomes in two HL-60 human leukemia sublines. Cancer Genet Cytogenet 1987;28:127–135.
37 Eckhardt SG, Dai A, Davidson KK, et al. Induction of differentiation in HL60 cells by the reduction of extrachromosomally amplified c-myc. Proc Natl Acad Sci 1994;91:6674–6678.
38 ENCODE Project Consortium TEP. The ENCODE (ENCyclopedia Of DNA Elements) project. Science (80-) 2004;306:636–640.
39 Au WW, Callaham MF, Workman ML, et al. Double minute chromatin bodies and other chromosome alterations in human myeloid HL-60 leukemia cells susceptible or resistant to induction of differentiation by phorbol-12-myristate-13-acetate. Cancer Res 1983;43:5873–5878.
40 Balajee AS, Hande MP. History and evolution of cytogenetic techniques: Current and future applications in basic and clinical research. Mutat Res Toxicol Environ Mutagen 2018;836:3–12.
41 Mendes-da-Silva P, Moreira A, Duro-da-Costa J, et al. Frequent loss of heterozygosity on chromosome 5 in non-small cell lung carcinoma. Mol Pathol 2000;53:184–187.
42 Lei Q, Li C, Zuo Z, et al. Evolutionary insights into RNA trans-splicing in vertebrates. Genome Biol Evol 2016;8:562–577.
43 Burton JN, Adey A, Patwardhan RP, et al. Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. Nat Biotechnol 2013;31:1119–1125.
44 Rickman DS, Soong TD, Moss B, et al. Oncogene-mediated alterations in chromatin conformation. Proc Natl Acad Sci 2012;109:9083–9088.
45 Coghill E, Eccleston S, Fox V, et al. Erythroid Kruppel-like factor (EKLF) coordinates erythroid cell proliferation and hemoglobinization in cell lines derived from EKLF null mice. Blood 2001;97:1861–1868.
46 Storlazzi CT, Fioretos T, Surace C, et al. MYC-containing double minutes in hematologic malignancies: Evidence in favor of the episome model and exclusion of MYC as the target gene. Hum Mol Genet 2006;15:933–942.
47 Alseraye F, Padmore R, Wozniak M, et al. MYC gene amplification in double minute chromosomes in an aggressive large B-cell lymphoma with leukemic presentation: a case report. Cancer Genet Cytogenet 2009;192:76–78.
48 Wong KF, Siu LLP, Wong WS. Double minutes and MYC amplification: A combined May-Grünwald Giemsa and fluorescence in situ hybridization study. Am J Clin Pathol 2014;141:280–284.
49 Poddighe PJ, Wessels H, Merle P, et al. Genomic amplification of MYC as double minutes in a patient with APL-like leukemia. Mol Cytogenet 2014;7:67.
50 Jin Y, Liu Z, Cao W, et al. Novel functional MAR elements of double minute chromosomes in human ovarian cells capable of enhancing gene expression. PLoS One 2012;7:e30419.
51 Bigner SH, Mark J, Bigner DD. Cytogenetics of human brain tumors. Cancer Genet Cytogenet 1990;47:141–154.
52 Petit T, Davidson K, Izbicka E, et al. Elimination of extrachromosomal c-myc genes by hydroxyurea induces apoptosis. Apoptosis 1999;4:163–167.
53 Sanborn JZ, Salama SR, Grifford M, et al. Double minute chromosomes in glioblastoma multiforme are revealed by precise reconstruction of oncogenic amplicons. Cancer Res 2013;73:6036–6045.
54 Smith G, Taylor-Kashton C, Dushnicky L, et al. c-Myc-induced extrachromosomal elements carry active chromatin. Neoplasia 2003;5:110–120.
55 Wiener F, Kuschak TI, Ohno S, et al. Deregulated expression of c-Myc in a translocation-negative plasmacytoma on extrachromosomal elements that carry IgH and myc genes. Proc Natl Acad Sci U S A 1999;96:13967–13972.

56 Kalkat M, De Melo J, Hickman KA, et al. MYC deregulation in primary human cancers. Genes (Basel) 2017;8:2–30.
57 Dang C V. MYC on the path to cancer. Cell 2012;149:22–35.
58 Kuzyk A, Mai S. c-MYC-induced genomic instability. Cold Spring Harb Perspect Med 2014;4:a014373.
59 Nora EP, Goloborodko A, Valton AL, et al. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. Cell 2017;169:930–944.e22.
60 D'Ippolito AM, McDowell IC, Barrera A, et al. Pre-established chromatin interactions mediate the genomic response to glucocorticoids. Cell Syst 2018;7:146–160.e7.
61 Wingett S, Ewels P, Furlan-Magaril M, et al. HiCUP: pipeline for mapping and processing Hi-C data. F1000Research 2015;4:1310.
62 Team RC. R: A language and environment for statistical computing. 2018.
63 Wikham H, François R, Henry L, et al. dplyr: A Grammar of Data Manipulation 2018.
64 Kerpedjiev P, Fudenberg G, Oullette S, et al. mirnylab/cooler 2018.
65 Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.