

# The life history of neochromosomes revealed

Anthony T Papenfuss<sup>1,2,3,4,\*</sup> and David M Thomas<sup>2,5,6,\*</sup>

<sup>1</sup>Bioinformatics Division; The Walter & Eliza Hall Institute of Medical Research; Parkville, Australia; <sup>2</sup>Peter MacCallum Cancer Centre; East Melbourne, Australia;

<sup>3</sup>Department of Medical Biology; University of Melbourne; Melbourne, Australia; <sup>4</sup>Sir Peter MacCallum Department of Oncology; University of Melbourne; Melbourne, Australia;

<sup>5</sup>The Kinghorn Cancer Centre, Garvan Institute of Medical Research; Sydney, Australia; <sup>6</sup>University of New South Wales; Sydney, Australia

**Keywords:** neochromosome, chromothripsis, breakage-fusionbridge mechanism, liposarcoma, cancer evolution, computational modelling

Neochromosomes are a little-studied class of chromosome-scale mutations that drive some cancers. By sequencing isolated neochromosomes from liposarcomas, we recently defined their structure at single-nucleotide resolution and proposed a model for their life history. Here, we summarize that work, highlighting significant aspects and providing historical context and insight into the discovery process.

Neochromosomes are giant extra chromosomes that are found in 3% of cancers overall<sup>1</sup> but are particularly common in certain cancer subtypes, including well-differentiated and de-differentiated liposarcomas (WD/DD-LPS). Neochromosomes may be circular or linear, possess centromeres and telomeres (in the linear form), and can harbor oncogenes at high copy number that drive tumor development. They were probably first observed in the 1950s by the Swedish geneticist Albert Levan as ring chromosomes in tumors,<sup>2</sup> but were not identified as a distinct cancer-causing entity until the 1990s.<sup>3</sup> How neochromosomes form has remained a mystery.

Low-resolution analyses of WD/DD-LPS neochromosomes using fluorescent *in situ* hybridization (FISH) and copy number arrays provided an approximate idea of DNA content, and an (under-) estimate of their size and copy number. This led to the hypothesis that the breakage-fusion-bridge (BFB) mechanism<sup>4</sup> might be involved in their formation,<sup>1,3</sup> but no solid proof of this existed. BFB was discovered by Barbara McClintock in maize in 1951. It is initiated by telomere loss: single telomere loss leads to the formation of a dicentric linear chromosome during replication, whereas loss of both telomeres

leads to the formation of a ring chromosome. During replication, DNA crossover events between the sister chromatids can occur. An odd number of crossovers leads to an interlocked and dicentric ring. Dicentric chromosomes form anaphase bridges and are torn apart during cell division. This leads to unequal inheritance of DNA by daughter cells and possible selection of some cell subpopulations.<sup>1</sup>

In our recent publication,<sup>5</sup> we describe the analysis of sequencing data from purified neochromosomes obtained from WD/DD-LPS cell lines by flow cytometric cell sorting and whole-genome sequence data from patient tumors. Flow sorting of chromosomes works by selecting DNA molecules based on the intensity of 2 fluorescent dyes that are markers of size and GC-content. We purified neochromosomal DNA in order to sequence the neochromosome at high coverage. Unavoidably, the rest of the genome was sequenced at low coverage. This can be seen in plots of the neochromosome copy number, which highlight multiple fascinating features in the background of the normal chromosomes (e.g., Fig. 1).

Analysis and interpretation of the neochromosome sequencing data required a high level of creativity. Almost every aspect of analysis (see Fig. 1A) required a

specialized method because of the unusual nature of the complex data generated. This included the copy number analysis, which integrated information from spectral karyotyping to calibrate the copy number on the neochromosome, fusion detection, classification of neochromosomal material, and finally a mathematical model of the mutational mechanisms that shaped the neochromosome over time. In a similar vein, we repeatedly found that we had to invent new terms to describe and discuss the data—including the term ‘neochromosome’ itself.

The sequencing data provided the first high-resolution view of a neochromosome, allowing us to describe their architecture in detail (Fig. 1B). Several observations suggested the involvement of chromothripsis, the shattering of a chromosome in a single catastrophic event<sup>6</sup> that was first identified in tumors that were otherwise genomically stable. The occurrence of a chromothriptic event is not observed directly, but rather inferred from the ruins of the chromosomes that remain. A set of statistical criteria have been proposed to test for chromothripsis,<sup>7</sup> but they are strongly motivated by the ideal case of a largely quiescent genome and must be adapted for more complex cases (see the

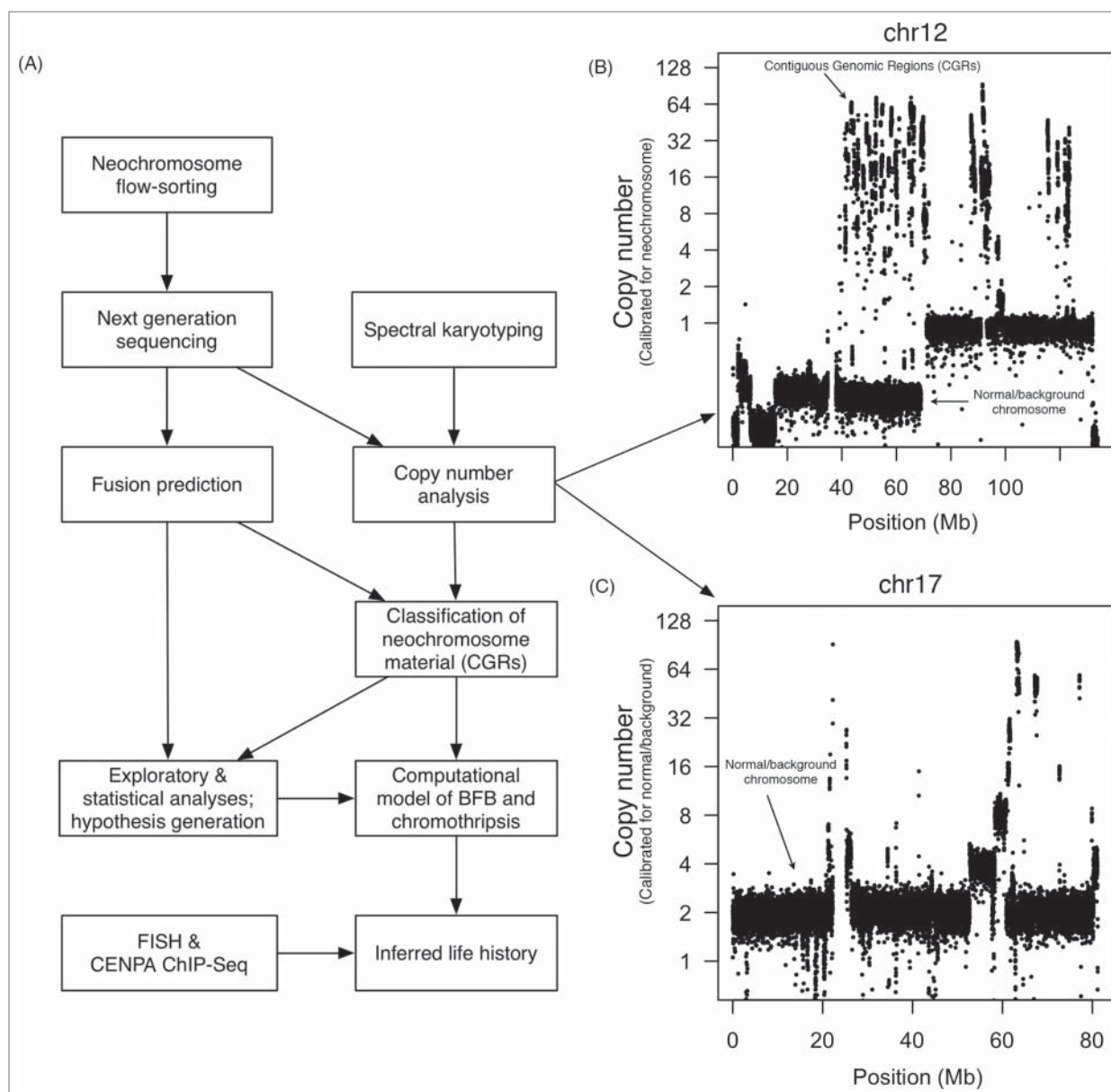
© Anthony T Papenfuss and David M Thomas

\*Correspondence to: Anthony T Papenfuss; Email: papenfuss@wehi.edu.au; David M Thomas; Email: d.thomas@garvan.org.au

Submitted: 12/10/2014; Revised: 12/17/2014; Accepted: 12/17/2014

<http://dx.doi.org/10.1080/23723556.2014.1000698>

This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.



**Figure 1.** Integration of multiple omics data and computational modeling provides insight into the life history of neochromosomes. **(A)** Schematic of the analysis of next-generation sequencing data from flow sorted neochromosomes, integration of fluorescent in situ hybridization (FISH) and ChIP-seq against centromere protein A (CENPA), and computational modeling. **(B)** Copy number profile for neochromosomal material derived from chromosome 12 in the T1000 cell line. This shows high levels of amplification and enrichment compared to a low background. Copy numbers are calibrated for the neochromosome; copy numbers less than one correspond to normal chromosomes. The highly amplified contiguous genomic regions (CGRs), which are the building blocks of the neochromosome, are indicated. **(C)** A "normal" or "background" chromosome from the T1000 cell line. Low coverage sequencing of this chromosome provides evidence for linear breakage-fusion-bridge.

Supplementary Information of references 5 and 8 for a detailed discussion). Our data were the antithesis of this: nearly every chromosome was involved and showed extreme copy number amplification. Nevertheless, the telomeric regions of the neochromosome showed the classic signature of chromothripsis and multiple

other lines of evidence supported the notion that chromothripsis underlies the highly amplified regions of chromosome 12 present in all neochromosomes. In particular, we discovered that the fundamental building blocks (we used the term contiguous genomic regions or CGRs; Fig. 1B) were rearranged and fused edge-to-edge in

an intrachromosomal walk, suggestive of an early chromothriptic event involving chromosome 12 only.

The key to demonstrating that both chromothripsis and BFB were involved in the formation of the neochromosome was the development of a mathematical model of these mechanisms. There were really 4

components to our model: a stochastic simulation of circular BFB, a simple model of selection, a model of chromothripsis, and a way to compare the model with the data. Because of the lack of detailed information about these mechanisms, our approach was to try to make few assumptions, keep the models simple, and use empirical data where possible. How we compared the model results to the data was the most critical part. We ran the model hundreds of times and compared summary statistics of the predictions from each run with summary statistics of the data. We found that we could not fit the BFB-only model to the data for any parameter choices, whereas the BFB plus

chromothripsis model easily fitted the data.

Through detailed analysis of sequencing data from liposarcoma neochromosomes, development of specialized methods for genotyping, hypothesis testing using computational modeling, and integration with observations of centromeres across multiple neochromosomes, we were able to infer the life history of these oncogenic chromosomal-scale mutations.

#### Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

#### References

1. Garsed DW, Holloway AJ, Thomas DM. Cancer-associated neochromosomes: a novel mechanism of oncogenesis. *Bioessays* 2009; 31:1191-200; PMID:19795405; <http://dx.doi.org/10.1002/bies.200800208>
2. Levan A. Self-perpetuating ring chromosomes in two human tumours. *Hereditas* 1956; 42:366-72; <http://dx.doi.org/10.1111/j.1601-5223.1956.tb03023.x>
3. Gisselsson D, Hoglund M, Mertens F, Johansson B, Dal Cin P, Van den Berghe H, Earnshaw WC, Mitelman F, Mandahl N. The structure and dynamics of ring chromosomes in human neoplastic and non-neoplastic cells. *Hum Genet* 1999; 104:315-25; PMID:10369161; <http://dx.doi.org/10.1007/s004390050960>
4. McClintock B. The Behavior in Successive Nuclear Divisions of a Chromosome Broken at Meiosis. *Proc Natl Acad Sci U S A* 1939; 25:405-16; PMID:16577924; <http://dx.doi.org/10.1073/pnas.25.8.405>
5. Garsed Dale W, Marshall Owen J, Corbin Vincent DA, Hsu A, Di Stefano L, Schröder J, Li J, Feng ZP, Kim BW, Kowarsky M, et al. The architecture and evolution of cancer neochromosomes. *Cancer Cell*; 26:653-67; PMID:25517748; <http://dx.doi.org/10.1016/j.ccell.2014.09.010>
6. Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, Pleasance ED, Lau KW, Beare D, Stebbings LA, et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* 2011; 144:27-40; PMID:21215367; <http://dx.doi.org/10.1016/j.cell.2010.11.055>
7. Korbel JO, Campbell PJ. Criteria for inference of chromothripsis in cancer genomes. *Cell* 2013; 152:1226-36; PMID:23498933; <http://dx.doi.org/10.1016/j.cell.2013.02.023>
8. Li Y, Schwab C, Ryan SL, Papaemmanuil E, Robinson HM, Jacobs P, Moorman AV, Dyer S, Borrow J, Griffiths M, et al. Constitutional and somatic rearrangement of chromosome 21 in acute lymphoblastic leukaemia. *Nature* 2014; 508:98-102; PMID:24670643; <http://dx.doi.org/10.1038/nature13115>

#### Funding

DMT and ATP were supported by a National Health and Medical Research Council (NHMRC) Project Grant (1004022) and the Liddy Shriver Sarcoma Initiative. DMT was supported by an NHMRC Senior Research Fellowship (1003929). ATP was supported by an NHMRC Career Development Fellowship (1003856) and an NHMRC Program Grant (1054618). The work benefitted from support by the Victorian State Government Operational Infrastructure Support and Australian Government NHMRC IRIISS.