

Chapter 5

Clonal haematopoiesis after childhood cancer treatment

1. Introduction

The findings of the preceding chapters demonstrate that pre-malignant CH is associated with clinical and genetic features that can help distinguish individuals at highest risk of developing certain blood cancers, particularly *de novo* AML. These experiments studied individuals from the general population without a known history of cancer or haematological disorder. Further work will be necessary to adapt AML predictive models to patient groups prone to CH with distinct genetic features. As discussed in the general introduction, CH is particularly common in certain clinical contexts, notably aplastic anaemia and following cytotoxic treatment for an unrelated malignancy (Bowman et al., 2018). CH in adult cancer patients has recently become an active area of research due to the increasing numbers of cancer survivors at elevated risk of CH-associated pathology, including therapy-related myeloid neoplasms (t-MN) and earlier onset of common non-malignant conditions, particularly cardiovascular disease (Bowman et al., 2018; Carver et al., 2007; Morton et al., 2018). CH has emerged as a potentially promising biomarker for the risk of t-MN and other late effects of cancer treatment (Bolton et al., 2019; Coombs et al., 2017; Gibson et al., 2017; Gillis et al., 2017; Takahashi et al., 2017).

Childhood cancer survivors display an earlier onset of ageing-associated cardiometabolic conditions (Armstrong et al., 2016; Bhakta et al., 2017; Rowland and Bellizzi, 2014) and an elevated risk of t-MN and other secondary malignancies (Bhatia et al., 2007; Pui et al., 1991; Turcotte et al., 2018). Predicting and mitigating long-term complications of treatment is emerging as a dominant challenge in an era where a large proportion of children with cancer can be cured of their primary malignancy (Oeffinger et al., 2006). However, the

prevalence, genetic landscape and clinical significance of CH in this population is largely unknown.

The aims of the experiments described in this chapter were the following:

- 1) Evaluate whether CH is prevalent in childhood cancer survivors who have received intensive cytotoxic treatment and/or radiotherapy.
- 2) Investigate the natural history of a case of paediatric t-MN lacking an MLL rearrangement.

The following introduction provides an overview of existing literature on cytotoxic therapy related CH and the pathogenesis of t-MN.

1.1 Therapy-related myeloid neoplasms

Epidemiology and risk factors

Therapy-related myeloid neoplasms comprise any AML or MDS arising after chemo and/or radiotherapy for a primary cancer, organ transplant or auto-immune condition (Arber et al., 2016). It constitutes one of the most challenging long-term complications of cancer treatment, with survival measured in months for most patients (Bhatia, 2013). Cytotoxic agents associated with the highest risk of t-MN are alkylating agents, topoisomerase II inhibitors and platinum-based drugs (Morton et al., 2018). The incidence and risk factors for t-MN have fluctuated as chemotherapy regimens for the commonest solid cancers have evolved (Bhatia, 2013; Morton et al., 2018). Over the past several decades, t-MN has accounted for a rising proportion of all newly diagnosed AML/MDS cases (Morton et al., 2018; Morton et al., 2014). Currently, t-MN constitutes around 10-20% of AML and MDS diagnoses, with an annual incidence of approximately 0.62/100,000 (De Roos et al., 2010; Hulegardh et al., 2015; Morton et al., 2018). A recent survey of all t-MN cases entered in the US SEER cancer registry between 2000 and 2014 found that nearly all solid tumour types were associated with t-MN, with the highest risk seen in patients treated for malignant bone tumours, followed by soft tissue sarcoma, testicular cancer, ovarian carcinoma and CNS malignancies (Morton et

al., 2018). These findings represent a modest departure from previous epidemiological trends showing highest t-MN risk among breast cancer and lymphoma patients (Morton et al., 2010; Morton et al., 2018). Several solid tumour types were newly associated with t-MN risk, most likely reflecting recent introduction or increase in use of platinum agents to treatment protocols (Morton et al., 2018). Younger age at chemo/radiotherapy exposure correlated with higher t-MN risk, with high cumulative incidence of t-MN observed in children treated for solid tumours (5% to 11%) (Bhatia et al., 2007; Kushner et al., 1998; Le Deley et al., 2003; Morton et al., 2018).

Around 16-20% of t-MN patients harbour penetrant germline variants implicated in cancer susceptibility (Churpek et al., 2016; Felix et al., 1996; Schulz et al., 2012; Voso et al., 2015), compared with 9.5-12.6% of cancer patients overall and 1-2.7% of individuals without cancer (Pritchard et al., 2016; Schrader et al., 2016; Zhang et al., 2015). Cancer-predisposing germline mutations in t-MN patients are frequently reported in genes involved in mediating cellular responses to DNA damage, such as *BRCA1*, *BRCA2*, *BARD1* and *TP53* (Felix et al., 1996; Felix et al., 1998; Schulz et al., 2012). This observation may help explain the notorious chemoresistance of t-MNs (Bhatia, 2013; McNerney et al., 2017). Germline factors may constitute a particularly powerful risk factor in children at highest risk of t-MN. For example, children with soft tissue or bone malignancies have an 11% cumulative 5-year risk of t-MN (Bhatia et al., 2007). This patient group appears to have an exceptionally high burden of germline variants predisposing to cancer, identified in nearly 50% of individuals in the most recent survey (Ballinger et al., 2016).

Genomic landscape and classification

The somatic genomic features of t-MN are similar overall to those seen in non-therapy related myeloid neoplasms, but with dramatic enrichment for high-risk changes, notably rearrangements involving *KMT2A* (*MLL*) and *RUNX1*, *TP53* mutations and chromosome 5 and/or 7 losses (Bhatia, 2013; Smith et al., 2003). In adults, two subtypes of t-MN are delineated based on chemotherapy exposure, genomic features and clinical behaviour (McNerney et al., 2017). The alkylating agent-related class of t-MN constitutes around 70% of cases and is characterised by the high-risk cytogenetic changes del(5q) and -7/del(7q) and *TP53* mutations (in around 33%) (Heuser, 2016), a relatively long latency (5-7 years from

cytotoxic exposure) and a tendency to initially present as MDS progressing towards AML (McNerney et al., 2017). In addition to alkylating agents (e.g., cyclophosphamide, melphalan), this class of t-MN is associated with exposure to platinum-based agents (e.g., cisplatin, carboplatin) and purine analogues (e.g., azathioprine, fludarabine) (McNerney et al., 2017; Offman et al., 2004; Waterman et al., 2012). The second broad category of t-MN is associated with topoisomerase II inhibitor exposure (e.g., anthracyclines and etoposide)(McNerney et al., 2017). The topoisomerase II (TOP2) inhibitor class of t-MN typically presents as frank AML and has a shorter latency (median 2-3 years) (Heuser, 2016; Smith et al., 2003). This may in part be driven by translocations that are common in these t-MN involving *KMT2A (MLL)*, *RUNX1* or *PML-RARA*, powerful oncogenic rearrangements that tend to require few cooperating events to trigger leukaemic transformation (Andersson et al., 2015; McNerney et al., 2017; Papaemmanuil et al., 2016; TCGA et al., 2013).

t-MN pathogenesis: chemotherapy-induced DNA damage or clonal selection?

Until recently, the conventional model of t-MN pathogenesis proposed that most cases were attributable to somatic driver events directly induced by cytotoxic agents (Bhatia, 2013). Many chemotherapy drugs associated with t-MN are mutagenic, and some are associated with particular patterns of genomic damage. For example, TOP2 inhibitors may increase the likelihood of reciprocal translocations by delaying ligation of double-strand breaks, thus prolonging the opportunity for recombination with DNA from another chromosome (Cowell and Austin, 2012). In keeping with this model, fusion oncogenes in t-MN arising post TOP2 inhibitor treatment tend to have breakpoints consistent with processing of 4-base staggered double-strand breaks from TOP2-mediated cleavage (Felix, 2001; Hasan et al., 2008; Mistry et al., 2005).

The alkylating agent class of t-MN is characterised by complex karyotypes, high numbers of copy number aberrations and *TP53* mutations in over a third of cases (Itzhar et al., 2011; Smith et al., 2003). Alkylating agents covalently modify DNA and promote DNA cross-linking double-strand breaks (Fu et al., 2012). It was thought that this genotoxicity induced structural changes and occasionally *TP53* mutations, with the latter contributing to genomic instability (Bhatia, 2013).

However, this model of t-MN pathogenesis was refuted by the work of Wong et al, who investigated the natural history of *TP53*-mutated t-MN (Wong et al., 2015b). Ultra-sensitive duplex sequencing demonstrated that the *TP53* driver mutation present (at clonal VAF) in the t-MN was usually detectable at very low levels (VAF 0.003-0.7%) in bone marrow samples taken prior to commencing cytotoxic treatment for the primary malignancy (Wong et al., 2015b). Furthermore, the point mutation burden and patterns did not differ between t-MN and *de novo* AML (Wong et al., 2015b). These findings suggested that cytotoxic treatment selected for pre-existing *TP53*-mutated HSCs, and that the cytogenetic complexity observed in the t-MNs reflected abrogation of the *TP53*-mediated DNA damage response and survival of cells that would otherwise have undergone apoptosis (Wong et al., 2015b). The clonal selection model was corroborated by a follow-up experiment in which *TP53*-mutated clones transplanted into mice only expanded if the animals were exposed to cytotoxic therapy (Wong et al., 2015b). Moreover, screening peripheral blood samples from a cohort of otherwise healthy elderly individuals (n=20) identified *TP53* mutations at very low VAF (<0.1%) in 37% (Wong et al., 2015b). These mutations persisted over time with little or no clonal expansion, suggesting that they conferred minimal selective advantage in the absence of unusual levels of genotoxic stress (Wong et al., 2015b). A contemporaneous study by Ok et al. compared *TP53* mutations in t-AML and *de novo* AML and found no evidence suggesting that *TP53* drivers in the former were induced by distinct chemotherapy-related mutational processes: there were no differences in mutation distribution, sequence context or proportion of transitions versus transversions (Ok et al., 2015). The finding that t-MN *TP53* drivers predate chemotherapy exposure has since been reproduced by other experiments (Schulz et al., 2015; Takahashi et al., 2017).

Clonal haematopoiesis as a biomarker for t-MN risk

An important role for clonal selection in t-MN pathogenesis was corroborated by recent studies investigating CH in cancer patients. CH is dramatically more prevalent in cancer survivors compared to individuals of the same age who have not been exposed to cytotoxic agents/radiotherapy and is enriched for mutations in *TP53* and its negative regulator *PPM1D* (Coombs et al., 2017; Gibson et al., 2017; Gillis et al., 2017; Takahashi et al., 2017). Numerous elegant studies in both mouse and human have demonstrated that cytotoxic agents and

radiotherapy promote expansion of HSCs harbouring *TP53* or *PPM1D* mutations (Bondar and Medzhitov, 2010; Hsu et al., 2018; Kahn et al., 2018; Wong et al., 2015b). In keeping with the clinical significance of CH in the general population, CH in cancer survivors is associated with higher risk of t-MN as well as with non-malignant adverse outcomes (Coombs et al., 2017; Gibson et al., 2017; Gillis et al., 2017; Takahashi et al., 2017).

Childhood t-MN

Although t-MN is a leading cause of death in paediatric cancer patients surviving their primary cancers, relatively little is known about its pathogenesis in children (Bhatia et al., 2007; Heuser, 2016; Kushner et al., 1998; Le Deley et al., 2003; Pui et al., 1991). The relative contributions of germline risk factors, chemotherapy-induced driver mutations and clonal selection are unclear. The genomic landscape of paediatric t-MNs has not been well characterised, complicating efforts to trace their clonal evolution. However, it is conceivable that the genetic basis overlaps with that of paediatric AML/MDS/MPN arising in the absence of cytotoxic therapy, possibly with enrichment for high-risk features as seen in adults. Compared to adult MDS, paediatric myeloid neoplasms are enriched for mutations in the RAS oncogenes as well as *RUNX1*, *SETBP1* and *ASXL1* (Locatelli and Strahm, 2018; Pastor et al., 2017). Furthermore, more than 30% of paediatric MDS patients have an inherited cancer predisposition or bone marrow failure syndrome compared to <5% of adults (Hasle, 2016). Deletions affecting chromosome 7 (-7/7q-) or chromosome 5 (-5/-5q) are present in around 25% and 1% of paediatric MDS cases, respectively (Hasle, 2016).

Allogeneic HSCT remains the only potential cure for paediatric t-MN (Locatelli and Strahm, 2018) and unlike their adult counterparts, most children with t-MN are HSCT candidates (Hasle, 2016; Locatelli and Strahm, 2018). Importantly, the only factor associated with improved overall survival in paediatric t-MN patients is shorter delay between t-MN diagnosis and transplant (Locatelli and Strahm, 2018; Maher et al., 2017). It is therefore conceivable that early detection and monitoring of patients at highest risk of progressing to t-MN could improve outcomes by minimising the interval between t-MN manifestations and allogeneic HSCT.

Current knowledge of paediatric t-MN natural history is limited to four case reports of children with MLL-rearranged (MLLr) t-MN after TOP2 inhibitor treatment (Blanco et al.,

2001; Megonigal et al., 2000; Ng et al., 2004; Robinson et al., 2008). As discussed above, there is some evidence that reciprocal fusions involving MLL may be directly induced by TOP2 inhibitors (McNerney et al., 2017). Consistent with this view, in each of these four cases, sensitive methods failed to detect the MLL fusion in blood or bone marrow samples taken before chemotherapy exposure (Blanco et al., 2001; Megonigal et al., 2000; Ng et al., 2004; Robinson et al., 2008). However, in three of the four case reports, the MLL fusion was detectable in blood and/or bone marrow over a year before t-MN presented clinically (17, 15.5, and 37 months latency in Blanco et al, Megonigal et al, and Robinson et al, respectively) (Blanco et al., 2001; Megonigal et al., 2000; Robinson et al., 2008). The shortest interval between MLLr detection and t-MN diagnosis (3 months) was reported by Ng et al in a child who developed t-MN only six months after diagnosis with hemophagocytic lymphohistiocytosis (Ng et al., 2004). These case reports offer some hope that even chemotherapy-induced fusion oncogenes generally associated with shorter latency to t-MN may be detectable early enough in disease evolution to enable monitoring and expedite definitive treatment. However, I could not identify any studies investigating the natural history of paediatric t-MN lacking an oncogenic fusion.

2. Results

2.1 Prevalence of CH-PD in childhood cancer survivors

To determine whether CH prevalence is elevated in children who have undergone intensive chemo/radiotherapy, we performed targeted deep sequencing of peripheral blood DNA from 84 paediatric cancer survivors to search for candidate driver mutations. The median age at cancer diagnosis was 4.5 years, and the commonest malignancies were acute lymphoblastic leukaemia (n=21), neuroblastoma (n=17) and non-Hodgkin lymphoma (n=10). Nineteen children had received a hematopoietic stem cell transplant (8 allogeneic and 11 autologous). The median interval between completion of cancer treatment and blood sampling was 6 years (range 2 – 25). Patient characteristics are summarised in Table 5.1 with details for each individual shown in Appendix 3.

Table 5.1 | Cohort summary

Diagnosis	Number of individuals	Mean age at diagnosis (years)	Mean time since last chemo/radiotherapy (years)
Neuroblastoma	17	3.0	11.4
Rhabdomyosarcoma	7	5.5	6.7
Acute lymphoblastic leukaemia	21	4.2	6.8
Non-Hodgkin lymphoma	10	6.9	8.3
Germ cell tumour	4	10.0	5.1
Lymphoblastic lymphoma	3	6.1	7.8
Hodgkin lymphoma	6	14.6	5.6
Nephroblastoma	5	3.5	7.6
Hepatoblastoma	1	0.3	9.4
Ewing sarcoma	4	8.0	8.0
Non-rhabdomyosarcoma soft tissue sarcoma	2	7.7	6.0
Choriocarcinoma	1	12.8	3.5
Nasopharyngeal carcinoma	1	15.9	3.0
Langerhans cell histiocytosis	2	3.4	6.6

Multiplex PCR was used to amplify 32 selected regions of 14 genes frequently mutated in CH and t-MN, including hotspots in the RAS oncogenes *NRAS* and *KRAS* (recurrently mutated in paediatric MDS/MPN) and all exons of *TP53* and *PPM1D* (Table 5.2; Methods section 2.3)(Coombs et al., 2017; Gibson et al., 2017; Locatelli and Strahm, 2018).

Table 5.2 | Genomic regions sequenced by multiplex PCR

Gene	Chromosome	Target codon/exon
<i>NRAS</i>	1	p.G12D
<i>SF3B1</i>	2	p.K666N; p.K700E
<i>DNMT3A</i>	2	p.R882/p.R693C
<i>IDH1</i>	2	p.R132H
<i>KIT</i>	4	exon 17
<i>NPM1</i>	5	p.L287fs*13
<i>JAK2</i>	9	p.V617F
<i>KRAS</i>	12	p.G12R
<i>IDH2</i>	15	p.R140Q; p.R172K
<i>PPM1D</i>	17	exons 1 - 6
<i>TP53</i>	17	exons 1 - 12
<i>SRSF2</i>	17	p.P95L
<i>ASXL1</i>	20	exon 12
<i>U2AF1</i>	21	p.S34F; p.Q157R

The median sequencing depth achieved across all regions of interest was 5,295X. No somatic mutations above the assay sensitivity threshold (VAF \geq 0.008) were observed in any of the 84 long-term paediatric oncology follow-up patients nor in 3 children with no history of cancer (Methods section 3.2).

2.2 Tracing the evolution of a paediatric t-MN with driver mutations in *PTPN11* and *SETBP1* to emergence in early neuroblastoma treatment

As discussed in the introduction, studies of paediatric t-MN evolution have thus far been limited to case reports of children presenting with MLLr t-MN (Blanco et al., 2001; Megonigal et al., 2000; Ng et al., 2004; Robinson et al., 2008). The aim of this experiment was to retrace the emergence of a paediatric t-MN with genetic features akin to the alkylating agent class of adult t-MN described earlier.

Case Report

A 4-year old girl presented with high-risk, metastatic (stage 4) neuroblastoma with bone marrow involvement. Apart from focal neuroblastoma involvement, the initial bilateral staging trephines and aspirates showed normal trilineage haematopoiesis. Pre-treatment blood counts were normal. She was enrolled on the high-risk neuroblastoma SIOPEN trial protocol (HR-NBL-1.7/SIOPEN, NCT01704716) and underwent Rapid COJEC induction chemotherapy consisting of ten weeks of treatment with a total of five chemotherapy agents: carboplatin, etoposide, vincristine, cisplatin and cyclophosphamide at cumulative doses of 1.5g/m², 1.4g/m², 12g/m², 320mg/m², 4.2g/m², respectively. Bilateral restaging bone marrow biopsies performed following completion of induction chemotherapy and count recovery (day 120 of treatment) remained positive for neuroblastoma infiltration. She therefore received additional induction chemotherapy to achieve metastatic remission, i.e., two cycles of TVD: Topotecan, Vincristine, Doxorubicin at cumulative doses of 15mg/m², 4mg/m² and 90mg/m², respectively. Platelet and neutrophil count recovery were unusually slow (3 months), though the child remained well with no infectious complications. Bone marrow examination following count recovery was normal, with cytomorphological examination negative for metastatic disease. Peripheral blood CD34+ stem cells (PBSC) were therefore harvested and she completed treatment, which included surgery, myeloablative therapy with busulfan and melphalan (BuMel), autologous PBSC rescue, irradiation of the site of primary disease (21 Gy), differentiation therapy (isotretinoin) and anti-GD2 immunotherapy. She remained well throughout, despite slow platelet and neutrophil count recovery after high-dose BuMel. Eight months after finishing treatment (32 months after diagnosis), she was incidentally noted on routine follow-up to have developed moderate peripheral cytopenia with Hb 102 g/dL, white cell count 2.3 x 10⁹/L, neutrophils 1.29 x 10⁹/L and platelets 91 x 10⁹/L. Bone marrow examination revealed <5% blasts and no evidence of neuroblastoma recurrence. G-banded bone marrow karyotyping revealed monosomy 7 in keeping with a developing t-MN. Two months later the patient suffered local neuroblastoma relapse and succumbed to disease progression soon thereafter.

Retracing molecular emergence of t-MN

We applied whole genome and deep targeted sequencing to identify driver events in the peripheral blood at the time of t-MN diagnosis (32 months after first chemotherapy). Sequences were analysed against the reference genome in order to call deleterious germline variants and to achieve maximum sensitivity for somatic changes (Methods section 3.5). In parallel, a matched analysis was performed using whole genome sequencing of parental blood samples. Median coverage of t-MN, maternal and paternal blood samples was 74X, 111X and 100X, respectively. Whole genome sequencing identified somatic complex changes in chromosome 7 (a major clone with 7q- and a subclone with complete monosomy 7) and canonical hotspot mutations in *PTPN11* and *SETBP1* (Figure 5.1 and Table 5.3). Both copy number and point mutation drivers variants were validated by deep targeted sequencing (Methods 3.3-3.6, Figure 5.1). Moreover, unmatched analysis identified a deleterious germline *BARD1* p.E652fs*69 mutation strongly associated with hereditary cancer predisposition (ClinVar accession numbers RCV000115621.5, RCV000200198.2) (De Brakeleer et al., 2010; Ramus et al., 2015; Schrader et al., 2016; Smith et al., 2016). Although this variant had not been detected by routine clinical genetics targeted screening for cancer predisposition during neuroblastoma work-up, it was also present at SNP VAF in the maternal blood sample.

In order to retrace the emergence of the t-MN clone, we performed ultradeep targeted sequencing (median coverage 25,000X) of bilateral bone marrow biopsies taken at the end of Rapid COJEC induction, 4.5 months into treatment and 29 months prior to t-MN presentation. Unfortunately, these were the earliest samples able to be sequenced, with no pre-treatment specimens available. The *PTPN11* p.G503E mutation was present in both left and right bone marrow biopsies at VAF of 0.12% and 0.09%, respectively. The *SETBP1* p.D868G mutation was detected in the left bone marrow biopsy at a lower VAF of 0.074%. These variants were detected by two algorithms, including shearwater, which accounts for the local error rate when calling subclonal mutations (Methods section 3.3) (Gerstung et al., 2012; Gerstung et al., 2014). Although sequencing of PBSC harvest is underway to further validate this finding, the depth of the sequencing and presence of the *PTPN11* in both marrow samples gives a reasonable degree of confidence in its validity. Furthermore, several reads supporting the bone marrow *PTPN11* mutation was subsequently identified by the clinical

diagnostic service using targeted sequencing on an orthogonal platform (Ion Torrent)(data not shown; personal communication from Dr Sam Behjati). The *SETBP1* mutation may be genuine in the left bone marrow and have escaped detection in the contralateral specimen due to rarity and stochastic molecule sampling, but nonetheless warrants additional validation. Copy number analysis of the targeted bone marrow sequencing revealed concordant changes consistent with recurrent copy number aberrations (CNAs) observed in neuroblastoma (Figure 5.1d,e) (Matthay et al., 2016), though lacked sensitivity to confidently call any chromosome 7 losses.

Table 5.3 | Summary of samples and genetic abnormalities

Sample ID	Month since NBL diagnosis	Clinical context	Sample type	Somatic driver events	
				Mutation	VAF (%)
PD31013c	32	t-MN diagnosis	Peripheral blood	<i>PTPN11</i> G503E	51.0
				<i>SETBP1</i> D868G	50.0
				-7/7q-	-
PD31013d	4.5	Staging post rapid COJEC induction	Bone marrow (right iliac crest)	<i>PTPN11</i> G503E	0.09
				<i>SETBP1</i> D868G	-
PD31013e	4.5		Bone marrow (left iliac crest)	<i>PTPN11</i> G503E	0.12
				<i>SETBP1</i> D868G	0.074

Figure 5.1

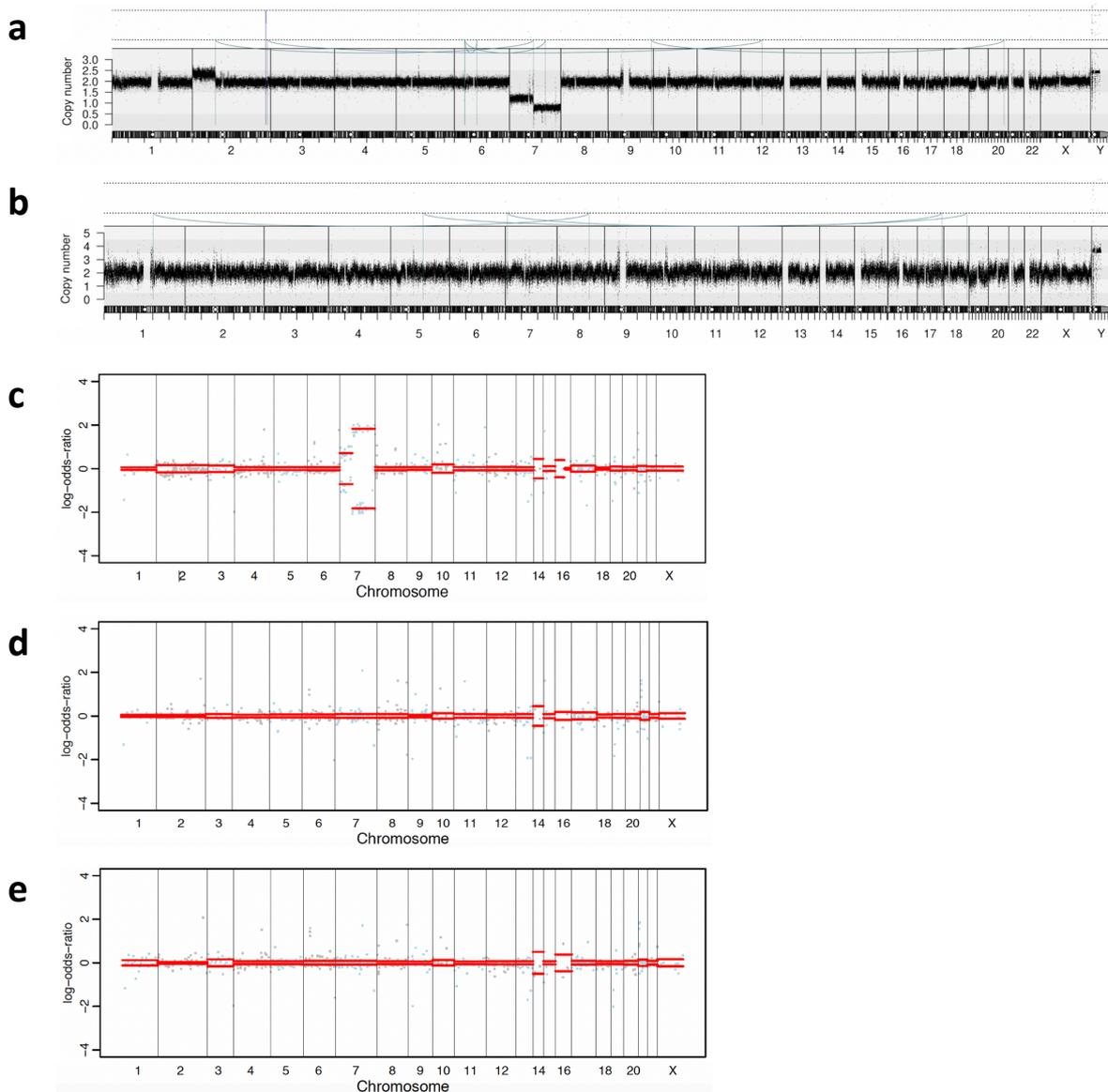


Figure 5.1 | Copy number profiles. a,b, Copy number changes and rearrangements detected from whole genome sequences of PD31013c (t-MN peripheral blood sample) (a) and PD31013d (right post-induction bone marrow biopsy) (b). The x axis shows chromosomal position and the y axis shows absolute copy number. Each dot in the plot represents the copy number of a particular genomic position (10 mega base bins). Coloured lines indicate breakpoints with rearrangements: brown, tandem duplication; blue, deletion; green and turquoise, inversion; grey, interchromosomal rearrangement. **c-e,** Copy number profiles derived from deep (>20,000x) targeted sequencing of t-MN (c) and bilateral bone marrow biopsies taken after induction chemotherapy, 15 months before t-MN emergence: PD31013d (right bone marrow) and PD31013e (left bone marrow) represented in panels (d) and (e), respectively. X-axis represents chromosome position. Y-axes represents allele-specific log-odds-ratio data with chromosomes alternating in blue and gray.

3. Discussion

The absence of any CH in the 84 heavily treated childhood cancer survivors screened stands in stark contrast to the situation recently observed in adults, where CH with candidate driver mutations is dramatically more common post chemo/radiotherapy than in the general population (Bowman et al., 2018; Gibson et al., 2017). Gibson et al. identified CH in over 25% of lymphoma survivors aged 30-39 and in over 40% of those aged 60-69 years (Gibson et al., 2017). Only 10 patients aged 20-29 were included in this study, none of whom had detectable CH, albeit using a less sensitive assay (detection threshold >2%)(Gibson et al., 2017). The most commonly mutated gene was *PPM1D*, which was captured in its entirety in our assay, followed by *DNMT3A* (most recurrent hotspot captured), *TET2* (not captured) and *TP53* (all exons captured) (Table 5.2)(Gibson et al., 2017). These findings have three plausible explanations. Firstly, somatic driver mutations may be extremely uncommon in the young even after exposure to chemotherapy, and hence the substrates for clonal selection are lacking. Secondly, it is possible that accrual of recognized 'driver' mutations is usually insufficient to trigger clonal expansion in the context of a very young haematopoietic niche. This hypothesis is supported by the fact that HSC mutations do begin accumulating early in life (Welch et al., 2012) and that the selective advantage of some CH drivers (most notably spliceosome gene mutations) appears to be age-dependent, implicating age-related changes in HSCs and/or their environment as key determinants of relative fitness (Link and Walter, 2016; McKerrell et al., 2015). This potential explanation is further supported by evidence that cancer-associated mutations are less able to drive clonal expansion in young compared to old stem cells (Zhu et al., 2016). Moreover, a recent study using ultra-sensitive sequencing of serially collected peripheral blood samples demonstrated that bona-fide driver mutations do not always lead to clonal expansion, even after several years (Young et al., 2016). Similar findings have been reported in other tissues, notably oesophagus and kidney, where oncogenic mutation acquisition has been timed to early childhood and adolescence, respectively (Mitchell et al., 2018; Yokoyama et al., 2019). The third potential explanation for our results is that the mutations under positive selection in paediatric cancer patients are so distinct from those observed in adult counterparts that this assay simply does not capture them. Our assay did include targets that are preferentially mutated in paediatric myeloid neoplasms – namely hotspots two RAS oncogenes and *ASXL1* exon 12 – but lacked other

genes and hotspots that are likely to be enriched in paediatric t-MN or CH, notably *SETBP1*, *PTPN11* and *RUNX1* (Hasle, 2016; Tartaglia et al., 2003). In summary, these results should not necessarily be taken to reflect absence of potentially oncogenic HSC mutations in young cancer survivors. Rather, it is possible that even canonical CH driver mutations may not commonly drive clonal outgrowth in children and young adults despite exposure to cytotoxic drugs. More sensitive DNA sequencing methods may enable detection of very rare mutated cells in this patient group, which would lend support to this hypothesis. Equally, future sequencing studies assessing larger cohorts with a broader gene panel are warranted to explore the genetic landscape of paediatric CH. Ideally such work would be informed by a comprehensive understanding of the genomic features of paediatric t-MN, which is currently lacking.

The second experiment described in this chapter traced the emergence of t-MN during treatment for high-risk neuroblastoma. We applied deep targeted sequencing to track missense driver mutations in *PTPN11* to bone marrow samples taken at the end of induction chemotherapy. This case adds to the limited existing knowledge of paediatric t-MN evolution in several ways. Firstly, these findings contribute a fifth case to the literature suggesting that paediatric t-MN evolution typically becomes detectable very early in the treatment for the primary malignancy (Blanco et al., 2001; Megoñigal et al., 2000; Ng et al., 2004; Robinson et al., 2008). In particular, this appears to be the first case reporting early molecular emergence of a non-MLLr case of paediatric t-MN. Although this patient was exposed to high doses of TOP2 inhibitors as well as platinum and alkylating agents, the clinical presentation and genomic features of this t-MN are reminiscent of the so-called alkylating agent class of adult t-MN with chromosome 7 loss, no fusion oncogene and an indolent clinical presentation with MDS rather than overt AML. In retrospect, the slow platelet and neutrophil count recovery following high-dose chemotherapy suggests early clinical manifestations of t-MN. This is in keeping with the tendency for paediatric MDS and MPN/MDS to present with neutropenia and/or thrombocytopenia (Hasle, 2016; Kardos et al., 2003; Niemeyer and Baumann, 2011), whereas adult MDS most frequently manifests with isolated anaemia (Locatelli and Strahm, 2018; Raza and Galili, 2012). As mentioned earlier, the most frequent cooperating point mutation drivers in adult t-MN occur in *TP53*, whereas drivers in *PTPN11* and *SETBP1* are relatively rare (observed in 3-9% and 3% of adult t-MN cases, respectively) (McNerney et al.,

2017). However, mutations in these genes are enriched in paediatric MPN/MDS (Hasle, 2016; Locatelli and Strahm, 2018). Somatic *PTPN11* mutations in particular are a feature of high-risk paediatric MDS warranting prompt allogeneic HSCT (Locatelli and Strahm, 2018).

Moreover, the incidental discovery of a deleterious germline *BARD1* mutation by whole genome sequencing provides further evidence that the contribution of germline predisposition to t-MN (and childhood cancer in general) may be underestimated. *BARD1* is a tumour suppressor involved in regulating the DNA damage response and TP53-mediated apoptosis (Irminger-Finger and Jefford, 2006). Loss-of-function germline mutations have been implicated in susceptibility to a variety of cancers, including t-MN (Irminger-Finger and Jefford, 2006; Schulz et al., 2012).

All discussion of clinical ramifications of these findings remains highly speculative at this point. However, current evidence indicates that the only factor clearly associated with improved childhood t-MN survival is shorter interval between t-MN diagnosis and allogeneic HSCT (Locatelli and Strahm, 2018; Maher et al., 2017). Hence it is possible that earlier detection of early t-MN clones could help address a major cause of mortality in children with cancer (Bhatia et al., 2007; Heuser, 2016; Kushner et al., 1998; Le Deley et al., 2003; Pui et al., 1991). With specific regard to neuroblastoma patients, it is conceivable that early identification of patients who may later require an allogeneic HSCT for t-MN could alter the risk/benefit analysis vis a vis proceeding with myeloablative treatment and autologous PBSC rescue (Fish and Grupp, 2008; Yalcin et al., 2015).

Collectively these findings propose several follow-up experiments. Firstly, scant knowledge of the genomic landscape of paediatric t-MN warrants collaborative efforts to whole genome sequence a sizeable cohort exposed to a range of treatment protocols. This in turn will inform future studies of the prevalence and prognostic significance of CH in childhood cancer patients. In the first instance, targeted sequencing assays should include genes preferentially mutated in paediatric myeloid neoplasms, enough heterozygous SNPs to call subclonal chromosomal arm-level copy number changes and sufficient intron tiling to detect recurrent AML-associated rearrangements. These are tractable goals even with panels small enough for routine clinical use (McKerrell et al., 2016). In the first instance, a retrospective case-control study could help assess the utility of CH-PD as a biomarker of t-MN risk. However, given the high cumulative incidence of paediatric t-MN (Bhatia et al., 2007;

Heuser, 2016; Kushner et al., 1998; Le Deley et al., 2003; Pui et al., 1991), a prospective approach in the context of clinical trial also warrants consideration, particularly for neuroblastoma and sarcoma protocols associated with the highest risk (Bhatia, 2013; Bhatia et al., 2007; Kushner et al., 1998; Morton et al., 2018).