GATA1-mutant clones are frequent and often unsuspected in babies with Down Syndrome: identification of a population at risk of leukemia

Irene Roberts[1], Kate Alford[2], Georgina Hall[3], Gaetan Juban[2], Helen Richmond[1,2], Alice Norton[2], Grant Vallance[2,4], Kelly Perkins[2], Emanuele Marchi[2], Simon McGowan[5], Anindita Roy[1], Gillian Cowan[1], Mark Anthony[6], Amit Gupta[6], John Ho[7], Sabita Uthaya[8], Anna Curley[9], Shree Vishna Rasiah[10], Timothy Watts[11], Richard Nicholl[12], Alison Bedford-Russell[13], Raoul Blumberg[14], Angela Thomas[15], Brenda Gibson[16], Chris Halsey[16], Pek-Wan Lee[17], Sunit Godambe[18], Connor Sweeney[2], Neha Bhatnagar[1], Anne Goriely[19], Peter Campbell[20] and Paresh Vyas[2,4] on behalf of the Oxford-Imperial Down Syndrome Cohort Study Group.

Centre for Haematology, Hammersmith Campus, Imperial College London[1]; MRC Molecular Haematology Unit[2], Computational Biology Research Group[5], Clinical Genetics Group[19], Weatherall Institute of Molecular Medicine, University of Oxford; Paediatric Haematology/Oncology Unit[3], Department of Haematology[4], Department of Neonatal Medicine[6], Oxford University Hospital, NHS Trust; Whipps Cross University Hospital NHS Trust London[7], Chelsea & Westminster Hospital NHS Foundation Trust, London[8], Cambridge University Hospital NHS Foundation Trust[9], Birmingham Women's NHS Foundation Trust[10], Guy's & St. Thomas' NHS Foundation Trust[11],The North West London Hospitals NHS Trust[12], Heart of England NHS Foundation Trust[13], The Whittington Hospital, London[14], Royal Hospital for Sick Children, Edinburgh[15], Royal Hospital for Sick Children, Glasgow[16], The Royal Wolverhampton Hospitals NHS Trust[17], Queen Charlotte's & St Mary's Hospitals, Imperial College Healthcare NHS Trust, London[18], Cancer Genome Group, Wellcome Trust, Sanger Centre, Cambridge UK[20].

Address for correspondence:

Professor Irene Roberts, Centre for Haematology Commonwealth Building, Hammersmith Campus Imperial College London, Du Cane Road, London W12 0NN, United Kingdom, irene.roberts@imperial.ac.uk; +44 208 383 2163 (fax +44 203 313 8223) and

Professor Paresh Vyas Department of Haematology and MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, Oxford University and Oxford Radcliffe Hospital, paresh.vyas@imm.ox.ac.uk; +44 1865 222 309 (fax +44 1865 222 500).

Word count: Abstract: 197 words. Total word count: 4,000 words. 3 figures and 3 tables.

Reference count: 40 references

Running title: High frequency of *GATA1* mutations in Down syndrome

Key words: Down Syndrome, TAM, ML-DS, trisomy 21, hematology in the newborn.

Key points

- *GATA1* mutations are common in neonates with Down syndrome but often unsuspected and detectable only with sensitive methods

- Multilineage blood abnormalities in all Down Syndrome neonates in absence of *GATA1* mutations suggests trisomy 21 itself perturbs hemopoiesis

Abstract

Transient Abnormal Myelopoiesis (TAM), a preleukemic disorder unique to neonates with Down syndrome (DS), may transform to childhood Acute Myeloid Leukemia (ML-DS). Acquired *GATA1* mutations are present in both TAM and ML-DS. Current definitions of TAM neither specify % blasts nor the role of *GATA1* mutation analysis. To define TAM, we prospectively analyzed clinical findings, blood counts and smears and *GATA1* mutation status in 200 DS neonates. All DS neonates had multiple blood count and smear abnormalities. Surprisingly, 195/200 (97.5%) had circulating blasts. *GATA1* mutations were detected by Sanger sequencing/Denaturing High Performance Liquid Chromatography (Ss/DHPLC) in 17/200 (8.5%); all with blasts >10%. Furthermore low-abundance *GATA1* mutant clones were detected by targeted next-generation-resequencing (NGS) in 18/88 (20.4%)(sensitivity ~0.3%) DS neonates without Ss/DHPLC-detectable *GATA1* mutations. No clinical or hematologic features distinguished these 18 neonates. We suggest the term Silent TAM for neonates with DS with *GATA1* mutations detectable only by NGS. To identify all babies at risk of ML-DS, we suggest *GATA1* mutation analysis, blood counts and smears, should be performed in DS neonates. Ss/DPHLC can be used for initial screening but where *GATA1* mutations are undetectable by Ss/DHPLC, NGS-based methods can identify neonates with small *GATA1* mutant clones.

Introduction

Children with Down syndrome (DS) have a 150-fold increased risk of acute myeloid leukemia (ML-DS) during the first 5 years of life, compared to children without DS, despite a lower incidence of other cancers[1,2]. ML-DS cells harbor acquired, N-terminal-truncating mutations in the key hematopoietic transcription factor gene *GATA1*[3-9]. *GATA1* mutations are also present in the neonatal preleukemic disorder Transient Abnormal Myelopoiesis (TAM), which is unique to neonates with trisomy 21[3-5,7-9]. TAM often precedes ML-DS and the same *GATA1* mutations are usually present in both disorders[3], clonally-linking the two conditions.

Retrospective clinical studies suggest that TAM affects ~10% of neonates with DS[10] and has a mortality of ~20%[11-13]. Furthermore, an estimated 20-30% of babies with TAM subsequently develop ML-DS[2,11,12,14]. Thus, TAM is an important clinical problem. Nevertheless, a key difficulty is the lack of clear clinical, hematologic and molecular diagnostic criteria for TAM.

WHO defines TAM as increased peripheral blood blast cells in neonates with DS[15]. Retrospective studies have used differing clinical and/or hematologic criteria to define TAM[11-14]. In these studies, presentation varied from disseminated leukemic infiltration with hepatic fibrosis to largely asymptomatic disease where diagnosis was based on various non-specific blood count abnormalities, such as circulating blasts, leukocytosis and/or thrombocytopenia, which have all been reported in DS neonates without TAM[16]. Importantly, no retrospective studies have systematically screened neonates for *GATA1* mutations[11-14]. Thus, current definitions of TAM neither specify the percentage of blasts considered abnormal in DS neonates nor the role of *GATA1* mutation analysis in diagnosis. As a result, asymptomatic TAM may be missed in some neonates, while in others TAM may be over-diagnosed by relying on non-specific clinical and hematologic features. Indeed, the only large systematic *GATA1* mutation screen, performed by Sanger sequencing (Ss) of PCR products from dried blood spots, found a prevalence of *GATA1* mutations in DS neonates of only 3.8%[17], in contrast to the 5-10% prevalence of TAM diagnosed by clinical and hematologic criteria[10].

The aims of this study were to more precisely define the population at risk of developing ML-DS and understand how to best define TAM. We prospectively, systematically determined *GATA1* mutation status, blood counts, blood cell morphology, blast cell frequency and clinical findings in 200 DS neonates recruited to the Oxford-Imperial Down Syndrome Cohort Study (OIDSCS). This provides a base to accurately study the natural evolution and impact of interventional treatment of this preleukemic disorder for the first time.

Patients, Materials and Methods

Study population: Neonates with a clinical diagnosis of DS confirmed by karyotyping were prospectively enrolled in OIDSCS between October 2006 and March 2012 in 18 UK hospitals. Of 213 DS babies born in OIDSCS hospitals during this period, 94% (200) were recruited. Reasons for non-recruitment were lack of available blood samples (n=7) or parental consent (n=6). OIDSCS recruits represent ~5% of total DS births in England over this period (http://www.wolfson.qmul.ac.uk/ndscr/reports/NDSCRreport11.pdf). Parents gave written informed consent. The study was approved by Thames Valley Research Ethics Committee (06MRE12-10; NIHR Portfolio No.6362).

Laboratory and clinical data: Complete blood counts (CBC) and blood smears were processed in treating hospitals. Peripheral blood smears were assessed by 2 independent observers (IR,GH) blinded to *GATA1* mutation status and recorded on OIDSCS proformas (Supplemental Figure 1). Differential counts were expressed as the percentage (%) of leukocytes on a 200-cell manual differential count. Results on DS neonates were compared with neonatal laboratory normal ranges used at Imperial College Healthcare NHS Trust derived from 80 healthy term (37-42 weeks) and 43 preterm (31-36 weeks) neonates screened for suspected sepsis, but found to have no clinical, microbiologic or hematologic evidence of sepsis. Morphologic evaluation of red cells and leukocytes was performed using European Working Group in Childhood (EWOG-MDS) criteria[18]. Giant platelets and megakaryocyte fragments were defined by their diameter: giant platelets- diameter >4μM but <8μM; megakaryocyte fragments- diameter >8μM[19]. Polycythemia was defined as hematocrit >0.65[20] and thrombocytopenia as platelets $<150 \times 10^9/L$[21,22] as per current guidelines.

TAM: Diagnosis of TAM (n=17) was made clinically by local teams in 13/17 or by OIDSCS review of CBC and blood smears (4/17). As the WHO classification for TAM[15,23] mentions only 'increased' peripheral blood blasts without defining normal ranges in neonates with or without DS, we used two approaches to define 'increased' blast cells in TAM. First, the

normal neonatal range for circulating blasts at Imperial College Healthcare NHS Trust (see above) for healthy term and preterm neonates without DS (range 0-4%; median 0%). Second, to allow for possible effects of sepsis on blast counts, we performed manual blast counts on blood smears from 80 sick preterm neonates without DS referred for hematologic review because of suspected sepsis at Imperial College Healthcare NHS Trust (range 0-8%; median 1%). We then prospectively defined TAM in OIDSCS as neonates with DS and peripheral blood blasts >10% and a *GATA1* mutation detected by Ss/DHPLC analysis followed by NGS.

Mutational analysis of the *GATA1* gene

Conventional Sanger sequencing (Ss) and Direct High Pressure Liquid Chromatography (DHPLC) Analysis: GATA1 analysis was performed on all consented peripheral blood samples (188/200; 94%) by Ss/DHPLC (WAVE; Transgenomic, Omaha, NE) as previously described[24].

Targeted next-generation-resequencing (NGS): NGS of *GATA1* exon 2 was performed on an independent aliquot of DNA from all cases with available DNA (n=104). Samples with targeted NGS-identified mutations were confirmed by pyrosequencing in a third independent DNA aliquot (Supplemental Methods). *GATA1* exon 2 was amplified using Phusion High Fidelity DNA Polymerase (NEB,UK) with forward 5'-TTTGAGAAGCTTAAAGGAGGGAAGAGGAGCAG-3' and reverse 5'-TTTGAGAAGCTTCCAGCCATTTCTGA-3' primers. PCR conditions were: 98°C for 30 seconds, 35 cycles of 98°C for 10 seconds, 66.3°C for 30 seconds and 72°C for 15 seconds. After the last cycle additional steps were 72°C for 5 minutes and 4°C for 10 minutes. PCR products were purified (PCR purification kit Qiagen, Hilden, Germany) and quality checked (Agilent 2100 Bioanalyser Agilent Technologies, Waldbronn, Germany). PCR products were digested overnight with *Hind*III followed by another round of purification. Purified product was ligated with T4 ligase overnight at 16ºC. 1 µg of ligated DNA was fragmented (Covaris S2, Covaris, Massachusetts, USA). After shearing, fragment distribution was measured (Agilent

2100 Bioanalyser). Libraries were constructed using the NEB Next DNA Sample Prep Master Mix Set 1 Kit (NEB) and ligated with 3 µl of Illumina Adapters. Ligated libraries were size-selected using Agencourt AMPure magnetic beads (Beckman Coulter, California, USA) and purified fragment distribution measured (Agilent 2100 Bioanalyser). Each library was PCR-enriched with 25 µM each of the following primers:

Multiplex PCR primer 1.0: AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT

Index primer: CAAGCAGAAGACGGCATACGAGAT[INDEX]CAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

Eight base-pair index tags were developed and validated at the Wellcome Trust Centre for Human Genetics (Oxford). Enrichment and adapter extension of each preparation was obtained using 5µl of size-selected library in a 50-µl PCR reaction. Cycling conditions were as recommended by Illumina Inc. (California, USA). After 10 cycles of amplification, product was purified using AMPureXp beads (Beckman Coulter, California, USA). Final size distribution was determined using a Tapestation 1DK system (Agilent Technologies, Waldbronn, Germany). Library concentration was determined by Agilent qPCR Library Quantification Kit on a MX3005P instrument (Agilent Technologies, Waldbronn, Germany). Sequencing was performed on an Illumina HiSeq2000 as 50 or 100 base paired end reads.

A custom Perl script was used to filter unprocessed reads (fastq files) by selecting reads with an Illumina Phred score for each base of >20. Filtered reads were mapped to *GATA1* exon 2 reference sequence (NCBI reference NT_079573.4 (*Homo sapiens* chromosome X genomic contig, starting position 11496706) using Novoalign (http://www.novocraft.com). SAMtools[25], R (http://www.r-project.org) and custom Perl scripts were used to generate and plot base pair frequencies within exon 2 using the mapped sequence reads. The resulting plots were used to identify possible mutations. Reads covering the region of mutation were then analyzed to establish their sequence. To quantitate mutant clone size, a custom Perl script was used to analyse original unprocessed reads by filtering using less stringent parameters (based on an average Phred score of >20 across the whole read) and counting numbers of reads containing mutated versus wild type sequence. Sensitivity and specificity of NGS were tested using serial DNA dilutions of a male ML-DS cell line with a hemizygous *GATA1* mutation and

normal cord blood (Supplemental Table 1). Mutation quantitation from NGS (3-5x10^5 mapping reads analysed/sample) was compared with pyrosequencing. The limit of detection of mutant *GATA1* sequence was ~1% by pyrosequencing and ~0·3% by NGS.

Blast cell immunophenotyping: Mononuclear cells (MNC) were isolated by Ficoll density gradient separation from peripheral blood samples from 7 DS neonates (2 with silent TAM and 5 with no *GATA1* mutations detected by NGS), 8 anonymised samples from neonates without DS and 7 samples with a confirmed diagnosis of TAM as defined in the OIDSC study (see Supplemental Methods).

Statistical analysis: Statistical analysis was performed using Prism software. Populations were compared using one-way ANOVA, 2-sided T-tests or Fisher's exact test; p-values <0·05 indicate statistically significant differences.

Results

Peripheral blood blasts and *GATA1* mutation analysis by Ss/DHPLC in DS neonates

Previous studies defined TAM using either 'increased' peripheral blood blasts, where blast % either was not defined or varied (as in the WHO classification[15]), or detection of *GATA1* mutations by standard Ss/DHPLC[-7,8]. Neither approach has been tested in prospective studies and the relation between blast % and *GATA1* mutation status is unknown. Therefore, to identify consistent criteria for diagnosis of TAM, we evaluated blasts on peripheral blood smears and correlated this with *GATA1* analysis by Ss/DHPLC.

Surprisingly, 97.5% of DS neonates (195/200) had blasts on blood smears (range 1-77%) (Figure 1A,B). By contrast, only 8.5% (17/200) had *GATA1* mutations by Ss/DHPLC (Figure 1C); all 17 had blasts >10% (Table 1; Figure 1A). Blast cell frequency in neonates without *GATA1* mutations was lower (median 4%; p<0.0001) but 6 neonates without *GATA1* mutations had blasts >10% (11-15%) (Figure 1A; Supplemental Table 2). None of these 6 neonates had clinical features associated with TAM and no exon 2 *GATA1* mutations were detected even using targeted NGS (Supplemental Table 3). While these cases may carry clones with large *GATA1* deletions or low abundance *GATA1* exon 3 mutations, the low frequency of such mutations reported in other studies[24,] makes this unlikely. None of these 6 individuals has developed ML-DS (median follow-up 35 months). Blast cell morphology was indistinguishable between those with or without *GATA1* mutations (Figure 1B). Thus, although a blast threshold of >10% had 100% sensitivity for detection of *GATA1* mutations by Ss/DHPLC, the specificity of this diagnostic criterion was only 74%. Since automated hematology analyzers also failed to identify blasts in 11/17 neonates with *GATA1* mutations, we suggest that a practical and sensitive definition of TAM is the presence of blasts >10% on blood smears and a *GATA1* mutation detected by Ss/DHPLC. Using this definition, we determined the clinical and hematologic features of DS neonates with TAM compared to those without TAM.

Clinical characteristics (Table 2)

13/17 (76.5%) neonates with TAM (as defined above) presented with clinical or hematologic signs that led to a clinician-diagnosis of TAM, which was subsequently confirmed by OIDSCS criteria. The remaining 4/17 cases were unsuspected by clinical teams and diagnosed through OIDSCS blood smear review with increased blasts (16-41%) and *GATA1* mutation analysis. Hepatosplenomegaly, jaundice and rash were more common in TAM but were not specific to TAM (Table 2). Thus, TAM cannot be recognised by clinical signs alone. The spectrum of congenital abnormalities was similar to previous reports[26].

Hematologic data (Table 2)

All DS neonates had abnormalities of $\geq$1 hematopoietic lineage and 95·0% had abnormalities in $\geq$3 lineages.

***Erythrocyte abnormalities**:* Overall, DS neonates had higher hematocrit (Hct), hemoglobin (Hb) and circulating erythroblasts than normal. Although median Hct (p=0.0122) and Hb (p=0.0011) were slightly lower in TAM than DS neonates without TAM, there was considerable overlap and only one neonate with TAM was anemic (Figure 1D; Table 2; Supplemental Table 2). Dyserythropoiesis was common in DS neonates (increased MCV, macrocytes, dyserythropoietic erythroblasts, target cells and basophilic stippling) (Figure 1E; Table 2) and was not significantly greater in TAM (Supplemental Table 4). There was no correlation between Hct, MCV or erythroblasts and heart disease, intrauterine growth restriction (IUGR) or maternal complications (Supplemental Table 5). This suggests that abnormal erythropoiesis in DS neonates is a consequence of trisomy 21 rather than *GATA1* mutation or secondary causes.

Platelet abnormalities: Median platelet counts were lower than normal in DS and not reduced further in TAM (p=0.9014) (Table 2; Figure 1F). The frequency of thrombocytopenia (platelets <150x10^9/L) was similar in neonates with and without TAM (p=0.6162) and in DS neonates with or without potential secondary causes (e.g. sepsis)[27] (Supplemental Table 2),

although median platelets were slightly lower in DS neonates with IUGR (Supplemental Table 5). Median MPV was similar in neonates with and without TAM (Table 2). Platelet morphology was abnormal in 193/200 (96.7%) DS neonates (giant platelets, circulating megakaryocytes and/or megakaryocyte fragments) (Figure 1G) consistent with trisomy 21-mediated effects on platelet production. Megakaryocyte fragments were strongly associated with TAM (16/17 [94.1%]) v 84/183 [45.9%] DS neonates without TAM (p=0.0002) (Supplemental Table 4). Although not specific, absence of megakaryocyte fragments had a negative predictive value for TAM of 99.0%.

Leukocyte abnormalities: Although leukocyte counts were higher in TAM than DS neonates without TAM (25.8 v 14·1x10^9/L;p<0·0001), in 9/17 (52.9%) neonates with TAM leukocyte counts were in the normal range (Figure 1H). Neutrophils, myelocytes, basophils, monocytes and blasts were increased in neonates with and without TAM (Table 3). Leukocytosis was not due to sepsis as only 10/200 had culture-positive sepsis (Supplemental Table 5). Dysplastic neutrophils and monocytes (EWOG-MDS criteria[18]) were common in DS neonates with and without TAM, including hypogranular neutrophils, pseudo-Pelger forms, stellate monocytes and dysplastic basophils (Figure 1I; Supplemental Table 4) implicating trisomy 21, rather than mutant *GATA1*, in these changes.

***GATA1* mutation analysis by next-generation-resequencing (NGS)**

The mutant/wild-type fraction required to identify *GATA1* mutant clones is ~30% by Ss and ~5% by DHPLC (Table 1 and unpublished data). To determine whether smaller-sized mutant *GATA1* clones explained the high prevalence of blasts in DS neonates without TAM we performed targeted NGS of *GATA1* exon 2 (97% of *GATA1* mutations are in exon 2[24]). The flow diagram in Figure 2A outlines the methodology. The bioinformatic pipeline for mapping reads (3-5 x 10^5 mapped reads analysed/sample), identifying mutations and quantitating mutant/wild-type fraction is described in the Methods. To confirm mutation identification/quantitation by NGS, independent DNA aliquots were analyzed by pyrosequencing. Detection limits of *GATA1* mutation were ~1-2% by pyrosequencing and

~0·3% by NGS (Supplemental Table 1). *GATA1* mutations were not detected in normal cord blood controls (n=2) included in each run (data not shown).

All samples with sufficient DNA were studied by targeted NGS. In DNA from 104 DS neonates (16/17 TAM; 88/169 non-TAM), mutations were detected in 34 cases: 16/16 with TAM and 18/88 (20.4%) neonates without TAM where direct Ss/DHPLC had not detected *GATA1* mutations (Figure 2B-D; Supplemental Table 3). Pyrosequencing confirmed *GATA1* mutations in 17/34 (data not shown). In the remaining 17/34, mean estimated mutant clone size was 1·5% (range 0·245%-5·96%) consistent with *GATA1* mutant fractions below the sensitivity of pyrosequencing detection. The size of the *GATA1* mutant clone correlated with blast % ($p<0.0001$; Figure 1E) from the same sample.

Clinical and hematologic features of 'silent TAM': Clinical and hematologic features of the 18/88 neonates with *GATA1* mutant clones identified only by NGS were indistinguishable from the 70 DS neonates without detectable *GATA1* mutations by NGS (Table 3). We therefore suggest the term 'silent TAM' where *GATA1* mutations are detectable by NGS but not by Ss/DHPLC. As expected, mutant clones were smaller and blast frequency lower in 'silent TAM' compared to TAM (Figure 1E). Since blasts % in 'silent TAM' (median 5%) was similar to DS neonates negative for NGS-detected *GATA1* mutations (median 4%) (Figure 2F), morphology-based enumeration of blasts cannot distinguish 'silent TAM' from babies without a *GATA1* mutation. Preliminary data suggest that immunophenotyping of circulating blasts from DS neonates using standard diagnostic panels[11,28] also fails to identify 'silent TAM' (Supplemental Table 6).

ML-DS in DS neonates with TAM or 'silent TAM': After a median follow-up of >33 months, ML-DS was diagnosed in 4 cases: 3/17 with TAM (DST5, DST10, DST12) and 1/18 with 'silent TAM' (DS108). None of the 70 neonates without NGS-detected *GATA1* mutations has developed ML-DS (Table 3). In 3/4 ML-DS cases a single mutant *GATA1* mutation was detected by NGS at birth and the same DNA substitution was present at diagnosis of ML-DS;

in 1/4 (DST12) both mutations detected at birth were also found at diagnosis of ML-DS (Table 1), confirming the clonal relationship between TAM ('silent' or overt) and ML-DS. Only 1 further case of ML-DS was diagnosed from the 18 study centers during the study period. It is not known whether or not this child had clinically or hematologically 'silent' TAM since they were not recruited to OIDSCS as the family moved abroad and neither *GATA1* mutation analysis nor blood smear review were performed at birth.

Discussion

OIDSCS is the first prospective study to systematically determine blood counts, blood cell morphology and *GATA1* mutation status in neonates with DS with the aims of (i) identifying the population at risk of developing ML-DS; (ii) defining the clinical and hematologic features associated with a mutant *GATA1* clone in DS neonates; and thereby, (iii) how best to define TAM. We demonstrate for the first time that almost all DS neonates have multiple quantitative and morphologic hematologic abnormalities, independent of their *GATA1* mutation status, providing strong correlative evidence that trisomy 21 itself causes trilineage perturbation of neonatal, as well as fetal, hematopoiesis[29-34].

GATA1 mutation analysis showed that 8.5% of babies had a *GATA1* mutation detected by Ss/DHPLC, similar to estimates from retrospective studies (5-10%)[10]. Since an additional 20.4% (18/88 neonates with sufficient available DNA) had *GATA1* mutations detectable only by targeted NGS, the overall frequency of *GATA1* mutations was 3-fold higher (29%) than previous estimates; importantly, two-thirds of cases were clinically and hematologically 'silent'. Thus, *GATA1* mutation analysis using NGS is the most reliable way of detecting all babies with mutant *GATA1* clones. In keeping with our previous work[3], a significant proportion of neonates (5/35) had >1 *GATA1* mutation suggesting that the N-terminal truncated GATA1 protein confers a selective growth advantage to mutant *GATA1*-containing clones or, alternatively, that trisomic cells have a 'mutator phenotype'[35,36]. Although the low frequency of non-hematologic malignancies in DS argues against this, trisomy 21 might induce high mutation rates at specific genomic loci (e.g. *GATA1*), as in the kataegis phenotype[37], rather than causing widely-dispersed mutations.

What implications do these data have for the definition and diagnosis of TAM? Most clinicians screen DS newborns with a CBC and use clinical and hematologic findings to flag possible diagnoses of TAM, which then have *GATA1* mutation analysis by conventional Ss/DHPLC. Our data show that TAM cannot be reliably diagnosed by clinical or hematologic features alone. Furthermore, no hematologic features are specific for TAM. Indeed 4/17 cases of TAM in our study were unsuspected by clinical teams either because a normal CBC led to routine smears not being evaluated or because blasts were wrongly attributed to 'prematurity'. This highlights the practical difficulty in defining TAM solely by a specific % of blasts, especially given inter-observer variation in blast cell enumeration on smears[38]. The threshold of 10% used in OIDSCS was a useful, rapid way of identifying all possible cases but was not specific and

our data (Figure 1) indicate that even if a larger cohort was studied, it would be difficult to define TAM with high sensitivity and specificity on a blast % threshold .

An alternative way of considering the issue of the diagnosis of TAM is to ask why it is important to diagnose TAM? First, an accurate definition of TAM allows the population at risk of transforming to ML-DS to be identified. This facilitates regular clinical and laboratory follow up by pediatric hemato-oncologists and ensures appropriate management of cytopenias that may precede ML-DS, including the timing of anti-leukemic therapy. Given the clear etiologic link between *GATA1* mutations and ML-DS[3-5,7-9], it is not surprising that babies with *GATA1* mutations detectable by Ss/DHPLC or NGS were both at risk of ML-DS. Our data clearly show that the best way to identify all those at risk is to comprehensively screen for *GATA1* mutations by Ss/DHPLC and NGS. The second reason for making a definitive diagnosis of TAM is that it may facilitate management of early complications of TAM, such as effusions and liver dysfunction. In fulminant cases, the diagnosis is usually clinically straightforward, the blast % is high and *GATA1* mutations should be easy to detect by Ss/DHPLC as the circulating disease burden is high. Our findings may have a modest impact on the immediate management of such cases of “classical” TAM but will prevent mis-diagnosis in neonates with clinical and hematologic features mimicking TAM (Table 3).

We favor a diagnostic algorithm (Figure 3) where evaluation of blood smears, as well as CBCs, as recommended by the American Academy of Pediatrics[39], is a useful, immediate screening step to identify DS neonates with "classical" TAM who may require early treatment (especially where *GATA1* analysis is unavailable or delayed). We suggest all DS neonates should also have *GATA1* mutation analysis by Ss/DHPLC to quickly identify those with large mutant *GATA1* clones. For DS neonates without mutations detected by Ss/DHPLC, we suggest NGS is used to identify low abundance *GATA1* mutations. By comprehensively detecting *GATA1* mutations, pediatric hematology follow-up can be limited to those at risk of transformation rather than all babies with blasts (i.e. ~all babies with DS).

Using this diagnostic approach, *GATA1* mutations will frequently be clinically and hematologically 'silent'. Therefore, we suggest the term 'silent TAM' for DS neonates with small *GATA1* mutant clones, detectable only by NGS. This approach illustrates the dilemma facing clinicians in the current era of high-throughput genetic sequencing. In describing a new disease entity, 'silent TAM', we acknowledge that we cannot yet know the best management for these infants. However, we know 'silent' TAM is

clinically important since it can give rise to ML-DS. Furthermore, TAM and 'silent' TAM offer a unique opportunity to address the significance of small preleukemic clones in DS. As ML-DS virtually always presents within a predictable time window spanning the first 5 years of life[40], we suggest that all DS neonates with *GATA1* mutations should be monitored until age 5 years with regular CBCs, smears and *GATA1* mutation analysis by NGS. We acknowledge that it is currently unknown whether this would allow early diagnosis and treatment of ML-DS or improve outcome. While CBC data, particularly platelet count, may provide a simpler, albeit less specific, early indicator of impending ML-DS, whether this would be equally effective in improving outcome is also unknown. Nevertheless, monitoring of mutant *GATA1* clones using such sensitive methods at last provides the opportunity to design effective protocols to eradicate *GATA1* mutant clones during the neonatal period and potentially prevent ML-DS. NGS, rather than Ss/DHPLC is required to pick up small clones at birth and at follow up. NGS technology is now widely available. Even though it is becoming cheaper, the suggested approach would limit NGS to those who do not have mutations detected by a simpler method. In the future, it may be that NGS becomes the most cost-effective method to detect *GATA1* mutations.

Finally, what is the risk of transformation conferred by harboring *GATA1* mutations? In previous retrospective studies of clinically diagnosed TAM, the risk of ML-DS was ~20-30%[11-14]. Given that population studies show that 1-2% of children with DS develop ML-DS[1], and that we found *GATA1* mutations in 29% of DS neonates, this suggests that *GATA1* mutation(s) may confer a risk of ML-DS of ~5-10%. This is consistent with our data in which 4/35 (11.4%) neonates with *GATA1* mutation(s) have developed ML-DS with median follow-up of >33 months. By contrast, none of the neonates without a *GATA1* mutation has developed ML-DS despite median follow-up of 40 months. Ultimately, longitudinal follow-up of all DS neonates in OIDSCS will allow accurate assessment of the relationship between mutant *GATA1* clone size and development of ML-DS and, thus, fully define the natural history of TAM and 'silent TAM'.

Acknowledgements

The authors thank all the parents who generously allowed their children to be included in the study and the haematology laboratory, nursing, medical and clerical staff in all the participating centres for their valuable contributions. IR, PV, KA, GJ, KP, GC and HR are funded by Leukaemia & Lymphoma Research (LLR) Specialist Programme Award 08030 and AN and AR by LLR Clinical Fellowships. IR is also funded by the Imperial College London Comprehensive Biomedical Research Centre and PV by the NIHR Biomedical Research Centres funding scheme, the Medical Research Council (MRC) Disease Team Award and the MRC Molecular Haematology Unit. NB is funded by the Kay Kendall Leukaemia Fund and Leuka. AG is funded by the Wellcome Trust ((Programme grant 091182). The funding sources had no role in study design or conduct and management of the study, and no role in data collection, data analysis, data interpretation, or writing of the report. IR and PV had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. IR, PV and GH contributed to all parts of the study; KA, GJ, HR, AN, GV, KP, EM, SMG, AR, GC, CS, NB, AG and PC performed the research, analysed data and revised the paper. All other authors contributed patient samples, acquired and interpreted data and revised the paper.

Conflict of interest:

The authors have no conflicts of interest to declare.

References

1. Hasle H, Clemmensen IH, Mikkelsen M. Risks of leukaemia and solid tumours in individuals with Down's syndrome. Lancet. 2000; 355: 165-9.

2. Lange BJ, Kobrinsky N, Barnard DR et al. Distinctive demography, biology, and outcome of acute myeloid leukemia and myelodysplastic syndrome in children with Down syndrome: Children's Cancer Group Studies 2861 and 2891. Blood. 1998; 91: 608-15.

3. Ahmed M, Sternberg A, Hall G et al. Natural history of GATA1 mutations in Down syndrome. Blood. 2004; 103: 2480-9.

4. Hitzler JK, Cheung J, Li Y, Scherer SW, Zipursky A. GATA1 mutations in transient leukemia and acute megakaryoblastic leukemia of Down syndrome. Blood. 2003; 101: 4301-4.

5. Rainis L, Bercovich D, Strehl S et al. Mutations in exon 2 of GATA1 are early events in megakaryocytic malignancies associated with trisomy 21. Blood. 2003; 102: 981-6.

6. Wechsler J, Greene M, McDevitt MA et al. Acquired mutations in GATA1 in the megakaryoblastic leukemia of Down syndrome. Nat Genet. 2002; 32: 148-52.

7. Xu G, Nagano M, Kanezaki R et al. Frequent mutations in the GATA-1 gene in the transient myeloproliferative disorder of Down syndrome. Blood. 2003; 102: 2960-8.

8. Groet J, McElwaine S, Spinelli M et al. Acquired mutations in GATA1 in neonates with Down's syndrome with transient myeloid disorder. Lancet. 2003; 361: 1617-20.

9. Taub JW, Mundschau G, Ge Y et al. Prenatal origin of GATA1 mutations may be an initiating step in the development of megakaryocytic leukemia in Down syndrome. Blood. 2004; 104: 1588-9.

10. Zipursky A. Transient leukaemia - a benign form of leukaemia in newborn infants with trisomy 21. Br J Haematol. 2003; 120: 930-8.

11. Klusmann JH, Creutzig U, Zimmermann M et al. Treatment and prognostic impact of transient leukemia in neonates with Down syndrome. Blood. 2008; 111: 2991-8.

12. Massey GV, Zipursky A, Chang MN et al. A prospective study of the natural history of transient leukemia (TL) in neonates with Down syndrome (DS): Children's Oncology Group (COG) study POG-9481. Blood. 2006; 107: 4606-13.

13. Muramatsu H, Kato K, Watanabe N et al. Risk factors for early death in neonates with Down syndrome and transient leukaemia. Br J Haematol. 2008; 142: 610-5.

14. Gamis AS, Alonzo TA, Ryan EA et al. Natural history of transient myeloproliferative disorder clinically diagnosed in Down syndrome neonates: a report from the Children's Oncology Group Study A2971. Blood 2011; 118: 6752-9.

15. Vardiman JW, Thiele J, Arber DA et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood. 2009; 114: 937-51.

16. Henry E, Walker D, Wiedmeier SE, Christensen RD. Hematological abnormalities during the first week of life among neonates with Down syndrome: data from a multihospital healthcare system. Am J Med Genet A. 2007; 143: 42-50.

17. Pine SR, Guo Q, Yin C et al. Incidence and clinical implications of GATA1 mutations in newborns with Down syndrome. Blood. 2007; 110: 2128-31.

18. Cantu Rajnoldi A, Fenu S, Kerndrup G et al. Evaluation of dysplastic features in myelodysplastic syndromes: experience from the morphology group of the European Working Group of MDS in Childhood (EWOG-MDS). Ann Hematol. 2005; 84: 429-33.

19. Noris P, Pecci A, di Bari F et al. Application of a diagnostic algorithm for inherited thrombocytopenias to 46 consecutive patients. Haematologica 2004; 89: 1219-25.

20. Sarkar S, Rosenkrantz TS. Neonatal polycythemia and hyperviscosity. Semin Fetal Neonatal Med. 2008; 13: 248-55.

21. Roberts I, Stanworth S, Murray NA. Thrombocytopenia in the neonate. Blood Rev. 2008; 22: 173-86.

22. Stanworth SJ, Clarke P, Watts T, et al. Prospective, observational study of outcomes in neonates with severe thrombocytopenia. Pediatrics. 2009; 124: e826-34.eonatal medicine. 2008; **13**(4): 248-55.

23. Hasle H, Niemeyer CM, Chessells JM, et al. A pediatric approach to the WHO classification of myelodysplastic and myeloproliferative diseases. Leukemia. 2003; 17: 277-82.

24. Alford KA, Reinhardt K, Garnett C et al. Analysis of GATA1 mutations in Down syndrome transient myeloproliferative disorder and myeloid leukemia. Blood. 2011; 118: 2222-38.

25. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25(:2078-9.

26. Rankin J, Tennant PW, Bythell M, Pearce MS. Predictors of survival in children born with Down syndrome: a registry-based study. Pediatrics. 2012; 129: e1373-81.

27. Chakravorty S, Roberts I. How I manage neonatal thrombocytopenia. Br J Haematol. 2012; 156: 155-62.

28. Langebrake C, Creutzig U, Reinhardt D. Immunophenotype of Down syndrome acute meyloid leukemia and transient myeloproliferative disease differs significantly from other diseases with morphologically identical or similar blasts. Klin Pediatr 2005; 217: 126-34.

29. Tunstall-Pedoe O, Roy A, Karadimitris A et al. Abnormalities in the myeloid compartment in Down Syndrome fetal liver precede acquisition of GATA1 mutations. Blood. 2008; 112: 4507-11.

30. Chou ST, Opalinska JB, Yao Y et al. Trisomy 21 enhances human fetal erythro-megakaryocytic development. Blood. 2008; 112: 4503-6.

31. Roy A, Cowan G, Mead AJ et al, Perturbation of fetal liver hematopoietic stem and progenitor cell development by trisomy 21, Proc Natl Acad Sci U S A, 2012; 109: 17579-17584.

32. Maclean, GA, Menne TF, Guo G et al, Altered hematopoiesis in trisomy 21 as revealed through in vitro differentiation of isogenic human pluripotent cells, Proc Natl Acad Sci U S A 2012; 109: 17567-17572.

33. Chou ST, Byrska-Bishop M, Tober MJM et al. Trisomy 21-associated defects in human primitive hematopoiesis revealed through induced pluripotent stem cells. Proc Natl Acad Sci U S A 2012; 109: 17573-17578.

34. Birger Y, Goldberg L, Chlon TM et al. Perturbation of fetal hematopoiesis in a mouse model of Down syndrome's transient myeloproliferative disorder. Blood 2013; Epub May 29

35. Amiel A, Goldzak G, Gaber E, Fejgin MD. Molecular cytogenetic characteristics of Down syndrome newborns. J Hum Genet. 2006; 51: 541-47.

36. Nizetic D, Groet J. Tumorigenesis in Down's syndrome: big lessons from a small chromosome. Nat Rev Cancer 2012;12: 721-32.

37. Nik-Zainal, S, Alexandrov LB, Wedge DC et al. Mutational processes molding the genomes of 21 breast cancers. Cell 2012; 149: 979-93.

38. Senent L, Arenillas L, Luno E, Ruiz JC, Sanz G, Florensa L. Reproducibility of the World Health Organization 2008 criteria for myelodysplastic syndromes. Haematologica 2013 Epub

39. Bull MJ. Health supervision for children with Down syndrome. Pediatrics. 2011; 128: 393-406.

40. Hasle H, Abrahamsson J, Arola, M et al. (2008). Myeloid leukemia in children 4 years or older with Down syndrome often lacks GATA1 mutation and cytogenetics and risk of relapse are more akin to sporadic AML. Leukemia *22*, 1428-1430.

Figure legends:

Figure 1: Hematologic abnormalities and *GATA1* mutation analysis by Ss/DHPLC in neonates with Down Syndrome

(A) Percentage of blasts on blood films from the first week of life in 200 neonates with DS, 17 with TAM (red circles), 183 without TAM (black circles). (B) Photomicrographs of typical blast cells in a neonate with TAM (top) and in a DS neonate without TAM (bottom). (C) *GATA1* mutation analysis in TAM by Sanger sequencing (Ss) and Denaturing High Pressure Liquid Chromatography (DHPLC). (Ci,ii) Mutation analysis of sample DST11. The mutation is detected by both Ss and DHPLC. (Ci) Sanger sequence trace. The arrow points the start of a double sequence trace indicative of an acquired *GATA1* mutation. (Cii) DHPLC trace from the same sample (Red line = mutant, Black line = normal). (Ciii, Civ) Mutation analysis of sample DST9. The mutation is detected by DHPLC but not by Ss. (Ciii) Sequence trace. (Civ) DHPLC trace from the same sample (Red line = mutant, Black line = normal). (D,F,H) Scatter graphs of hematocrit (D), platelet counts (F) and leukocytes (H) in 200 DS neonates in the first week of life, 17 with TAM (red circles) and 183 without TAM (black circles). The horizontal lines show the upper and/or lower limits of the normal neonatal laboratory range (see Supplemental methods). (E,G,I) Photomicrographs of erythrocyte (E), platelet (G) and leukocyte (I) morphologic abnormalities in neonates with DS. (E) Top left: macrocytes (black arrowheads); top right: target cells (white arrowheads); bottom left: dyserythropoietic erythroblasts (fine black arrow); bottom right: basophilic stippling (grey arrow). (G) Examples of giant platelets (GP)(black arrowhead) and megakaryoblasts (white arrowhead), megakaryocyte fragments (MK fragment) and circulating megakaryocytes (MKs) in blood films from DS neonates without TAM (top row) and with TAM (bottom row). (I) Top left: hypogranular neutrophil; top right pseudo-Pelger neutrophil; bottom left: monocyte with stellate nucleus; bottom right dysplastic basophil. Scale bars indicate 10 microns.

Figure 2: *GATA1* mutation analysis in DS neonates with TAM and 'silent TAM'

(A) Flow diagram of preparation and analysis of samples for deep sequencing. (B) Pie charts of *GATA1* mutation analysis of the 200 babies in the cohort by standard Ss/DHPLC (left) and next-generation-resequencing (NGS) (right). (C) Examples of base pair plots from next-generation-resequencing analysis of patient samples (mutation indicated by arrows) with (D) corresponding pyrosequencing traces below

(mutant peaks indicated by arrows). On the X-axis is the position along the *GATA1* exon 2 amplicon (432 base pairs). On the Y-axis is the read depth at different positions along the amplicon. Therefore, the black line trace shows the number of reads mapping to *GATA1* sequence at different positions along the amplicon. At the position of the black arrowhead a mutation was introduced into the PCR primer (mapping outside *GATA1* exon 2) so that all PCR products would have a unique tag. This introduced mutation is detected by the blue line. All PCR products have this introduced mutation as the height of the blue line is to the level of black trace (total number of mapping reads). Sequence analysis also shows there are two common single nucleotide polymorphism (SNPs) at positions rs62600348 T>G and rs66717003 T>G (indicated by the star) in the amplicon that map to position 48649449 and 48649456 within GATA1 exon 2. Nucleotide 0 is the first nucleotide of *GATA1* exon 1 including exons and introns. NCBI reference NT_079573.4 (*Homo sapiens* chromosome X genomic contig, starting position 11496706) was used. The location of *GATA1* mutation (indicated by black arrow). (Ci, Di) Patient sample DST11 with a 7 bp duplication at position 48649625 previously detected by Ss/DHPLC. (Cii, Dii) Patient sample DST9 with an insertion of 7 bp at position 48649670 previously detected by DHPLC only. (Ciii, Diii) Patient sample DS158 with a 2 bp deletion at position 48649552 detected by NGS only and confirmed by pyrosequencing. (Civ, Div) Patient sample DS79 with a point mutation at position 48649565 detected by NGS only but not detectable by pyrosequencing. (E) Relationship between *GATA1* mutant clone size (Y-axis) as determined by next-generation-resequencing and % blasts detected by morphology. (F) Distribution of % blasts in TAM (n=17) (filled red circles (left), Silent TAM (n=88) (open red cell circles (middle) and in samples without a *GATA1* mutation detected by next-generation-resequencing (n=70) (filled black circles) right.

Figure 3: Algorithm for diagnosis and monitoring of mutant *GATA1* clones in DS neonates

Suggested algorithm for diagnosis and monitoring of mutant *GATA1* clones in DS neonates. Evaluation of a blood smear and CBC can be used as an immediate screening step to identify DS neonates with "classical" TAM who may require early treatment (especially where *GATA1* analysis is unavailable or delayed). As a next step *GATA1* mutation analysis by Ss/DHPLC will quickly identify DS neonates with large mutant *GATA1* clones. For DS neonates without mutations detected by Ss/DHPLC, NGS is the most reliable way of identifying low abundance *GATA1* mutations, allowing pediatric hematology follow-up to be limited to those at risk of transformation rather than all DS babies with peripheral blood blasts. <u>Monitoring of all DS</u>

children with *GATA1* mutations until age 5 years is recommended. This can be done using serial CBC/smears with *GATA1* mutation analysis as indicated, eg for persistent cytopenias. For the small number of DS babies with blasts >10% who have no detectable *GATA1* mutations by NGS, more detailed studies to exclude the presence of rare *GATA1* deletions are suggested.

DS: Down syndrome

CBC: Complete blood count

Ss/DHPLC: Sanger sequencing/ Denaturing High Performance Liquid Chromatography

NGS: Targeted next-generation-resequencing

Table 1:
***GATA1* mutation analysis and peripheral blood blasts in DS neonates with TAM**

Patient	Mutation detection by Ss/DHPLC	*Nature and position of mutation detected by Ss	Nature and position of mutation detected by NGS[a]	Effect of mutation	Blood blast %	Clone size by NGS (%)	Clinical features
DST1	Ss and DHPLC	G>C point mutation, position 48649737	Confirmed	Loss of splice donor site	17	56.7	Jaundice Hepatospleno-megaly IUGR A&W 48m
DST2	Ss and DHPLC	G>A point mutation, position 48649737	Confirmed	Loss of splice donor site	77	45.28	CHD Jaundice Hepatomegaly A&W 30m
DST3	Ss and DHPLC	Deletion 17bp GCGGCA CTGGCCT ACTA, position 48649688	Confirmed	Frame shift	37	49.63	CHD Jaundice Rash Bacterial sepsis A&W 33m
DST4	Ss and DHPLC	dup 20bp 48649670	20 bp duplication ACAGCCACC GCTGCAGCT GC, position 48649670	Frame shift	35	5.96	Jaundice Preterm (Gestation at birth 34 weeks) A&W 32m
DST5	Ss and DHPLC	A>T mutation at position 48649739	Confirmed	Loss of splice donor site	40	27.9	CHD Jaundice Hepatospleno-megaly Coagulopathy Transient spontaneous remission Sudden clinical deterioration with increasing blasts age 2m Diagnosed as TAM/ML-DS Rx: AraC CCR 24m
DST6	DHPLC only	N/A	Clone1: T>A mutation at position 48649738 Clone 2: 1 bp deletion (G) at position 48649666	Clone1 = Loss of splice donor site Clone2 = frame shift	32	10.49	CHD Jaundice A&W 57m

DST7	Ss and DHPLC	G>A point mutation, position, 48649520	Confirmed	Prem-ature stop codon	41	3.79	Jaundice Hepatospleno-megaly Effusions Bacterial sepsis IUGR Rx Ara C A&W 30m
DST8	DHPLC only	N/A	G>T point mutation, position 48649715	Prem-ature stop codon	15	15.7	IUGR A&W 45m
DST9	DHPLC only	N/A	7 bp insertion GGTGAGC position 48649670	Frame shift	20	3.81	Nil of note A&W 22m
DST 10	Ss and DHPLC	Insertion CAGTGCCTACT, position 48649704	Confirmed	Frame shift	42	15.48	Jaundice Spontaneous remission by 6 weeks ML-DS at 22 m Rx: 4 cycles AML chemoRx CCR 28m
DST 11	Ss and DHPLC	Duplication 7bp CCCCTCT, position 48649625	Confirmed	Frame shift	38	17.06	Hepato-splenomegaly Rash CHD A&W 41m
DST 12[b]	Ss and DHPLC	2 bp deletion AG, position 48649606	Clone1 = 2 bp deletion (AG), position 48649606 Clone2 = 8 bp duplication CACCGCTG position 48649675	Clone1 = Frame shift Clone2 = Frame shift	23	Clone 1: 20.21 Clone 2: 0.54	CHD Jaundice Hepatomegaly Rash Liver failure Spontaneous remission by 2 weeks. Developed ML-DS age 4 m Rx: 4 cycles AML chemoRx CCR 34m
DST 13	Ss and DHPLC	Insertion GCAGCTGGAGCACAGCC, position 48649676	Confirmed	Frame shift	50	17.07	Jaundice Hepatomegaly Liver failure Rx AraC A&W 36m
DST 14	Ss and DHPLC	C>T point mutation, position 48649565	Confirmed	Premature stop codon	73	82.31	CHD Jaundice Hepato-splenomegaly Coagulopathy Rx AraC A&W 49m
DST 15	Ss and DHPLC	Deletion 14bp GTAACTCCATTGAG, position 48649737	Confirmed	Loss of splice donor site	17	33	CHD Jaundice Hepatomegaly Coagulopathy Rx AraC A&W 43m

DST 16	Ss and DHPLC	2 bp deletion (AG), position 48649600	ND	Frame shift	33	ND	CHD A&W 51m
DST 17	DHPLC only	N/A	2 bp deletion (AG) at position 48649552	Frame shift	16	3	Jaundice A&W 37m

[a] Coordinates refer to human genome, build GRCh37 (hg19)

[b] DST12: 4 copies of *RUNX1* in 6% of BM cells at diagnosis of ML-DS by FISH (no other additional cytogenetic abnormalities were detected at diagnosis of TAM or ML-DS in the other cases).

IUGR- intrauterine growth restriction

Ss- direct Sanger sequencing

DHPLC- direct high performance liquid chromatography

CHD- congenital heart disease

A&W- alive and well

Rx- treatment

AraC- cytosine arabinoside

ML-DS- Myeloid Leukemia of Down Syndrome

CCR- complete clinical remission

N/A – Not applicable

ND - Not done

Table 2. Clinical and hematologic data of neonates recruited to the Oxford Imperial Down syndrome Cohort (OIDSC) Study

Number of neonates with DS				
	TAM (n=17)	DS without TAM (n=183)	p	
Clinical characteristics				
Gender M:F	9:8 (1·1:1)	96:87 (1·1:1)	1·0000	
Gestation at birth, weeks (range)	37·0 (34·4-39·6)	38·0 (30·9-42·6)	0·1717	
Preterm (<37 weeks)	5 (29.4%)	39 (21·3%)	0·5408	
Small-for-gestational-age (SGA)[a]	2 (11.8%)	10 (5·5%)	0·2749	
Hepato(spleno)megaly	7 (41.2%)	6 (3·3%)	<0·0001	
Jaundice	13 (76.4%)	78 (42·9%)	0·0098	
Rash	3 (17.6%)	1 (0·6%)	0·0020	
Pleural/pericardial effusion/ ascites	1 (5.9%)	3 (1·6%)	0·3010	
Congenital heart disease	8 (47.1%)	90 (49.2%)	1.0000	
Other congenital anomalies[b]	2 (11.8%)	20 (10·9%)	1.0000	
Hematologic characteristics				Normal range[c]
Median Hb g/dL (range)	18.4 (11·7-22·2)	20·5 (7·7-28·0)	0·0011	16·6 (12·7-20·3)
Median hematocrit (range)	0·562 (0·357-0·647)	0·599 (0·243-0·822)	0·0122	0·503 (0·408-0·610)
Anemia	1 (5.9%)	3 (1.6%)	0.3010	Hct ≥0.400
Median MCV (fL)	(93·3-133)	108 (88·7-123·8)	0·2104	103·5 (89·6-117·7)
Nucleated red cells/100 WBC	8 (1-122)	5 (0-186)	0·0187	1 (0-29)
Median Platelets $x10^9$/L[d]	117 (36-446)	148·5 (9-432)	0·9014	253 (150-388)
MPV (fL)[d]	9·75 (7·4-13·5)	10·6 (7·3-12·8)	0·6448	10·4 (8·1-12·5)
Thrombocytopenia	10 (54.6%)	91 (50.6%)	0.6152	Platelets. <150$x10^9$/L
WBC $x10^9$/L	25.8 (19·7-73·2)	14·1 (4·6-51·72)	<0·0001	11·6 (4·9-26·7)
Blasts (%)	35 (15-77)	4 (0-15)	<0·0001	0 (0-4)
Neutrophils $x10^9$/L	13·80 (7·0-31·0)	9·57 (0·12-38·1)	0·0372	6·2 (1·36-20·1)
Metamyelocytes $x10^9$/L	0.79 (0-5.66)	0.35 (0-5.18)	0.2986	N/A
Myelocytes $x10^9$/L	0.25 (0-3.96)	0 (0-1.94)	<0.0001	N/A

Monocytes x10^9/L	1.41 (0·3-6·7)	1·11 (0·08-6·96)	0·3082	0·78 (0-2·67)
Basophils x10^9/L	0·37 (0-2·0)	0·17 (0-2·3)	<0·0001	0 (0-0·15)
Eosinophils x10^9/L	0·24 (0-1·12)	0·22 (0-1·18)	0·9543	0·20 (0-1·87)
Lymphocytes x10^9/L	6.0 (0.23-19.2)	2.94 (0.4-8.5)	<0.0001	3.8 (0.86-9.7)

[a] SGA defined as birthweight <10th percentile for gestational age

[b] Duodenal atresia, oesophageal atresia, tracheo-oesophageal fistula, imperforate anus, Hirschsprung's, renal anomalies

[c] Normal range: derived from anonymized data from 80 normal healthy term (37-42 weeks) and 43 preterm (31-36 weeks) neonates (see Supplementary Methods)

[d] Platelet count not available in 3 neonates due to platelet clumping and automated MPV measurable for 6/17 neonates with TAM, 59/183 DS neonates without TAM and 122 neonates without DS.

N/A- not applicable

Table 3: Clinical and hematologic features of DS neonates with silent *GATA1* mutations ('silent TAM') compared to DS neonates without *GATA1* mutations by targeted NGS

	Number (%) of neonates with DS				
Clinical and hematologic characteristics	No *GATA1* mutations (by targeted NGS) (n=70)	Silent TAM (n=18)	p-value Silent TAM v no *GATA1* mutations	TAM (n=17)	p-value Silent TAM v TAM)
Gender (M:F)	30:40	10:8	0.4282	9:8	1.00
Median gestation at birth (weeks)	38.1	38.3	0.7122	37.0	0.0431
Hepatosplenomegaly	4 (5.7%)	0	0.5775	7 (341.2%)	0.0072
Jaundice	33 (47.1%)	11 (61.1%)	0.6024	13 (76.4%)	0.7283
Rash	1 (1.4%)	0	1.00	3 (17.6%)	0.2273
Pleural/pericardial effusion and/or ascites	1 (1.4%)	0	1.00	1 (5.9%)	0.4848
Congenital heart disease	37 (52.9%)	9 (50%)	1.000	8 (47.1%)	0.4935
Death ML-DS[a] Median follow up (mths)	1 (1.4%) 0 34	1 (5.6%) 1 (5.6%) 33	0.3691 0.2069 0.3587	0 (0%) 3 (17.6%) 40	1.000 0.3377 0.1043
Hct Median (range)	0.592 (0.243-0.80)	0.617 (0.509-0.736)	0.2414	0.562 (0.357-.0.65)	0.0024
MCV (fL) Median (range)	108.0 88.7-122.1)	108.4 (94.4-122.2)	0.1715	108 (93.3-133)	0.3777
Platelets $x10^9/L$ Median (range)	166 (26-432)	127 (50-253)	0.1840	117 (36-1208)	0.5274
WBC $x10^9/L$ Median (range)	14.8 (4.7-44.2)	13.5 (5.5-29.1)	0.1074	25.8 (19.7-73.2)	<0.0001
Blasts (%) Median (range)	4 (0-15)	4.5 (1-10)	0.9842	35 15-77)	<0.0001
Neutrophils $x10^9/L$ Median (range)	10.58 (1.5-38.1)	8.5 (2.1-23.3)	0.0427	13.80 (7.00-31.00)	0.0285
Monocytes $x10^9/L$ Median (range)	1.19 (0.38-6.0)	1.10 (0.23-2.28)	0.1650	1.41 (0.67-5.28)	0.1398
Basophils $x10^9/L$ Median (range)	0.2 (0-1.07)	0.19 (0-0.6)	0.6835	0.37 (0-1.27)	0.0197

[a] ML-DS was diagnosed in 3 neonates with TAM (DST5, age 2 months; DST10, age 22 months: DST12, age 4 months) and in one neonate with silent TAM (DS 108). This neonate had 5% blasts and mild thrombocytopenia at birth ($79 \times 10^9/L$) but had a normal CBC and smear at age 9 months. Isolated thrombocytopenia ($23 \times 10^9/L$) was noted at age 11 months shortly after a viral illness and was attributed to ITP (no blasts were seen on the blood smear). Thrombocytopenia persisted and by age 15 months occasional blasts were seen on the smear. Progressive pancytopenia lead to the diagnosis of ML-DS at age 18 months (BM blasts 35%; no additional cytogenetic abnormalities). All patients who developed ML-DS remain in complete clinical remission after treatment with modified AML chemotherapy.