Published in final edited form as:

Radiat Res. 2018 July ; 190(1): 88–97. doi:10.1667/RR15053.1.

Molecular Cytogenetics Guides Massively Parallel Sequencing of a Radiation-Induced Chromosome Translocation in Human Cells

Michael N. Cornforth^{a,c,1}, Pavana Anur^b, Nicholas Wang^b, Erin Robinson^c, F. Andrew Ray^{c,d}, Joel S. Bedford^{c,d}, Bradford D. Loucas^a, Eli S. Williams^e, Myron Peto^b, Paul Spellman^b, Rahul Kollipara^f, Ralf Kittler^f, Joe W. Gray^b, and Susan M. Bailey^{c,d}

^aDepartment of Radiation Oncology, University of Texas Medical Branch, Galveston, Texas 77555

^bDepartments of Molecular and Medical Genetics, Biomedical Engineering, Oregon Health and Science University, Portland, Oregon 97201

^cKromaTiD Inc., Fort Collins, Colorado 80523

^dDepartment of Environmental and Radiological Health Sciences, Colorado State University, Fort Collins, Colorado 80523

^eDepartment of Pathology, University of Virginia School of Medicine, Charlottesville, Virginia 22908

^fMcDermott Center, University of Texas Southwestern Medical Center, Dallas, Texas 75235

Abstract

Chromosome rearrangements are large-scale structural variants that are recognized drivers of oncogenic events in cancers of all types. Cytogenetics allows for their rapid, genome-wide detection, but does not provide gene-level resolution. Massively parallel sequencing (MPS) promises DNA sequence-level characterization of the specific breakpoints involved, but is strongly influenced by bioinformatics filters that affect detection efficiency. We sought to characterize the breakpoint junctions of chromosomal translocations and inversions in the clonal derivatives of human cells exposed to ionizing radiation. Here, we describe the first successful use of DNA paired-end analysis to locate and sequence across the breakpoint junctions of a radiation-induced reciprocal translocation. The analyses employed, with varying degrees of success, several well-known bioinformatics algorithms, a task made difficult by the involvement of repetitive DNA sequences. As for underlying mechanisms, the results of Sanger sequencing suggested that the translocation in question was likely formed via microhomology-mediated non-homologous end joining (mNHEJ). To our knowledge, this represents the first use of MPS to characterize the breakpoint junctions of a radiation-induced chromosomal translocation in human cells. Curiously, these same approaches were unsuccessful when applied to the analysis of inversions previously identified by directional genomic hybridization (dGH). We conclude that molecular cytogenetics

¹Address for correspondence: Department of Radiation Oncology, University of Texas Medical Branch, 301 University Blvd., Galveston, TX 77555-0884; mcomfor@utmb.edu.

continues to provide critical guidance for structural variant discovery, validation and in “tuning” analysis filters to enable robust breakpoint identification at the base pair level.

INTRODUCTION

Large-scale structural rearrangements in DNA, such as chromosomal translocations and inversions, figure prominently in a number of human disease states, especially cancer, and their importance as driver aberrations in leukemias is well established (1, 2). The cells of virtually all cancers harbor chromosomal rearrangements (3, 4). In some cases these are known to be the underlying causative event; in others, they are a major source of carcinogenic genomic instability. All agents capable of producing chromosomal aberrations are carcinogens, and conversely, virtually all carcinogens are also clastogens (5). In addition, because the analysis of structural chromosomal aberrations provides precise quantifiable data on a cell-by-cell basis, it has become the “gold standard” for use in radiation biodosimetry (6–8).

The large majority of chromosomal aberrations caused by ionizing radiation are exchange-type aberrations. Asymmetrical exchange products include dicentrics, microscopically visible interstitial deletions and ring chromosomes. Rings and larger interstitial deletions transmit poorly to cell progeny because they are likely to contain vital genetic information necessary for cell survival (9, 10). Consequently, they would play a role in radiotherapy. Although smaller interstitial deletions transmit with higher frequencies, of primary concern to carcinogenesis and mutagenesis are the transmissible symmetrical products of misrepair: chromosomal translocations and inversions.

Radiation-induced chromosomal aberration formation has been a focus of study for several decades (11–13). From a biophysics standpoint, quite a bit is known about the relationship between radiation exposure and cytogenetic damage, including the effects of absorbed dose, dose rate, ionization density and (more broadly speaking) energy deposition/track structure (14). By comparison, it is only relatively recently that information has begun to emerge about chromosomal aberration formation and underlying DNA repair/misrepair processes at the molecular level.

Most translocations (e.g., in leukemias) were first discovered utilizing classical cytogenetic methods, such as chromosome banding applied to metaphase chromosomes. Molecular cytogenetic approaches have advanced extensively to embrace more sensitive and specific hybridization strategies for detecting structural rearrangements. These include the combinatorial whole-chromosome painting methods of multicolor fluorescence *in situ* hybridization (mFISH) (15) and spectral karyotyping (SKY) (16) aimed at detecting rearrangements between different chromosomes (e.g., translocations) and, more recently, techniques like directional genomic hybridization (dGH) (17) designed to detect intrachromosomal rearrangements (i.e., inversions). These approaches complement other methods of genomic analyses, such as comparative genomic hybridization (CGH) and gene expression profiling that have facilitated discovery of translocations like the fusion of TMPRSS2 with ETS family transcription factors in prostate cancer (18). Together, these

techniques are revealing a growing number of structural variants (SVs) that would benefit from analysis at the DNA sequence level.

Two broad avenues of research have provided molecular insight into the formation of radiation-induced chromosome exchanges at the sequence level. In the first approach, DNA double-strand break (DSB)-inducing activities of rare-cutting nucleases are used to mimic the effects of radiation-induced damage. An example is the pivotal work of Richardson and Jasin (19, 20), who inserted exogenous recognition sites of I-SCE1 into the mammalian genome, and then used I-SCE1 to produce DSBs at those specific sites. The inserted sequences were constructed to allow selection of recombinants that restored function to an otherwise inactive gene, thereby allowing for its selection among a population of treated cells. Their results supported the notion that DSBs are the principle initiating substrates for the formation of chromosomal translocations, which form via (mis)repair process involving non-homologous end joining [NHEJ; e.g. (21)].

The above view is consistent with long-standing cytogenetic dogma that it takes a DSB on each of the chromosomes involved to form an exchange-type aberration (22). Important results from this work also include identification of a prominent role of the multifunction resection endonuclease CtIP (also known as RBBP8) in the formation of chromosomal translocations (23–32), which interacts with the MRE11-RAD50-NBS (MRN) complex in processing of DSBs. CtIP plays a role in homologous recombination repair, where it shares with NHEJ the same initial step in the processing of DSBs (33). CtIP-mediated NHEJ is thought to represent an alternative (alt) pathway that differs from the more common canonical form of NHEJ (cNHEJ), whereby it is not dependent on Ku70 (34). Whereas both alt-NHEJ and cNHEJ processes typically rely on short stretches of homology (a few base pairs) at recombinational junctions, the alt-NHEJ pathway tends to result in somewhat longer microhomologies at nascent breakpoint junctions. The majority of translocation junctions were formed via end joining with short microhomologies. These and other studies support the view that alt-NHEJ is the primary mediator of translocation formation in mammalian cells (35–38).

A second experimental approach is exemplified by seminal work conducted by Thacker and colleagues (39, 40), who studied intragenic breakpoints associated with deletion mutations in the native human HPRT locus. Rather than using enzymatic cutting to produce DSBs, they exposed primary human fibroblasts to 3.3 MeV alpha particles. They found that approximately one-half of the radiation-induced mutants recovered had suffered large intrachromosomal deletions of part or all of the HPRT gene and reasoned that some portion of the latter were probably large enough to be seen under the microscope (i.e., as interstitial deletions). One-half of these had at least one breakpoint within the HPRT gene sequence located on the X chromosome. One of the mutants recovered had a breakpoint within HPRT, and the other breakpoint on chromosome 9, forming a translocation that was visible by mFISH. Sequencing across the resulting breakpoint junctions led to the following observations. Breakpoint junctions of large intrachromosomal deletions typically displayed short microhomologies (2–5 base pairs) characteristic of underlying NHEJ processes. Notably, the breakpoints themselves frequently occurred within repetitive DNA [e.g., *Arthrobacter luteus* (Alu), short and long interspersed nuclear elements (SINE and LINE,

respectively)]. While these approaches have brought undeniable power and clarity to bear on the study of radiation-induced chromosomal aberrations, neither approach is immune from critique.

Regarding experimental approaches that employ enzymatically-induced DSBs, it is debatable whether these can ever fully mimic the complexity, molecular diversity and local lesion density associated with radiation exposure (41, 42). Moreover, for a given target construct, the enzyme recognition sequences flanking the cut site, as well as their integration placement in the genome are fixed, as opposed to the presumably random distribution of initial lesions produced by radiation. This raises the possibility that the selection of enzymatic target sequence and its placement in the genome may affect results.

It should also be mentioned that some experimental approaches rely on strong selective pressure for cells to survive in the presence of otherwise-lethal agents (e.g., 6TG or G418). As a result, the frequencies of analyzable cells can be orders of magnitude lower compared to the recovery of transmissible aberrations in cells without such selection pressure applied. Consequently, unintentional selection bias is always a possibility.

Here, we introduce an experimental model system ostensibly unaffected by these issues. Our approach employed paired-end MPS of DNA associated with the breakpoints of chromosomal translocations and inversions that were first discovered using mFISH and dGH, respectively. In principle, MPS should be sufficiently robust for the analysis of large-scale SVs, and several computational algorithms have been designed for this purpose (43–46). These are generally based on detection of genomic SVs from split and/or discordant read pairs in paired-end sequencing data. However, the sensitivity and specificity of such sequencing-based methods for detection of large-scale genomic rearrangements like translocations and inversions is not yet fully established.

Using a suite of paired-end sequencing analysis tools (47), we describe our attempts to detect and characterize chromosomal aberration/SV breakpoints in clonal populations selected from hTERT-immortalized fibroblasts whose progenitor cells were irradiated with either ^{137}Cs gamma rays or high-energy ^{56}Fe ions. Clones from irradiated populations were expanded from single cells and assessed for translocations and/or inversions using mFISH and dGH. We detail our experiences involving the attempted analysis of five clones carrying SVs and describe strategies used to successfully sequence across the breakpoint junctions of a (3:4) translocation in one of these five clones.

MATERIALS AND METHODS

Cytogenetics

Cell culture and irradiation—These were conducted jointly at the University of Texas Medical Branch (UTMB, Galveston, TX) and Colorado State University (CSU, Fort Collins, CO). Early on in the development of our experimental system, BJ-1 hTERT-immortalized human fibroblasts (N.I.A cell repository; Coriell Institute, Camden, NJ) were first subcloned. These were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum, penicillin and streptomycin, and maintained at

37°C in a humidified atmosphere of 5% CO₂ and air (48). Clone K1 was chosen among these on the basis of its high plating efficiency, robust growth and karyotypic stability. They represent the parent culture from which all irradiated clones were derived.

Confluent K-1 cultures were gamma irradiated with either 1.0 or 4.0 Gy (¹³⁷Cs source, 662 MeV Mark I irradiator; JL Shepherd & Associates, San Fernando, CA) at a dose rate of 1.3 Gy/min at UTMB. Cultures exposed to 1.0 GeV/amu ⁵⁶Fe ions were irradiated with 1.0 Gy at NASA's Space Radiation Laboratory (NSRL, Brookhaven National Laboratory, Upton, NY). After a 12-h postirradiation recovery interval at 37°C, cells were released from the density-inhibited state by trypsinization and plated at low density for clonal expansion. Individual clones derived from the growth of single cells were isolated using cloning cylinders without intentional regard to colony size or morphology. (For subsequent studies, aliquots were cryopreserved in liquid nitrogen for future use.) Colcemid (0.10 µg/ml final concentration) was added to individually-expanded cultures (clones) 4 h prior to hypotonic treatment and fixation in 3:1 methanol:glacial acetic acid.

mFISH hybridization and image collection—Previously fixed cells were spread onto glass microscope slides by standard cytogenetic methods. Slides were then treated with acetone, RNase A and proteinase K before another fixation in 3.7% formaldehyde. Slides were dehydrated through an ethanol series (70, 85 and 100%) and air dried. They were then incubated in 72°C formamide (70%) in 2× SSC (0.3 M NaCl, 0.03 M sodium citrate) for 2 min to denature chromosomal DNA. After dehydration through another ethanol series, 10 µl of denatured (10 min at 72°C) SpectraVision 24-color mFISH Assay probe (Vysis; Abbott Molecular, Des Plaines, IL) was applied to each slide. Slides were covered with a 22 × 22-mm glass coverslip sealed into position with rubber cement. Samples were allowed to hybridize for 48 h in a 37°C incubator. After hybridization, coverslips were removed and the slides were washed for 2 min in 0.4× SSC containing IGEPAL® (0.3%) nonionic detergent at 72°C. This was followed by a 30-s wash in 2× SSC (0.1% IGEPAL) at room temperature. Prior to image capture, 15 µl of DAPI (0.14 µg/ml) dissolved in Vectashield® anti-fade mounting media (Vector® Laboratories, Burlingame, CA) was applied to each slide and covered with a 24 × 40-mm coverslip. For each cell, six fluor-dependent images were captured using a Zeiss AxioPhot epifluorescence microscope (Thornwood, NY) interfaced with a SensSys black-and-white CCD camera. mFISH karyotypes were constructed from the six images using MetaSystems image analysis software (Newton, MA).

Directional genomic hybridization (dGH)—The creation of directional strand-specific probes for human chromosome 3 has been previously described elsewhere (17). Briefly, contiguous DNA sequences were downloaded from the NCBI genomic database (GRCh37.p2 primary assembly), which were masked using Genetic Information Research Institute (GIRI) software (Mountain View, CA) to remove repeat sequences. Single-stranded oligonucleotides were designed to unique sequences using Array Designer version 4.2 (PREMIER Biosoft International, Palo Alto, CA), and later with proprietary software (KromaTiD Inc., Fort Collins, CO), by tiling to small target regions at specified non-overlapping locations along the length of each chromosome 3 contig at ~1-Mb intervals. Oligo design criteria demanded similar directionality and length (~40 mers), as well as

uniform melting temperatures ($70.0^{\circ}\text{C} \pm 5.0^{\circ}\text{C}$). Some 17,000 individual oligos were synthesized (InvitrogenTM, Carlsbad, CA), hydrated and pooled into appropriate subsets for end-labeling with fluorescent dNTP analogs (Cy-3, Fluorescein; General Electric/PerkinElmer[®] Inc., Waltham, MA) using terminal transferase (New England Biolabs[®] Inc., Ipswich, MA). Pools of 90 individually labeled oligos constituted a probe set, which together spanned relatively short unique regions (~5–14.5 kb); probe sets were hybridized individually to confirm strand specificity. The complete chromosome 3-specific chromatid paint consisted of 190 probe sets.

dGH cell culture, slide pretreatment and hybridization—Clonally-derived cell cultures were grown for one round of DNA synthesis in the presence of $5.0\text{ }\mu\text{M}$ BrdU (5-bromo-2'-deoxyuridine) plus $1.0\text{ }\mu\text{M}$ BrdC (5-bromo-2'-deoxycytidine) (17, 49). At CSU, microscope slides containing spread metaphases were incubated in PN buffer for 10 min at room temperature, rinsed in PBS, then dehydrated through an ethanol series (75, 85 and 100%) for 2 min each. Slides were air dried, stained with Hoechst 33258 ($0.5\text{ }\mu\text{g/ml}$ in $2\times$ sodium citrate; SSC) for 15 min in the dark, then rinsed with deionized distilled water (ddH_2O). Slides were air dried, flooded with $2\times$ SSC, coverslipped and exposed to 365 nm ultraviolet light (UV Stratalinker[®] 2400; Stratagene, La Jolla, CA) for 35 min, then rinsed in ddH_2O to remove coverslip, air dried and dehydrated in the ethanol series, as above.

For each pretreated slide, a mixture of hybridization buffer ($25.3\text{ }\mu\text{l}$) and chromatid paint ($2.25\text{ }\mu\text{l}$) (both from KromaTiD Inc.) and ddH_2O ($2.45\text{ }\mu\text{l}$) was prepared and heated to 75°C for 5 min, then pipetted onto pretreated slides, which were coverslipped and sealed with rubber cement. Slides were heated at 73°C for 3 min, then transferred to individual hybridization chambers (Corning[®] Inc., Corning, NY) and incubated at 37°C overnight. After hybridization, the slides were washed five times in $2\times$ SSC at 42°C for 15 min each. Slides were rinsed in PN buffer, counterstained with DAPI/anti-fade ($18\text{ }\mu\text{l}$; Vector Laboratories) and coverslipped. Image capture and analysis was performed on an Olympus[®] Bx41 microscope outfitted with fluorochrome-appropriate excitation/barrier filters, running MetaVueTM 7.1 software and equipped with a Photometrics Cool-SNAP ES² camera.

Sequencing and Bioinformatics

Whole genome DNA libraries were generated using the Agilent SureSelectXT kit. For each clone, $1\text{ }\mu\text{g}$ of genomic DNA was sheared using a Covaris E220 to a peak target size of 150 base pairs (bp). Fragmented DNA was concentrated using AMPure XP beads (Beckman Coulter[®] Inc., Fullerton, CA), and DNA ends were repaired using T4 DNA polymerase, Klenow polymerase and T4 polynucleotide kinase. A-tailing ($3'$) with exo-minus Klenow polymerase was followed by ligation of Agilent paired-end oligo adapters to genomic DNA fragments. Ligated DNA was polymerase chain reaction (PCR)-amplified for eight cycles, purified using AMPure XP beads and analyzed using an Agilent Bioanalyzer DNA 1000 Sensitivity chip. Qualified libraries were quantitated by qPCR and sequenced on an Illumina HiSeq 2000 (San Diego, CA).

At the University of Texas Southwestern (Dallas, TX) whole genome sequence reads were mapped to the human reference genome (hg19) using Burrows-Wheeler Aligner software

(BWA; version 0.7.12) (50), with default parameters. The mapped reads were sorted using SAMtools (v. 1.2.). Meerkat (46) (v. 0.185), was used to identify SVs that were supported by at least two discordant read pairs and one split read. To filter false positives, split reads sequences were extracted, and their mate pair sequences from fastq files for all SVs called with Meerkat. These split reads and their mate pairs were mapped to the human reference genome (hg19) using the BLAST-like Alignment Tool (BLAT) (v. 35) (51). BLAT results that matched the breakpoints provided by Meerkat were reported. At Oregon Health and Science University (Portland, OR), in addition to aligning whole genome reads to hg19 using BWA (50), duplicate PCR reads were removed using MarkDuplicates from Picard Tools. Realignment around single nucleotide variants and indels and quality score recalibrations were done using the Genome Analysis Toolkit (GATK) software library (52, 53). To call SVs we used Geometric Analysis of Structural Variants (GASV) (v. 1.4) (54), DELLY (versions 0.1.1 and 0.6.7) and BreakDancer (v. 0.0.9) tools (55). DELLY mapping quality (q) flag was set to 30.

RESULTS

BJ-1 hTERT-immortalized fibroblasts were irradiated with either ^{137}Cs gamma rays or ^{56}Fe ions, and metaphases analyzed using 24-color mFISH (15) and dGH (17, 49) to detect chromosomal translocations and inversions, respectively. These cells, which are effectively immortal, were expanded in culture after irradiation to provide clonal material for subsequent molecular analysis. Several of the irradiated clones contained reciprocal translocations and other rearrangements, five of which are discussed here. mFISH karyotypes of the chromosomes in four of these clones are shown in Fig. 1. By comparison, >99% of nonirradiated cells showed a normal euploid 46, XY. In addition to these clones, Fig. 2A shows mFISH and chromosome-3-specific dGH analyses of a fifth clone, K1-400C4 (DNA sample 4A) that contained a reciprocal (3;4s) translocation. Notably, the particular chromosomal translocation(s) found in each of the five clones were homogenous, meaning they were present in every cell within a given clone. As discussed below, K1-400C4 became the main focus of our studies, primarily because it also contained a prominent paracentric inversion involving chromosome 3, as shown in Fig. 2B. A schematic of dGH is provided in Supplementary Fig. S1 (<http://dx.doi.org/10.1667/RR15053.1.S1>).

We analyzed the five clones carrying rearrangements described in Figs. 1 and 2 (and one control clone showing no aberrations) using paired-end sequencing to identify exact breakpoints for each SV. Whole genome libraries were prepared using DNA from each clone. Each DNA sample was sonicated, ligated to Agilent paired-end oligonucleotides and PCR amplified. The amplified fragments were purified and were initially sequenced to ~8× genome coverage on an Illumina HiSeq 2000. We compared reads from the aberration-bearing clones to those from the nonirradiated control clone to identify SVs. We applied several SV detection tools, as these evolved during the course of this effort, which spanned some five years.

We aligned 100 base pair, paired-end sequences to the hg19 human reference genome using BWA (50). We used SAMtools to sort and index the alignments and Mark-Duplicates from Picard Tools (43) to remove duplicate reads generated during the PCR amplification stage.

After duplicate removal, we realigned reads around single nucleotide variants and indels using the GATK Software library (52, 53). We called local positions to target for realignment using RealignerTargetCreator and the reads were realigned using IndelRealigner. Finally, we recalibrated quality scores using GATK BaseRecalibrator and PrintReads, which binned reads based on the original quality score, the dinucleotide and the position in the read. We first used the GASV (v. 1.4) (54), DELLY (v. 0.1.1) (45) and BreakDancer (v 0.0.9) (43) tools to detect breakpoint junctions but were unable to identify the SVs expected from the cytogenetic studies. We then reduced the read length to 50, realigned the sequences and used the same tools to identify SVs. In a parallel effort, we used Meerkat (46) to identify partially mapped reads and pairs in which one of the mates is unmapped. The unmapped portions of the soft clipped reads and cleaved ends of unmapped reads were then remapped to identify candidate SVs.

None of the above efforts identified the breakpoint junctions of the rearrangements that were apparent by mFISH or dGH in any of the five clones. Next, we used an updated DELLY (v. 0.6.7) tool, one of the top scoring tools in the International Cancer Genome Consortium DREAM challenge (ICGC-TCGA) (56) for SV identification. This tool identified many rearrangements and inversions, but nearly all were predicted at low confidence and supported by only two or three reads. This failure to detect aberrations led us to focus subsequent aberration detection efforts on clone K1-400C4 carrying a reciprocal translocation, which was classified as t(3;4)(q26.3;q31) by G-banding (Supplementary Fig. S2; <http://dx.doi.org/10.1667/RR15053.1.S1>). As mentioned, this clone also contained various inversions in chromosome 3 that were readily detected by dGH. Interestingly, but perhaps not unexpectedly (given their small sizes), none of these were visible by conventional G-banding or multi-fluor combinatorial banding (mBAND; not shown) (57, 58).

We resequenced this clone to 30× coverage and again called rearrangements using DELLY (v. 0.6.7). This enabled us to identify a candidate t(3;4) reciprocal translocation, but there were also many other likely false positive breakpoints. We filtered out false positive SVs that occurred from reference mapping artifacts by comparing DELLY calls from clone K1-400C4 to those made in a panel of five normal samples, comprised of the nonirradiated clone 2B (from this study) and four normal samples from ICGC. This reduced the total number of potential translocations identified by DELLY from 2,279 to 31. We further reduced the number of candidate translocations to the six listed in Table 1 by requiring >2 supporting paired-end reads for each SV. This same strategy reduced the candidate inversion count from 845 to 19 (Table 2).

Using the data from whole genome sequencing, the best supported candidate was chosen for validation by Sanger sequencing. PCR primers were designed to amplify both anticipated products of a reciprocal translocation between chromosomes 3 and 4. Sanger sequencing (Fig. 3) of the amplified PCR products revealed a translocation event that resulted in a 1-bp deletion of chromosome 4 and a 6-bp deletion of chromosome 3, with the translocation occurring within a 4-bp (AAGG) overlap between chromosome 3 and 4 (Fig. 4). We took this as confirmation of the t(3;4)(q26.3;q31) translocation detected by G-banding, mFISH and dGH. A BLAT alignment showed that chromosome 3-flanking sequences mapped to a

long terminal repeat (LTR) sequence and that the chromosome 4 sequence mapped within a LINE (Supplementary Figs. S3 and S4 <http://dx.doi.org/10.1667/RR15053.1.S1>). The chromosome 4 mapping within the LINE explains our observation of many other highly homologous matches within this region.

Interestingly, the existence of t(6;9) and t(1;10) translocations received significant bioinformatics support, despite these rearrangements not being visible (i.e., validated) by mFISH or dGH analyses (Table 1). Conversely, we failed to find bioinformatics support for the prominent paracentric inversion observed within chromosome 3 by dGH, even after filtering against SVs found in the panel of normal samples (Table 2).

DISCUSSION

The goal of this study was to determine the extent to which paired-end MPS could be used to determine the sequences across the breakpoints of radiation-induced SVs, specifically chromosomal translocations and inversions present in clones isolated from irradiated cells. Analyses initially focused on five clones in which aberrations were demonstrated by mFISH and dGH to be homogeneously present in essentially all cells. The aberrations in four clones were induced by exposure to ^{137}Cs gamma rays, a radiation of low-linear energy transfer (LET). Aberrations in the fifth clone were induced by exposure to 1 GeV/amu ^{56}Fe ions, a high-LET radiation.

Summarizing our results for the sake of discussion, initial studies were performed using information obtained from paired-end analyses in which the genome coverage achieved during sequencing averaged 8 $\times$. This level of coverage was initially deemed adequate, given the homogeneous nature of chromosomal aberrations within affected cells. Structural variant analyses were performed using a suite of analysis tools at two different institutions: BreakDancer (43), CREST (44), DELLY (45) and Meerkat (46). Over the course of the analysis period, SV discovery tools continued to improve, so several cycles of analysis were performed. Surprisingly, none of the cytogenetically apparent aberrations were found using the collection of SV discovery tools.

We then explored the possibility that failure to detect the known SVs was due to insufficient read depth by resequencing clone K1-400C4 to achieve 30 $\times$ genome coverage. Analyses of these data using a current version of DELLY (v. 0.6.7) suggested the presence of thousands of potential translocations and inversions. These were reduced to a manageable number by comparing candidate SVs to those in normal cells, to eliminate mapping artifacts and polymorphic SVs, and by increasing the number of breakpoint reads required for an SV call. One of the candidate SVs that passed these filters involved breakpoints at 3q26.2 and 4q31.1. Within the resolution associated with breakpoint assignment by G-banding (59), we consider these locations identical to those of the t(3;4)(q26.3;q31) translocation previously identified cytogenetically. This SV was validated by Sanger sequencing of PCR products generated using primers designed to allow amplification across the presumptive breakpoint. These studies showed that the sequences flanking both breakpoints involved elements of LTR and LINE, which almost certainly complicated SV detection. This would seem to

suggest that MPS analysis of larger-insert libraries may help to resolve this problem in future studies.

Massively parallel sequencing analysis failed to identify sequences spanning inversion breakpoints on chromosome 3, despite compelling cytogenetic evidence for their existence provided by dGH. Note, for example in Table 2, that while a number of prospective breakpoint locations received significant bioinformatics support, none were within chromosome 3. It is possible this failure stems from extensive polymorphism within the population, although it seems more likely that this is due to higher levels of rearrangement complexity not apparent at the cytogenetic level. Likewise, we note that clone K1-400H6 (Fig. 1) contained a complex translocation involving three chromosomes and surmise that such complexity may well extend into the submicroscopic level. Alternatively, failure to locate the inversion may be due to the extensive involvement of repetitive sequences at the breakpoint sites. Here again, generation and analysis of larger insert libraries may be necessary to discover and sequence across the breakpoints.

Figure 4 shows that the otherwise reciprocal t(3;4) breakpoint junctions of clone K1-400C4 involved the resection of 1–6 base pairs, as well as a 4-bp overlap of homology. (It is worth noting this result required sequencing across both breakpoints of the reciprocal translocation, and that this level of information is not easily obtainable with other experimental systems.) These results support the conclusions of others who, despite using fundamentally different experimental approaches, concluded that translocation formation is driven by microhomology-mediated/alternative non-homologous end joining (mmNHEJ/alt-NHEJ) processes (20, 34, 39, 40, 60). The finding that both the breakpoints in this translocation occurred within DNA repeats is also noteworthy, but not altogether surprising, given the collective abundance of DNA repeats in the human genome (61, 62). Although it is premature to make generalizations at this early stage, this result is currently explainable on the basis of randomness of radiogenic DSBs serving as substrates for interchromosomal recombination. If so, then we can expect radiation-induced exchange breakpoints within DNA repeat sequences to be commonplace, quite unlike that for the vast majority of other studies using MPS to examine large-scale SVs, where the exchange breakpoint locations are sought in connection with a phenotypic outcome and are therefore more likely to occur in coding regions.

DNA repeats are problematic for all methods of sequencing, but perhaps especially so for paired-end approaches, which employ relatively short sequence reads (< 100 bp). Strategies making use of larger insert libraries (and longer reads) promise to circumvent many of the problems related to repetitive DNA. It is worth contemplating the use of mate-pair sequencing (63) as a prelude to whole genome paired-end sequencing (64) in future studies. Other possibilities include more recently developed approaches, such as single molecule real-time sequencing (65) or linked-read sequencing (66).

Hopefully, these would also alleviate the problems we incurred with inversion analysis, which is important, since radiation-induced inversions are almost certainly more abundant than translocations (49). Additionally, longer reads may aid the analysis of complex exchanges of the type represented by clone K1-400H6 (Fig. 1).

Whereas MPS has been recently applied to the analysis of radiation-induced intrachromosomal SVs (67), we are unaware of studies that have employed our particular approach, which involved MPS applied to the analysis of radiation-induced interchromosomal SVs that were first identified by molecular cytogenetic techniques. We conclude that modern molecular cytogenetic analysis tools like mFISH and dGH provide an important complement to paired-end MPS in SV discovery and analysis. Prior cytogenetic demonstration of structural aberrations/variants allows development of analysis filters to remove false positive SV calls, selection of the proper sequencing read depth and, in some cases, may indicate the need to switch to larger insert libraries to cross repeat-rich regions. As they continue to evolve, MPS-based strategies may well become sufficiently robust for use as a stand-alone approach to SV analysis. For now, molecular cytogenetics continues to provide crucial guidance in SV discovery and validation efforts.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Marco Durante at GSI for mBAND analysis, Dr. Yongjia Yu at UTMB for DNA preparation, and Professor George Iliakis at the University of Essen for thoughtful discussions. This work has received support from the following grants: National Institutes of Health: NIAID R01080486-02 (SMB, MNC); National Aeronautics and Space Administration: NNX15AG74G (MNC), NNX14AB02G (SMB), NNX15AK13G (JSB); The Center for Spatial Systems Biomedicine, Knight Cancer Institute (JWG, PS). MNC, ER, FAR, JSB, JWG and SMB are shareholders in KromaTiD, Inc., of which ER is an employee.

References

- 1Rowley JD. Molecular cytogenetics: Rosetta stone for understanding cancer—Twenty-Ninth G. H. A. Clowes Memorial Award Lecture. *Cancer Res.* 1990; 50:3816–25. [PubMed: 2191766]
- 2Rowley JD. Chromosome abnormalities in leukemia. *J Clin Oncol.* 1988; 6:194–202. [PubMed: 3422259]
- 3Mitelman F. *Catalog of Chromosome Aberrations in Cancer* 4. New York: Wiley-Liss; 1991
- 4Solomon E, Borrow J, Goddard AD. Chromosome aberrations and cancer. *Science.* 1991; 254:1153–60. [PubMed: 1957167]
- 5Cornforth MN. DNA Damage and repair Nickoloff JA, , Hoekstra M, editorsTotowa, NJ: Humana Press; 1998 55985
- 6Durante M, Bonassi S, George K, Cucinotta FA. Risk estimation based on chromosomal aberrations induced by radiation. *Radiat Res.* 2001; 156:662–7. [PubMed: 11604089]
- 7Edwards AA. The use of chromosomal aberrations in human lymphocytes for biological dosimetry. *Radiat Res.* 1997; 148:S39–44. [PubMed: 9355855]
- 8Bauchinger M, Braselmann H, Savage JR, Natarajan AT, Terzoudi GI, Pantelias GE, et al. Collaborative exercise on the use of FISH chromosome painting for retrospective biodosimetry of Mayak nuclear-industrial personnel. *Int J Radiat Biol.* 2001; 77:259–67. [PubMed: 11258840]
- 9Cornforth MN, Bedford JS. A quantitative comparison of potentially lethal damage repair and the rejoining of interphase chromosome breaks in low passage normal human fibroblasts. *Radiat Res.* 1987; 111:385–405. [PubMed: 3659275]
- 10Revell SH. Radiation-induced chromosome damage in man Ishihara T, , Sasaki MS, editorsNew York: Liss; 1983 215233
- 11Catcheside DG, Lea DE, Thoday JM. Types of chromosome structural change induced by the irradiation of *Tradescantia* microspores. *J Genetics.* 1946; 47:113–36. [PubMed: 21010995]
- 12Lea DE. *Actions of radiations on living cells* London: Cambridge University Press; 1946

- 13Lea DE, Catcheside DG. The mechanism of the induction by radiation of chromosome aberrations in *Tradescantia*. *J Genetics*. 1942; 44:216–45.
- 14Levy D, Vazquez M, Cornforth M, Loucas B, Sachs RK, Arsuaga J. Comparing DNA damage-processing pathways by computer analysis of chromosome painting data. *J Comput Biol*. 2004; 11:626–41. [PubMed: 15579235]
- 15Speicher MR, Gwyn Ballard S, Ward DC. Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat Genet*. 1996; 12:368–75. [PubMed: 8630489]
- 16Schrock E, du Manoir S, Veldman T, Schoell B, Wienberg J, Ferguson-Smith MA, et al. Multicolor spectral karyotyping of human chromosomes. *Science*. 1996; 273:494–7. [PubMed: 8662537]
- 17Ray FA, Zimmerman E, Robinson B, Cornforth MN, Bedford JS, Goodwin EH, et al. Directional genomic hybridization for chromosomal inversion discovery and detection. *Chromosome Res*. 2013; 21:165–74. [PubMed: 23572395]
- 18Tomlins SA, Laxman B, Varambally S, Cao X, Yu J, Helgeson BE, et al. Role of the TMPRSS2-ERG gene fusion in prostate cancer. *Neoplasia (New York)*. 2008; 10:177–88.
- 19Richardson C, Jasin M. Recombination between two chromosomes: implications for genomic integrity in mammalian cells. *Cold Spring Harb Symp Quant Biol*. 2000; 65:553–60. [PubMed: 12760073]
- 20Richardson C, Jasin M. Frequent chromosomal translocations induced by DNA double-strand breaks. *Nature*. 2000; 405:697–700. [PubMed: 10864328]
- 21Godwin AR, Bollag RJ, Christie DM, Liskay RM. Spontaneous and restriction enzyme-induced chromosomal recombination in mammalian cells. *Proc Natl Acad Sci U S A*. 1994; 91:12554–8. [PubMed: 7809076]
- 22Cornforth MN. Perspectives on the formation of radiation-induced exchange aberrations. *DNA Repair*. 2006; 5:1182–91. [PubMed: 16807139]
- 23Yu X, Baer R. Nuclear localization and cell cycle-specific expression of CtIP, a protein that associates with the BRCA1 tumor suppressor. *J Biol Chem*. 2000; 275:18541–9. [PubMed: 10764811]
- 24Li S, Ting NS, Zheng L, Chen PL, Ziv Y, Shiloh Y, et al. Functional link of BRCA1 and ataxia telangiectasia gene product in DNA damage response. *Nature*. 2000; 406:210–15. [PubMed: 10910365]
- 25Yu X, Chen J. DNA damage-induced cell cycle checkpoint control requires CtIP, a phosphorylation-dependent binding partner of BRCA1 C-terminal domains. *Mol Cell Biol*. 2004; 24:9478–86. [PubMed: 15485915]
- 26Yu X, Fu S, Lai M, Baer R, Chen J. BRCA1 ubiquitinates its phosphorylation-dependent binding partner CtIP. *Genes Dev*. 2006; 20:1721–6. [PubMed: 16818604]
- 27Sartori AA, Lukas C, Coates J, Mistrik M, Fu S, Bartek J, et al. Human CtIP promotes DNA end resection. *Nature*. 2007; 450:509–14. [PubMed: 17965729]
- 28Huertas P, Jackson SP. Human CtIP mediates cell cycle control of DNA end resection and double strand break repair. *J Biol Chem*. 2009; 284:9558–65. [PubMed: 19202191]
- 29Yuan J, Chen J. N terminus of CtIP is critical for homologous recombination-mediated double-strand break repair. *J Biol Chem*. 2009; 284:31746–52. [PubMed: 19759395]
- 30You Z, Bailis JM. DNA damage and decisions: CtIP coordinates DNA repair and cell cycle checkpoints. *Rends Cell Biol*. 2010; 20:402–9.
- 31You Z, Shi LZ, Zhu Q, Wu P, Zhang YW, Basilio A, et al. CtIP links DNA double-strand break sensing to resection. *Mol Cell*. 2009; 36:954–69. [PubMed: 20064462]
- 32Durante M, Bedford JS, Chen DJ, Conrad S, Cornforth MN, Natarajan AT, et al. From DNA damage to chromosome aberrations: joining the break. *Mutat Res*. 2013; 756:5–13. [PubMed: 23707699]
- 33Truong LN, Li Y, Shi LZ, Hwang PY-H, He J, Wang H, et al. Microhomology-mediated end joining and homologous recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proc Natl Acad Sci U S A*. 2013; 110:7720–5. [PubMed: 23610439]
- 34Weinstock DM, Brunet E, Jasin M. Formation of NHEJ-derived reciprocal chromosomal translocations does not require Ku70. *Nat Cell Biol*. 2007; 9:978–81. [PubMed: 17643113]

- 35McVey M, Lee SE. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet.* 2008; 24:529–38. [PubMed: 18809224]
- 36Kass EM, Jasin M. Collaboration and competition between DNA double-strand break repair pathways. *FEBS Lett.* 2010; 584:3703–8. [PubMed: 20691183]
- 37Simsek D, Jasin M. Alternative end-joining is suppressed by the canonical NHEJ component Xrcc4-ligase IV during chromosomal translocation formation. *Nat Struct Mol Biol.* 2010; 17:410–6. [PubMed: 20208544]
- 38Chiarle R, Zhang Y, Frock RL, Lewis SM, Molinie B, Ho YJ, et al. Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. *Cell.* 2011; 147:107–19. [PubMed: 21962511]
- 39Morris T, Masson W, Singleton B, Thacker J. Analysis of large deletions in the HPRT gene of primary human fibroblasts using the polymerase chain reaction. *Somat Cell Mol Genet.* 1993; 19:9–19. [PubMed: 8460402]
- 40Singleton BK, Griffin CS, Thacker J. Clustered DNA damage leads to complex genetic changes in irradiated human cells. *Cancer Res.* 2002; 62:6263–9. [PubMed: 12414656]
- 41Ward JF. The complexity of DNA damage: relevance to biological consequences. *Int J Radiat Biol.* 1994; 66:427–32. [PubMed: 7983426]
- 42Goodhead DT. The initial physical damage produced by ionizing radiations. *Int J Radiat Biol.* 1989; 56:623–34. [PubMed: 2573657]
- 43Chen K, Wallis JW, McLellan MD, Larson DE, Kalicki JM, Pohl CS, et al. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat Methods.* 2009; 6:677–81. [PubMed: 19668202]
- 44Wang J, Mullighan CG, Easton J, Roberts S, Heatley SL, Ma J, et al. CREST maps somatic structural variation in cancer genomes with base-pair resolution. *Nat Methods.* 2011; 8:652–4. [PubMed: 21666668]
- 45Rausch T, Zichner T, Schlattl A, Stutz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics.* 2012; 28:i333–9. [PubMed: 22962449]
- 46Yang L, Luquette LJ, Gehlenborg N, Xi R, Haseley PS, Hsieh CH, et al. Diverse mechanisms of somatic structural variations in human cancer genomes. *Cell.* 2013; 153:919–29. [PubMed: 23663786]
- 47Baker M. Structural variation: the genome's hidden architecture. *Nat Methods.* 2012; 9:133–7. [PubMed: 22290183]
- 48Loucas BD, Durante M, Bailey SM, Cornforth MN. Chromosome damage in human cells by gamma rays, particles and heavy ions: track interactions in basic dose-response relationships. *Radiat Res.* 2013; 179:9–20. [PubMed: 23198992]
- 49Ray FA, Robinson E, McKenna M, Hada M, George K, Cucinotta F, et al. Directional genomic hybridization: inversions as a potential biodosimeter for retrospective radiation exposure. *Radiat Environ Biophys.* 2014; 53:255–63. [PubMed: 24477407]
- 50Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009; 25:1754–60. [PubMed: 19451168]
- 51Kent WJ. BLAT—the BLAST-like Alignment Tool. *Genome Res.* 2002; 12:656–64. [PubMed: 11932250]
- 52DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genetics.* 2011; 43:491–8. [PubMed: 21478889]
- 53McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010; 20:1297–1303. [PubMed: 20644199]
- 54Sindi SS, Onal S, Peng LC, Wu HT, Raphael BJ. An integrative probabilistic model for identification of structural variation in sequencing data. *Genome Biol.* 2012; 13:R22. [PubMed: 22452995]

- 55Chen K, Wallis JW, McLellan MD, Larson DE, Kalicki JM, Pohl CS, et al. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat Methods*. 2009; 6:677–81. [PubMed: 19668202]
- 56Ewing AD, Houlihan KE, Hu Y, Ellrott K, Caloian C, Yamaguchi TN, et al. Combining tumor genome simulation with crowdsourcing to benchmark somatic single-nucleotide-variant detection. *Nat Methods*. 2015; 12:623–30. [PubMed: 25984700]
- 57Hada M, Wu H, Cucinotta FA. mBAND analysis for high- and low-LET radiation-induced chromosome aberrations: a review. *Mutat Res*. 2011; 711:187–92. [PubMed: 21232544]
- 58Johannes C, Chudoba I, Obe G. Analysis of X-ray-induced aberrations in human chromosome 5 using high-resolution multicolour banding FISH (mBAND). *Chromosome Res*. 1999; 7:625–33. [PubMed: 10628663]
- 59Savage JR. Assignment of aberration breakpoints in banded chromosomes. *Nature*. 1977; 270:513–4. [PubMed: 593369]
- 60Zhang Y, Jasin M. An essential role for CtIP in chromosomal translocation formation through an alternative end-joining pathway. *Nat Struct Mol Biol*. 2011; 18:80–4. [PubMed: 21131978]
- 61de Koning APJ, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet*. 2011; 7:e1002384. [PubMed: 22144907]
- 62Treangen TJ, Salzberg SL. Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat Rev Genet*. 2012; 13:36–46.
- 63Metzker ML. Sequencing technologies [mdash] the next generation. *Nat Rev Genet*. 2010; 11:31–46. [PubMed: 19997069]
- 64Yang R, Chen L, Newman S, Gandhi K, Doho G, Moreno CS, et al. Integrated analysis of whole-genome paired-end and mate-pair sequencing data for identifying genomic structural variations in multiple myeloma. *Cancer Inform*. 2014; 13:49–53.
- 65Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, et al. Real-time DNA sequencing from single polymerase molecules. *Science*. 2009; 323:133–8. [PubMed: 19023044]
- 66Greer SU, Nadauld LD, Lau BT, Chen J, Wood-Bouwens C, Ford JM. Linked read sequencing resolves complex genomic rearrangements in gastric cancer metastases. *Genome Med*. 2017; 9:57. [PubMed: 28629429]
- 67Behjati S, Gundem G, Wedge DC, Roberts ND, Tarpey PS, Cooke SL, et al. Mutational signatures of ionizing radiation in second malignancies. *Nat Commun*. 2016; 7:12605. [PubMed: 27615322]

FIG. 1. mFISH karyotypes of four individual clones that contained stable translocations resulting from ionizing radiation exposure. Clone K1-F18 is derived from exposure to 1.0 GeV/amu Fe ions; the remaining three clones are derived from exposure to ^{137}Cs gamma rays. The rearrangements shown were observed in every cell within a particular clone. All clones harbored simple reciprocal translocations. In addition to a simple translocation, K1-400 H6 also contained a complex exchange that simultaneously involved three chromosomes. mPAINT nomenclature (62) is used to describe the various simple and complex translocations observed.

FIG. 2.

Gamma-irradiated clone 4A, as analyzed by both mFISH and by chromosome-3-specific dGH. Panel A: A t(3;4) translocation by mFISH. Panel B: The same clone probed by dGH showing a prominent inversion (green arrow), also confirming the (3;4) translocation involving the same chromosome 3 homologue (yellow arrows).

FIG. 3.
Sanger sequencing results of PCR products displaying the translocation site containing a shared “AAGG” sequence between chromosomes 3 and 4.

Chromosome 4

CTCTGGAAGCTTCTTCCCAG**AAGG**GCACTTGCCAGTTGCCACTCAGA

GGAGGTCCCCACTGTTTTAATTCTAT**AAGG**AAACAGCTTTGTAATGA

Chromosome 3

FIG. 4.

Schematic drawing of the translocation showing the “AAGG” sequence in red and deleted bases in the translocations in underlined italics.

TABLE 1
Prospective Translocations in Clone K1-400C4; 30× Depth; Filtered against Panel of Five Normal Samples

First chromosome	POS	Read depth filter	Second chromosome	End position of variant	PE support	Mapping quality	Connection type	Breakpoint estimate
4	139706155	PASS	3	170398644	PE = 25	MAPQ = 60	CT = 3 to 5	Precise
4	139706159	PASS	3	170398648	PE = 22	MAPQ = 60	CT = 5 to 3	Precise
9	33130549	PASS	6	43655575	PE = 6	MAPQ = 60	CT = 5 to 3	Imprecise
10	65424406	PASS	1	162690280	PE = 8	MAPQ = 60	CT = 5 to 5	Precise
10	65432504	PASS	1	162690281	PE = 5	MAPQ = 60	CT = 5 to 5	Imprecise
GL000192.1	183009	PASS	16	70977727	PE = 7	MAPQ = 52	CT = 3 to 5	Imprecise

TABLE 2
Prospective Inversions in Clone K1-400C4, 30× Depth, after Filtering against Panel of Normal Samples

First chromosome	POS	Read depth filter	Second chromosome	End position of variant	PE support	Mapping quality	Connection type	Breakpoint estimate
1	157331077	PASS	1	157478431	PE = 8	MAPQ = 60	CT = 5 to 5	Imprecise
1	248610266	PASS	1	248808345	PE = 6	MAPQ = 44	CT = 5 to 5	Imprecise
2	195979381	PASS	2	195984619	PE = 15	MAPQ = 60	CT = 3 to 3	Precise
4	154057300	PASS	4	154206834	PE = 5	MAPQ = 39	CT = 5 to 5	Precise
4	154057401	PASS	4	154222227	PE = 4	MAPQ = 37	CT = 3 to 3	Imprecise
4	183570302	PASS	4	183572494	PE = 11	MAPQ = 42	CT = 3 to 3	Imprecise
6	158059416	PASS	6	158064640	PE = 6	MAPQ = 60	CT = 5 to 5	Precise
6	167634246	PASS	6	167699656	PE = 8	MAPQ = 49	CT = 5 to 5	Imprecise
10	18847591	PASS	10	18858517	PE = 6	MAPQ = 60	CT = 3 to 3	Imprecise
10	27192130	PASS	10	37135628	PE = 4	MAPQ = 36	CT = 3 to 3	Imprecise
11	25076632	PASS	11	25083536	PE = 12	MAPQ = 60	CT = 3 to 3	Precise
11	49843892	PASS	11	49909098	PE = 4	MAPQ = 44	CT = 5 to 5	Imprecise
13	19767008	PASS	13	21769416	PE = 5	MAPQ = 49	CT = 3 to 3	Imprecise
13	19921797	PASS	13	25112377	PE = 4	MAPQ = 56	CT = 3 to 3	Imprecise
13	75735136	PASS	13	75735572	PE = 6	MAPQ = 60	CT = 3 to 3	Imprecise
14	58259295	PASS	14	58412164	PE = 3	MAPQ = 50	CT = 5 to 5	Precise
15	49412023	PASS	15	49412468	PE = 7	MAPQ = 49	CT = 5 to 5	Imprecise
16	20449340	PASS	16	20601186	PE = 5	MAPQ = 46	CT = 3 to 3	Imprecise
21	46294684	PASS	21	47312137	PE = 6	MAPQ = 51	CT = 3 to 3	Imprecise