Childhood Cancer 2016

children with cancer uk
fighting the UK's biggest child killer

Registered Charity No. 298405

Day three: Wednesday 7th September 2016

What do we know about influenceable causes of childhood cancer?

Chaired by Professor Denis Henshaw and Professor Tariq Enver

Submitted abstracts selected for poster/oral presentation
(* denotes abstracts selected for oral presentation)

Diet, transplacental carcinogenesis and risk to children: results from the NewGeneris project

Soterios Kyrtopoulos
Institute of Biology, Medicinal Chemistry and Biotechnology,
National Hellenic Research Foundation, Athens
skyrt@eie.gr

NewGeneris ("Newborns and Genotoxic Exposure Risks") is a European project which utilized biomarker methodology to study the relationship between exposure of pregnant mothers to dietary genotoxic and immunotoxic chemicals, on one hand, and health effects in children, on the other, with emphasis on cancer risks. The project made use of biological material, as well as exposure and other data, from 1,151 pregnant women and their newborns coming from mother-child cohorts based in 5 regions across Europe (Norway, Denmark, UK, Spain, and Greece). Basic questions addressed in the project related to the dietary habits of pregnant women and the associated exposure to a series of genotoxic and immunotoxic chemicals, the ability of these chemicals to cross the placenta and reach the fetus, and the induction in the fetus of potentially carcinogenic events. The main research tools employed for this purpose included the measurement of biomarkers of exposure (DNA and protein adducts) in maternal and cord blood, biomarkers of early biological effects (changes in gene expression) or cancer risk (micronuclei) and in vitro placental perfusion. The dietary chemicals on which the project focused included organochlorinated persistant organics (dioxins, polychlorinated biphenyls - PCBs), polycyclic aromatic hydrocarbons - PAHs, heterocyclic amines, acrylamide, nitrosamines, mycotoxins and fat oxidation-derived reactive aldehydes. Some of these chemicals are found in food as contaminants whereas others are mainly formed in situ during food processing/cooking. Analysis of food frequency questionnaires was conducted for the purpose of identifying the main sources of exposure to these dietary chemicals and their variation in the different regions involved in the study.

In vitro placental perfusion studies showed that, while the transfer rates of the different compounds varies, all chemicals tested readily cross the placental barrier, with N-nitrosodimethylamine passing through most rapidly, followed by acrylamide and PAHs such as benzo[a]pyrene, with PCB 52 and dioxin being the slowest. The extent to which these chemicals reach the fetus was evaluated by measuring DNA or protein adducts in leukocytes isolated from umbilical cord blood. For the evaluation of exposure to dioxins and PCBs, plasma levels of species which activate the Ah, estrogen or androgen receptors were measured using different versions of the CALUX assay. Significant levels of exposure biomarkers were found for all chemicals, along with a strong correlation with the corresponding biomarkers measured in maternal blood, i.e. with maternal dietary intake. Large inter-individual variations in biomarker levels were also observed.

The potential carcinogenic risks associated with fetal exposure to carcinogens were examined by measuring in cord blood lymphocytes the frequency of micronuclei, a validated biomarker predictive of cancer risk. Significant associations with the micronucleus frequency were found for malonedialdehyde deoxyguanosine (M1dG), a biomarker of exposure mainly to oxidative fat metabolites, and ER- and AR-CALUX, related to exposure to dioxins and PCBs. The relationship between these exposure biomarkers and micronucleus frequency was further explored by examining the correlation of each biomarker with the expression of a series of genes selected on the basis of the results of a separate genome-wide expression profiling study conducted on a fraction of the studied newborns (see below). It was found that, six out of seven genes whose expression was associated significantly with the frequency of micronuclei were also significantly associated with the levels of at least one of the above exposure biomarkers. The affected genes have roles in cell cycle regulation and apoptosis.

Global gene expression profiling was conducted on lymphocytes from cord blood of 84 neonates from the Norwegian cohort and the correlation of gene expression with selected exposure biomarkers (DR-, ER- and AR-CALUX, acrylamide–hemoglobin adducts) was evaluated. The most notable finding relates to the fact that, although exposure levels did not differ significantly between sexes, in gender-stratified analyses major differences in the numbers of correlations with gene expression were observed, with little overlap between correlating genes. Moreover, opposite correlations with relevant biomarkers of exposure were found for leukemia/lymphoma genes, an observation which may be relevant to the known male-specific predisposition to develop these cancers in childhood. For example, in boys dioxin exposure was associated with activation of the proinflammatory transcription regulator TNFα, and acrylamide exposure was associated with activation of the Wnt pathway which is associated with uncontrolled cell growth.

Finally, a genome-wide SNP analysis was conducted on cord blood DNA from 435 subjects to look for polymorphisms which interact with the association between the exposures of interest and micronucleus frequency. A strong signal was found in the folate hydrolase 1 gene in relation with the association of dioxin-related androgen exposure with micronuclei. A polymorphism in this gene may cause low serum levels of folate, thus leading to chromosome instability and micronucleus formation. Additional signals were found in the epoxyhydrolase 1/2 and cytochrome p450 2E1 genes which are involved in the bioactivation of carcinogens.

In conclusion, the results of the NewGeneris project support the idea that the exposure of pregnant women living in the European region to dietary carcinogens results in fetal exposures which give rise to measurable molecular and functional changes associated with increased risk of developing cancer during childhood and later life.

Acknowledgements: This research was funded by the European Union 6th Framework Programme (Priority 5: Food Quality and Safety; grant no. FOOD-66 CT-2005-016320)

Origins and clonal evolution of childhood leukaemia

Tariq Enver
UCL Cancer Institute, London
t.enver@ucl.ac.uk

Childhood acute lymphoblastic leukaemia is thought in many cases to initiate *in utero.* The factors that influence the formation of the initiating lesions are not well understood although the nature of these genetic aberrations is well documented and include in approximately a quarter of children the presence of the t12;21 chromosomal translocation that fuses the transcription factors Tel (ETV6) and Aml1 (RUNX1). The Tel-Aml1 fusion gene produces a pre-leukaemic clone but in and of itself is insufficient to produce frank leukaemic transformation. For this, additional mutations are required and it remains unclear what factors influence their acquisition. These additional mutations tend to arise in loci that are normally involved in cell fate control in the B lineage. Cell fate regulators act in a context dependent manner and thus understanding the nature of the target cell or cells in which the initiating and subsequent mutations arise is of importance. To gain insight with these issues we have been exploring the target genes of TEL-AML1 and associated second hits as well as developing new foetal specific models in which to examine the biological impact of TEL-AML1. In this regard we have identified candidate target cells for TEL-AML1 within the human foetal liver and modeled these using IPS cells *in vitro.* Leukaemic clones appear to evolve in a branching manner such that at presentation the marrow is replete with multiple variegated subtypes providing a diverse substrate for selection in response to therapy. Beyond genetic heterogeneity, leukaemic cells exhibit epigenetic heterogeneity in respect of their immunophenotypes and functional properties including cell cycle status and niche residence. To obtain a "real-time" longitudinal analysis of subclonal dynamics through treatment we have established an *in vivo* mouse model that allows a patient tumour to be independently exposed to treatment multiple times, enabling us to distinguish between deterministic and stochastic mechanisms of selection during therapy.

What do we know about risk factors for childhood cancer and what are the challenges in ongoing research?

Kurt Straif MD, MPH, PhD
Section Head, IARC Monographs, International Agency for Research on Cancer, Lyon, France
straifk@iarc.fr

By way of introduction the presentation will provide an overview on procedures and main results of the IARC Monographs programme in identifying carcinogenic hazards to humans, from nomination of agents for systematic review and evaluation by the Monographs programme, prioritisation of agents by an international Advisory Group, selection of topics and of experts, including Declaration of Conflicts of Interest and the defined roles of different types of participants to safeguard against bias from vested interests, pre-meeting preparations, evaluation process at the 8-days consensus meeting and post-meeting communication of results and finalisation of the Monographs. The IARC Handbooks of Cancer Prevention follow procedures similar to those of the Monographs and identify what works in cancer prevention. The second part will focus on agents with important evaluations on childhood cancer, such as ionizing and non-ionizing radiation, parental smoking, occupational exposures for the Monographs and avoidance of body fatness for the Handbooks. Given the relatively scarce conclusive classifications to date, the final part will explore particular challenges in the causal inference linking environmental exposures to childhood cancer, such as exposure assessment, latency, and relative rarity of childhood cancers. With these challenges in mind research options in cancer epidemiology (childhood cancer registries, analytic study designs), taking into account evidence on cancer in experimental animals and use of mechanistic data for prioritisation will be explored and supplemented with examples of recent studies.

Primary Prevention: Can we reduce exposure to risk factors associated with childhood leukemia and other cancers?

Mark D Miller MD, MPH

Co-Director, Western States Pediatric Environmental Health Specialty Unit at the University of California, San Francisco; Director, Community Outreach and Translation Core, Center for Integrative Research on Childhood Leukemia and the Environment; Director, Children's Environmental Health Program at the California EPA (comments do not represent the State of California)

ucsfpehsumiller@gmail.com

A growing body of literature has implicated the role of several environmental hazards in the etiology of childhood leukemia. Exposures to solvents, traffic, pesticides, and tobacco smoke have consistently demonstrated positive associations with the risk of developing childhood leukemia including in pooled and meta-analyses. Intake of vitamins and folate supplementation during the pre-conception period or early pregnancy has been demonstrated to have a protective effect. Though less well studied, evidence is beginning to accumulate about environmental exposures and risk for brain tumors in children. Many of these risk factors overlap with those associated with leukemia. Despite the strength of these findings, the dissemination of current research to clinicians has been limited. Currently there is an absence of programs directed at reducing exposures to these identified risk factors. The incidence of childhood leukemia is high and increasing in Hispanic children in California, demonstrating a particular vulnerability in certain populations not entirely explained by genetics. It would be prudent to establish programs to alter exposure to those factors with well-established associations with cancer risk rather than to suspend judgment until no uncertainty remains. I will discuss perspectives on what constitutes adequate evidence for action and benefits vs. risks of potential preventative actions.

The Oxford Survey of Childhood Cancers: Perspective and Prospects

John Bithell[1] and Gerald Draper
[1]University of Oxford
john.bithell@spc.ox.ac.uk

The Oxford Survey of Childhood Cancers was started in Oxford in the 1950s and collected data on nearly all British children dying of malignant disease under the age of fifteen between 1953 and 1981. This survey was the largest case-control study of childhood malignancy ever undertaken and it resulted in many papers on a variety of topics concerned with the aetiology of the diseases; the most important was probably the observed association between ante-natal irradiation and malignant disease in the child. The study material has been preserved and is currently being documented and edited, with a view to making it available for further analyses, principal among which is a re-analysis of the X-ray data in the light of new insights, including the significance of birth weight. Although the study suffers from a number of obvious weaknesses – including the possibility of differential recall between cases and controls – there is a wealth of information, some of which has not yet been explored. For example, there is information on congenital abnormalities for both the probands and their sibs. Cross-tabulating the frequencies against other variables provides the possibility of observing associations that are free from case-control bias. Thus for example, we find an excess of children diagnosed with Wilms' tumour who also have one of a number of developmental abnormalities of the genito-urinary system; the fact that there is no corresponding excess among their sibs has implications for the nature of the genetic mutation and whether some cases could be due to an early somatic mutation. This paper will give a brief sketch of the available material in the survey and its potential value for aetiological investigations.

NQO1, GSTM1 and GSTT1 Polymorphisms Are Associated With Childhood Acute Myeloid Leukemia Subtypes and Type I Mutations

Gisele Brisson[1], Francianne Andrade[1], Filipe Bueno[1], Ingrid Cezar[1], Bruno Lopes[1], Bruno Gonçalves[1], Eugênia Terra-Granado[1], Maria Pombo-de-Oliveira[1]

[1]Pediatric Hematology-Oncology Program, Instituto Nacional de Câncer (INCA), Rio de Janeiro, Brasil.
gbrisson@inca.gov.br

Introduction. Childhood Acute Myeloid Leukemia (cAML) is a rare disease and little is known about its etiology. In Brazil, median age-adjusted incidence rate is about 10.4 cases per million, with a higher frequency of acute promyelocytic leukemia (APL). To try to identify causative factors underlying those unique epidemiological features, and considering that genetic susceptibility can modulate the risk of DNA damage from exogenous exposures, a case-control study to investigate the *NQO1, GSTM1* and *GSTT1* genetic polymorphisms in cAML risk in a Brazilian population was conducted.

Methods. There were 1,232 samples tested, being 450 AML (<21 years old) and 782 controls (healthy samples). Genomic DNA was extracted from bone marrow aspirates, peripheral or cord blood samples using QIAamp DNA Blood Mini Kit (QIAGEN). Genotyping for *NQO1* 609C>T (rs1800566) was performed by real time PCR (C_2091255_30, Applied Biosystems); *GSTT1* and *GSTM1* homozygous deletion (null or non-null genotypes) by multiplex PCR. AML subtypes were characterized according to World Health Organization recommendations. Type I mutations – *KRAS/NRAS* (G12D, G13D), *FLT3* (D835 or ITD) and *KIT* (exons 8 and 17) were investigated by direct sequencing. Statistical analysis was performed to estimate odds ratio (OR), age-adjusted OR (adjOR), and 95% confidence interval (95% CI), with chi-square tests, considering P-value < 0.05 as statistically significant.

Results. Cases and controls were compared in relation to demographic variables and genotype frequencies. High frequency of AML with 2–10 years-old (40.9%), males (56.2%) and non-whites (54.2%) were found. There were no statistical differences of AML and controls in relation to gender (*P* = 0.597) or ethnicity (*P* = 0.758), but differed in age groups (P < 0.001). Associations tests were adjusted by age. Concerning AML as a whole, there were 28.7% of cases with somatic abnormalities (*CBFβ/MYH11, RUNX1/RUNX1T1,* or *KTM2A-MLL-r* fusions genes); APL was the most frequent subtype (20.7%). Somatic mutations in *FLT3* (24.2%), *NRAS* (12.5%), *KRAS* (5.8%) and *KIT* (7.3%) were observed. *NQO1* 609C>T genotype frequencies were in accordance with Hardy-Weinberg equilibrium (P = 0.999); *NQO1* T allele was 0.25 - 0.27 among control and cases; *GSTT1* and *GSTM1* null genotype frequencies were 23.8% and 39.8%; 23.0% and 39.4% among controls and cases, respectively. Under recessive model (CC+CT versus TT), *NQO1* 609C>T was associated with increased risk to APL (adjOR 2.76, 95% CI 1.12 - 6.78, P = 0.027) and myelomonocytic or monoblastic/monocytic leukemia (adjOR 2.35, 95% CI 1.17 - 4.72, P = 0.017). *GSTM1* null genotype was associated with increased risk to AML with PML/RARα (adjOR 3.06, 95% CI 1.17 - 7.98, P = 0.022) and AML with $FLT3^{mut}$ (adjOR 2.42, 95% CI 1.17 - 5.00, P = 0.017). *GSTM1* or *GSTT1* null conferred increased risks to AML with $FLT3^{mut}$ (adjOR 2.66, 95%CI 1.05 - 6.76, P = 0.040) or any type I mutation (adjOR 2.14, 95%CI 1.02 - 4.52, P = 0.045). Association between *GSTM1* or *GSTT1* null genotypes and AML with $NRAS^{mut}$ (adjOR 2.76, 95% CI 0.99 - 7.75, P = 0.053) were observed, and it was statistically significant on crude analysis (OR 2.31, 95%CI 1.06 - 5.04, *P* = 0.042). Results show that *NQO1* and *GSTs* variants are associated with specific AML subtypes [APL and $FLT3^{mut}$].

Conclusions. Since *NQO1* and *GSTs* encode for detoxifying enzymes that are essential for protecting bone marrow from reactive metabolites, like benzene derivatives, genetic variants can modulate the risk of DNA damage, mainly in genes involved in signaling pathways, which are disturbed in leukemia cells.

Making Sense of Genetic Associations with Childhood Leukaemia Risk

Sandeep K Singh[1], Philip J Lupo[2], Michael E Scheurer[2], Anshul Saxena[1], Amy E Kennedy[3], Ertan Kanbur[4], Mehmet Fatih Okcu[2], **Mehmet Tevfik Dorak[4]**

[1]Florida International University, Miami, FL, USA; [2]Baylor College of Medicine, Houston, TX, USA; [3]National Cancer Institute, Bethesda, MD, USA; [4]Liverpool Hope University, Liverpool, U.K.

dorakm@hope.ac.uk

Besides environmental risk factors, genetic risk markers can also be used to gain insight into aetiology and disease biology. Childhood acute lymphoblastic leukaemia (ALL) has been repeatedly examined for genetic risk markers and very consistent results have been reported. We also did our own genome-wide association study (GWAS) designed to unravel sex-specific risk markers with novel results. Here, we report our functional assessments of the most established childhood ALL genetic risk markers (*ARID5B* rs4245595, rs10821936, rs7089424, rs10821936, and *IKZF1* rs1110701 & rs6964969, rs4132601, rs11978267) as well as the statistically most significant result (*RASSF2* rs4813720) from our GWAS. The *ARID5B* and *IKZF1* SNPs were all in non-coding regions, and associations with different SNPs within the same gene were not independent (they were all correlated with one another with r^2 >0.80). *ARID5B* and *IKZF1* SNPs all showed correlations with the expression levels of ARID5B and IKZF1, respectively (i.e., they were expression quantitative trait loci or eQTLs). They were not trans-eQTLs for any other gene. *IKZF1* and *ARID5B* are both upregulated in ALL, and are both targets of MIR218. Extensive functional annotations of these risk markers and their statistically similar SNP sets (or proxies) did not identify any target gene other than *ARID5B* and *IKZF1*. In our GWAS, *RASSF2* non-coding region SNP rs4813720 showed the strongest result for a genetic association with childhood ALL risk in interaction with sex (protection for males, risk for females). *RASSF2* is a tumour suppressor gene and inactivates the oncogene *KRAS* by physically binding to it. rs4813720 is known to be an eQTL for *RASSF2* by virtue of location in a transcriptionally very active part of the gene. Screening of mQTLdb for methylation-QTL (mQTL) effects revealed that rs4813720 is an mQTL for the CpG island cg22485289 within *RASSF2* at birth, during childhood and adolescence as well as during pregnancy at the statistical threshold ($P < 10^{-14}$) set for that study. Most crucially, the CpG island cg22485289 is one of the methylation sites reported to be crucial for B-lymphocyte differentiation and found to be hypomethylated in ALL. Thus, the sex-specific association with childhood ALL appears to be due to *RASSF2* methylation differences modulated by rs4813720. The overall data suggest that males possessing the minor allele of rs4813720 are protected by increased methylation of this site. We also identified *RASSF2* as a target of MIR17HG-derived microRNAs (MIR17/MIR19) which are produced in response to oestrogen-induced *MYC* activation. This would suggest that females are at high risk for MYC-induced transformation and promotion events, especially if they are *KRAS*-mutation-positive (a common occurrence in childhood ALL). We also obtained lists of co-expressed genes (Spearman's rho >0.80) for *ARID5B, IKZF1* and *RASSF2* and subjected the gene sets to enrichment analyses. Interesting findings with FDR $<5x10^{-8}$ included enrichment for SP1 binding sites (*ARID5B, IKZF1* and *RASSF2*), targets for MIR17 (*RASSF2*) and MIR19 (*IKZF1*), MLLT7 and LEF1 binding sites (*ARID5B* and *IKZF1*), and ELK1 binding sites (*IKZF1*). We will continue with *in silico* analyses of genetic association results. Our future work will include replication of genetic association results in a larger cohort with known KRAS mutation status and experimental verification of these results in childhood ALL cells to produce results with clinical utility.

Progress in Interpretation of Childhood ALL Risk Associations Identified in Candidate Gene Studies

Amy E Kennedy[1], Philip J Lupo[2], Michael E Scheurer[2], Ertan Kanbur[3], Mehmet Fatih Okcu[2], **Mehmet Tevfik Dorak[3]**

[1]National Cancer Institute, Bethesda, USA; [2]Baylor College of Medicine, Houston, USA;
[3]Liverpool Hope University, Liverpool, UK
dorakm@hope.ac.uk

Having recently reported replications of our previously reported associations of the extended HLA region polymorphisms (*HLA*-DRB9 rs2395185, *HFE* rs1800562, rs9366637 and rs807212), and the transferrin receptor gene (*TFRC*) polymorphism (rs3817672) with birth weight and childhood acute lymphoblastic leukaemia (ALL) risk, we have now revisited these statistical associations to explore their potential mechanisms. We performed an extensive survey of available results from unbiased omics studies. We used PhenoScanner to check cumulative genome-wide association study (GWAS) results and expression/methylation quantitative trait locus (eQTL/meQTL) status of SNPs and their proxies (r^2 >0.6); and Blood eQTL Browser, BIOS QTL Browser and mQTLdb for additional eQTL/meQTL results. We examined effects of environmental chemicals on relevant gene expression levels in Comparative Toxicogenomics Database (CTDbase).

rs2395185: Since we first reported the *HLA-DRB4* lineage, which is now represented by rs2395185, as a male-specific risk marker, this result has been observed in several other studies. It is now also associated with Hodgkin lymphoma (protection), lung cancer in never-smoked women (risk), and height. rs2395185 and its statistically similar SNP (ssSNP) set cumulatively are among the strongest trans-eQTLs in the genome, that is, they correlate with expression levels of genes on different chromosomes. rs2395185 and three of its ssSNPs are also the strongest meQTLs in the human genome ($P = 3x10^{-310}$; false discovery rate (FDR) = 0).

rs1800562: As a missense mutation (C282Y) responsible for iron overload causing hereditary haemochromatosis (HH), rs1800562 is associated with male-specific childhood ALL risk and birth weight in our studies. rs1800562 is also a trans-eQTL for the X-linked gene iron-related gene *ALAS2*, a cis-eQTL for HIST1H4C and cis-meQTL for four CpG sites. rs1800562 shows positive correlation with height, and patients with HH are reported to be taller than non-HH controls.

rs9366637: This SNP was identified in our original work funded by *Children with Leukaemia* as a haplotype tagging *HFE* SNP, and showed an association with birth weight as well as childhood ALL risk (also in a Korean study). This SNP is associated with increased iron absorption via decreased *HFE* expression, and correlates with some iron indices. We have recently reported its association with preterm birth as well. We have now identified its meQTL effect on cg18576210.

rs807212: This SNP is the one reported by us that tags a haplotype that is devoid of all pathogenic *HFE* mutations that cause iron overload. Its protective association is stronger than the original *HFE* rs1800562 risk association, and in a multivariable model, accounts for it. rs807212 is one of the strongest meQTLs in the genome for multiple CpG sites, including cg18576210, in all five stages examined (pregnancy, birth, childhood, adolescence and middle-age). It is a very strong eQTL for *HIST1H4C*, one of the histone genes in the vicinity of *HFE*, which is also an eQTL target for other HFE region SNPs that are childhood ALL risk markers. CTDbase showed that *HIST1H4C* expression is modified by benzo[a]pyrene and aflatoxin B1.

rs3817672: This *TFRC* SNP only showed associations in interaction with *HFE* SNPs in our studies. It also is an eQTL for *TFRC* and strong meQTL for multiple CpG sites. One of those (cg23245007) is reported to be a differentially methylated site in both ALL and CLL.

Conclusion: One common feature of these risk markers was that they are all strong trans-eQTLs and -meQTLs with some of their targets on other chromosomes. In addition to the obvious effects on iron homeostasis, the nearby histone gene cluster is implicated as the common eQTL target for all *HFE* region risk markers. *HIST1H4C*, which is cancer-related, is a common eQTL target for *HFE* region risk markers. All *HFE* region SNPs examined are also strongly associated with height like *HLA-DRB9* rs2395185. Adult height has recently been identified as a major contributor to the gender effect in cancer susceptibility. Our results draw attention to the involvement of large networks of trans-eQTL and -meQTL targets as well as the pathways regulating height for their involvement in childhood ALL susceptibility.

Developing a model to study co-operating oncogenic mutations in CEBPA-mutated AML

Catherine Hockings[1], Jasmine Rowell[1], Greg Contento[1], Oscar Pena[1], Elspeth Payne[1]
[1]UCL Cancer Institute, 72 Huntley Street, London WC1E 6BT
Catherine.hockings@ucl.ac.uk

Acute myeloid leukemia (AML) accounts for 20% of leukaemias seen in children. In contrast to acute lymphoblastic leukaemia (ALL), the prognosis of paediatric AML remains poor, with an overall survival at 5 years of just over 60%. Furthermore toxicity from both conventional chemotherapy and stem cell transplant for paediatric AML can be life-long.

The current model for the pathogenesis of AML involves a normal haematopoietic stem or progenitor cell (HSPC) acquiring successive genetic abnormalities that lead to clonal expansion and malignant transformation. Next generation sequencing (NGS) has facilitated the identification of more than 250 recurrent genetic changes in AML, with each child possessing between 5 - 20. Only a fraction of these have been validated as driver events, and the epistatic relationships and cellular consequences of combined mutational genotypes remain largely unexplored. A unique insight into this is afforded by a subset of children who have inherited a mutation that predisposes to the development of AML, so-called familial leukaemia with predisposition to AML (FPD-AML) since these mutations are unequivocally driver mutations. One such mutation occurs in the N-terminal domain of the transcription factor CEBPA. These mutations also occur in 10 -15% of sporadic paediatric AML cases. To further understand the role of N-terminal CEBPA mutations and co-operating secondary driver mutations we have generated a zebrafish model carrying an analagous N-terminal frame-shift mutation in Cebpa ($Cebpa^{Nterm}$).

Our initial analysis shows that $Cebpa^{Nterm/Nterm}$ mutants have severe defects in developmental myelopoeisis. By 5 days post fertilisation (dpf), homozygous $Cebpa^{Nterm/Nterm}$ mutants show a complete absence of Sudan Black (SB) staining myeloid cells. $Cebpa^{Nterm/+}$ siblings display a minor delay in myeloid development but no significant difference in numbers of SB staining cells at 5 dpf.

$Cebpa^{Nterm/Nterm}$ mutants are viable during the larval and juvenile period but show reduced growth and survival compared to siblings. At 4 weeks $Cebpa^{Nterm/Nterm}$ mutants show markedly reduced length and weight compared to age matched controls ($Cebpa^{+/+}$ and $Cebpa^{Nterm/+}$). All $Cebpa^{Nterm/Nterm}$ mutants die by 6 weeks of age. Assessment of myeloid cells at 5 weeks of age demonstrated a continued absence of myeloid cells by flow cytometric analysis of blood and whole kidney marrow.

The occurrence of biallelic N-terminal CEBPA-mutations in human AML is however extremely rare. By contrast the most frequent “second hit” mutation occur in the C-terminal DNA binding domain of the other allele of CEBPA. These mutations are universally in frame deletions or insertions, suggesting retained dimerization capability but abrogation of DNA binding. In addition to CEBPA C-terminal mutations, a number of other recurrently mutated genes have been found in CEBPA mutated AML. Our ongoing work is generating a knock-in C-term Cebpa model along with Crispr-targeting of additional hits to define their role in leukaemogenesis in CEBPA- mutated AML.

MNX1 has Role in the Development of Childhood Acute Leukemia

Fang Ju[1], Stefano Tonin, Olivier Binda, Christine J Harrison, Gordon Strathdee
[1]Northern Institute for Cancer Research, Newcastle University
f.ju2@newcastle.ac.uk

Several studies suggest that *MNX1 (HLXB9)* plays an important role in leukaemogenesis. MNX1 is a transcription factor expressed in haematpoietic progenitor cells. It is fused with *ETV6* as a part of a recurrent t(7;12))(q36;p13) in childhood AML. This translocation has been detected in about 25% of infant AML patients and is associated with a very poor outcome. The WT copy of *MNX1* is also over-expressed in many patients with this translocation. This led to interest in *MNX1* as on oncogene, either as part of a fusion gene or independently when over-expressed. However, only a fraction of t(7;12) patients appear to express a detectable fusion gene and some cases with the translocation exhibit deletion of the whole of the *MNX1* locus on the translocated chromosome 7. In addition, subsequent studies have shown that *MNX1* is frequently targeted by hypermethylation and is epigenetically inactivated in about a third of childhood ALL cases, suggesting that *MNX1* could actually play a suppressive role in leukaemia. Thus the role of *MNX1* in the causation of leukaemia is still unclear.

To attempt to address this question we have performed functional studies of MNX1 in human leukaemia cell lines. Using a lentiviral based transduction system we have re-expressed or over-expressed MNX1 in a variety of cell lines with both loss of MNX1 expression (due to promoter hypermethylation) or maintenance of MNX1 expression. In lymphoid leukaemia cell lines re-expression of MNX1 into deficient cell lines (NALM6 and Reh) resulted in induction of apoptosis and cell cycle arrest. In contrast, over-expression in two cell lines which maintain endogenous MNX1 expression (MOLT4 and Raji) had little or no impact on cell growth. Similar results were obtained in AML cell lines, with rapid induction of apoptosis in three cell lines with hypermethylated MNX1 (HL60, Kasumi-1 and MV4;11) but only limited impact of over-expression in the MNX1 positive GDM-1 cell line (although MNX1 positive THP1 cell lines were also highly sensitive to additional MNX1 expression). The results are consistent with MNX1 playing a predominantly negative role in the development of leukaemia, but also suggest its specific impact may be dependent on genetic background.

Studies in pancreatic cells have suggested that induction of apoptosis by MNX1 may be dependent on its phosphorylation status. To determine if the differential sensitivity to MNX1 expression could be related to its phosphorylation status we created two mutant clones with mutations of known MNX1 phosphorylation sites; mutated to alanine in one (non-phosphorylatable mutant) and mutated to aspartate in the other (phospho-mimetic mutant). Both mutants were essentially equally effective at inducing anti-proliferative effects in sensitive cell lines such as NALM6 and THP1. However in the resistant MOLT4 cell line, the phosphomimetic mutant had no impact on cell growth, exactly as seen for WT MNX1, but the non-phosphorylatable version negatively impacted cell growth and survival. To confirm that this effect was dependent on phosphorylation of MNX1, MOLT4 cells over-expressing MNX1 were treated with an inhibitor of GSK3beta (known to phosphorylate MNX1 on ser77 and ser79). While this inhibitor had no impact in cells expressing either of the two mutant versions, in cells expressing WT MNX1, treatment with the inhibitor reproduced the anti-proliferative effects seen with the non-phosphorylatable mutant, suggesting that tolerance for MNX1 expression may be dependent on phosphorylation or ser77 and/or ser79 by GSK3B. MNX1 has also been reported to interact with Menin, a key co-factor for MLL and current studies aim to determine if this interaction is also dependent on MNX1 phosphorylation status.

Childhood Leukaemia and Natural Background Radiation

Gerry M Kendall[1], Mark P Little[2], Richard Wakeford[3], Kathryn J Bunch[1], Jon CH Miles[4], Tim J Vincent[1], Jill R Meara[5] and Michael FG Murphy[1]

[1]Childhood Cancer Research Group, University of Oxford; [2]Radiation Epidemiology Branch, NCI, DHHS, NIH, DCEG, Bethesda, USA; [3]Institute of Population Health, University of Manchester; [4]Formerly Public Health England, CRCE, Didcot; [5]Public Health England, CRCEH, Didcot

Gerald.Kendall@ceu.ox.ac.uk

Introduction

Epidemiological studies of natural background ionising radiation and childhood cancer must be huge to have sufficient power to detect the very small radiation effects predicted by standard models. Interview-based case-control studies with sufficient study subjects are impracticable and would also be liable to various important biases. Record-based studies, with cases and controls drawn from existing registers, can be potentially large enough and can avoid obvious biases. However, they generally lack direct measurements of radiation levels in subjects' homes, and other relevant variables.

Background (Phase I Study)

A published record-based case-control study from Great Britain involving 27 447 cases and 36 793 controls investigated links between childhood cancer risk and two components of naturally occurring radiation: radon and gamma rays (including the directly ionising component of cosmic rays). It detected a statistically significant association between indoor gamma ray exposures and the risk of childhood leukaemia; the Excess Relative Risk (ERR) per unit dose was compatible with that extrapolated from higher dose studies, in particular those of the Japanese Atomic Bomb Survivors, and also with a UK study of cancer in children and young adults following diagnostic computerised tomographic imaging. For gamma rays, relative risks for other childhood cancers, and for radon, relative risks for both leukaemia and other cancers, were above one but not close to statistical significance.

Phase II Study

A criticism of Phase I was that it used small area averages as measures of gamma ray doses, based on a set of 2283 measurements. The study in preparation will use estimates of individual doses rather than small area averages, based on model extrapolations from a larger measurement set (n = 10 199). It will also be substantially larger including provisionally 54 462 cases and 69 992 controls. This study should allow radiation-related risks from Phase I to be refined.

Cell-specific genome-wide dysregulation of gene expression in fetal haematopoietic stem and progenitor cells precedes the development of leukaemia in Down syndrome

Binbin Liu[1], Anindita Roy[1], Katerina Goudevenou[2], Georg Bohn[2], Sarah Filippi[3], Natalina Elliott[1], Gillian Cowan[2], David O'Connor[2], Neha Bhatnagar[2], Anastasios Karadimitris[2], Irene Roberts[1, 4]

[1]Department of Paediatrics, University of Oxford, Oxford, [2]Centre for Haematology, Imperial College London, London, [3]Department of Statistics, [4]Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, United Kingdom.
Binbin.liu@paediatrics.ox.ac.uk

Background: Trisomy 21 (T21) is a major risk factor for childhood leukaemia, both in children with Down syndrome (DS) and, as an acquired abnormality, in leukaemias in children without DS. We have previously shown that T21 causes multilineage perturbation of fetal and neonatal haematopoiesis. In particular, T21 fetal liver samples have increased numbers of immunophenotypic haematopoietic stem cells (HSC) and megakaryocyte-erythroid progenitors (MEP) together with a severe reduction in committed B progenitors (CBP). Functional studies and Fluidigm gene expression assays support lineage bias favouring megakaryocyte-erythroid over B lymphoid differentiation and proliferation but did not identify causative genes on chromosome 21 (Hsa21) reflecting the limited number of Hsa21 genes which can be included on the panel1. Determining the mechanisms underlying these defects is likely to be important for understanding why young children with DS are more prone to myeloid and lymphoid leukaemias. **Aims:** To investigate the impact of the additional copy of chromosome 21 (Hsa21) on the transcriptome of primary fetal liver HSC and progenitor cells (HSPC). Specifically, we developed a bottom-up approach to determine (a) whether T21 causes consistent patterns of genome-wide dysregulation of gene expression in HSPC and (b) whether there is a T21 'signature' perturbed on the localised chromosomal level and whether this is consistent across the lineage. Methods: Snapshots of HSPC transcriptome were captured by RNA-seq of 8 flow-sorted populations, HSC, MPP, LMPP, CMP, MEP, GMP, ELP and CBP, from 7 gestation-matched 2nd trimester FL samples: DS (n = 4) and normal (n = 3). Indexed cDNA libraries were multiplexed and sequenced using Illumina HiSeq2500. Raw reads generated were subjected to an in-house developed analysis pipeline including adaptor trimming, QA, filtering and alignment of genome and transcriptome. Batch effect and other unwanted experimental variation was modelled and eliminated prior to differential expression analysis in each HSPC population between the DS and normal FL samples. Local polynomial regression fitting (LOESS) was used to define the boundaries of up/down regulated chromosomal domains. Data analysis and visualisation was performed with R and Bioconductor packages.

Results: Although lineage-specific gene expression patterns were largely maintained in DS HSPC, differential expression analysis identified around 100 to 800 significant genes (FDR<0.1) in T21 HSPC. Perturbation of lineage-associated gene expression was particularly prominent in HSC, MPP and CBP and the impact of trisomy 21 on haematopoiesis was specific to each lineage. More than half of protein coding Hsa21 genes was expressed in both DS and normal FL HSPC. Within each HSPC population, a consistent Hsa21 gene expression profile was seen in all samples and most of the differences in Hsa21 gene expression between T21 and normal HSPC were an exaggeration of the 'normal' HSPC profile rather than an aberrant expression of a different set of Hsa21 genes or a single Hsa21 gene. Identified dysregulation domains in T21 HSPC shows Hsa21 gene expression was consistently higher in T21 HSPC compared to normal HSPC along the entire length of Hsa21. In contrast, up and down-regulated local domains were identified on other autosomes in T21.

Summary/Conclusion: These data show that global perturbation of fetal haematopoiesis by T21 is matched by genome-wide dysregulation of gene expression affecting most chromosomes in all HSPC. Although DS FL HSPC populations all showed a 'T21 gene expression signature', this reflected more pronounced expression of the same Hsa21 genes as in normal HSPC rather than being driven by aberrant expression of a small subset, or single, Hsa21 gene(s). Dysregulation domains have been constructed based on the transcriptome for study of trisomy 21 impact to fetal haematopoiesis. We hope this model will provide a platform for understanding the role of somatic GATA1 and cohesin/CTCF mutations in DS leukaemias.

Anthropometric measurements at birth and risk of primary central nervous system tumors: a systematic review and meta-analysis

Marios K Georgakis[1], Eleni I Kalogirou[1], Athanassios Liaskas[1], Maria A Karalexi[1], Paraskevi Papathoma[1], Kyriaki Ladopoulou[1], Maria Kantzanou[1], Georgios Tsivgoulis[2], **Eleni Th Petridou[1]**
[1]Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, National and Kapodistrian University of Athens, Athens, Greece; [2]Second Department of Neurology, Attikon University General Hospital, School of Medicine, National and Kapodistrian University of Athens, Athens, Greece
epetrid@med.uoa.gr

Background

The aetiology of primary central nervous system (CNS) tumors remains largely unknown, albeit their peak in childhood points to perinatal parameters as tentative risk factors. In this meta-analysis we opted to quantitatively synthesize published evidence on the association between birth anthropometrics and risk of primary CNS tumors.

Methods

Pairs of independent reviewers searched Medline through May 18th, 2016, performed data abstraction and evaluated study quality using Newcastle-Ottawa scale. Random-effects meta-analyses were conducted to explore the association of birth weight and size-for-gestational-age separately for primary CNS tumors diagnosed in childhood and those in adulthood. Subgroup, sensitivity, meta-regression and dose-response analyses were also performed.

Results

A total of 41 articles were included. High birth weight (>4,000 g) was associated with increased risk of childhood CNS tumors (OR: 1.14, [1.09 - 1.20]; 22,330 cases). The risk was significantly higher for astrocytomas (OR: 1.22, [1.13 - 1.31]; 7,456 cases) and variable for the main embryonal CNS tumors subtypes (OR: 1.15, [0.99 - 1.33]; 2,352 cases); notably sizeable for primitive neuroectodermal tumors (PNETs; OR: 1.32, [1.05 - 1.64]; 890 cases) and practically null for medulloblastomas (OR: 1.03, [0.83 - 1.67]; 1,037 cases). All of these associations were non-linear with an attenuation of risk for birth weights <2,500 g. No effect was noted for ependymomas. Infants, large-for-gestational-age, were also at increased risk for childhood CNS tumors (OR: 1.12, [1.03 - 1.22]); the paucity of evidence did not allow in depth exploration of histological subtypes though. No heterogeneity or publication bias were evident, whereas the effects were robust across high quality and registry-based studies, studies adjusting for gestational age and, as expected, and those using alternative sources for birth weight assessment. No meta-analysis was possible for other anthropometric factors, whereas the limited available evidence on adult CNS tumors did not reveal significant associations.

Conclusions

The largest ever meta-analysis shows a robust association of the easy-to-assess high birth weight, independently of gestational age, with risk of childhood CNS tumors, especially astrocytomas and possibly with PNETs. Further research is needed to explore the role of specific mechanisms underlying this association.

What the differential risk for acute leukemia subtypes in subsequent offspring by history of maternal miscarriage of stillbirth may imply? A systematic review and meta-analysis

Maria A Karalexi[1], Marios K Georgakis[1], Alkistis Skalkidou[2], Eleni I Kalogirou[1], Stylianos-Iasonas Biniaris-Georgalis[1], Eric Herlenius[3], Dimitrios Loutradis[4], Georgios P Chrousos[5], **Eleni Th Petridou[1]**

[1]Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, National and Kapodistrian University of Athens, Athens, Greece; [2]Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden; [3]Neonatal Research Unit, Department of Women's and Children's Health, Astrid Lindgren Children's Hospital both at Karolinska Institutet, S-171 76, Stockholm, Sweden; [4]1st Department of Obstetrics and Gynecology of the University of Athens, Alexandra Hospital, Athens, Greece; [5]First Department of Pediatrics, Athens University Medical School, Aghia Sophia Children's Hospital, Athens, Greece

epetrid@med.uoa.gr

Background

Maternal history of fetal loss, notably miscarriage and stillbirth, has been linked to adverse outcomes in subsequent offspring, including childhood (0 - 14 years) leukemogenesis. We, therefore, aimed to systematically assess and quantify the potential impact of maternal fetal loss history on leukemogenesis risk in subsequent offspring by disease subtype.

Methods

PubMed (end of search: March 2016) and the reference lists of the relevant studies were searched following PRISMA guidelines and included studies were evaluated using Newcastle-Ottawa scale. Random-effects meta-analyses were conducted on the association of history of miscarriage and stillbirth with acute leukemia (AL) and its major subtypes, notably acute lymphoblastic (ALL) and myeloid (AML) leukemia. Sensitivity and subgroup analyses by age and ALL subtypes were thereafter undertaken.

Results

Thirty-four eligible studies were identified amounting over 5,149 AL cases and 10,582 controls. History of previous fetal loss (miscarriage or/and stillbirth) was associated with increased AL risk overall [Odds Ratio (OR): 1.09, 95% Confidence Intervals (CI): 1.01 - 1.18] and ALL, in particular (16,466 cases and 38,917 controls; OR: 1.11, 95% CI: 1.04 - 1.18). Marginally significant were the associations of previous stillbirth with ALL or previous miscarriage with both ALL and AML (2,182 cases and 27,022 controls). Noticeable was the sizable association of history of miscarriage with infant ALL (OR: 2.27, 95% CI: 1.18 - 4.39).

Conclusions

In the largest ever meta-analysis, the risk for infant ALL was more than double in case of miscarriage history and marginally increased for ALL in all age groups in case of fetal loss. Less robust are the findings for the less frequent AML. Future research on the potential interaction of genetic polymorphisms with environmental factors is needed, however, to unveil plausible underlying biological mechanisms and render these observations a useful tool in clinical practice.

The Incidence Rate and Risk Factors Associated With Early Age Acute Leukaemia

Rejane de Souza Reis[1], Marceli de Oliveira Santos[1], Beatriz de Camargo[2], Julio Fernando Pinto Oliveira[1], Neimar de Paula Silva, Luiz Claudio Santos Thuler[3], **Maria S Pombo-de-Oliveira[2]**

[1]Surveillance and Cancer Prevention Coordenation, Instituto Nacional do Câncer, Rio de Janeiro, Brazil; [2]Pediatric Hematology-Oncology Program, Research Center, Instituto Nacional de Câncer, Rio de Janeiro, Brazil; [3]Clinical Research program at Reseach Center, Instituto Nacional do Cancer, Rio de Janeiro, Brazil

mpombo@inca.gov.br

Introduction

Childhood leukemias (CL) are originated during foetal life, arising from somatic clonal cells in which the oncogenic potential changes might be associated with environmental risk factors. The population-based registries (PBCR and Information System on Live Births-SINASC in Brazil) have information that enables us to test risk factors associated with leukemia in early age. The aim of this study was to identify birth characteristics and maternal variables that would be associated with acute leukemia in early age.

Methods

The study design explored secondary dataset information on population bases: 1) the first analysis concern the incidence rate (IR) and trends of 18 PBCR in Brazil, which gather CL subtypes (children ≤ 5 years at diagnosis); 2) a case-cohort study using data from the same PBCR cities and the SINASC of the same cities of PBCR. The risk associations variables to test were grouped into (i) characteristics of maternal exposure during pregnancy of the child (age, occupation, education) and (ii) characteristics of the child at birth (birth order, type of delivery, weight, APGAR notes, syndrome phenotype and/or congenital malformation). The case-control ratio was 1:4. Standardize, database structuring and statistical analyzes were performed throughout the Excel, SPSS, StatCalc-Epinfo and R-Studio statistical package. The magnitude of odds ratio (OR), 95% confidence interval (CI) were performed through logistic regression models (univariate) and multivariate analysis for whose raw OR suggest risks.

Results

The overall IR of all leukemia were 80.64/million for children at 2 years of age, and varied according to Brazilian regions; the lowest IR in the Northeast region (37.00/million). In the analysis of trends, increasing IR were observed in eight localities, mainly in the Northeast region were the João Pessoa city showed a significant increase (AAPC = 20.11%; CI 95%: 3.5 - 39.4) per annum for ten years period. In the opposite direction, in São Paulo there was a decline in acute leukemia IR (AAPC = -4.02%; CI 95%: -6.1 to -1.9). The median incidence rate is increasing and directly proportional to age in childhood AL. For the second analysis, 1,215 children being 272 cases and 1,088 controls were randomly selected. There were null association with socioeconomic indicators in the whole PBCR settings. Maternal age ≥ 30 years, delivery as cesarean section (AdjOR, 1.33 CI, 0.98 - 1.79) and the high birth weight ≥3,200 g were associated with increased risk of CL; 3% of acute leukemia had congenital anomaly. Regarding maternal exposures, occupation analysis it was found a positive association with agricultural work (OR, 2.45; 95% CI,1.33 - 4.51).

Conclusions

The sum up of this study with the identification of CL risk factors in population-based design, strengthens the exchange of knowledge and improvement of databases, and contributes to investigations on leukemia causation around the world.

DNA methylation as a mediator of in utero exposures on risk of childhood acute lymphoblastic leukaemia (ALL)

Jessica A Timms[1], Caroline L Relton[2], Gordon Strathdee[3], Judith Rankin[1], Jill A McKay[1]
[1]Institute of Health and Society, Newcastle University, UK, [2]MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, [3]Northern Institute for Cancer Research, Newcastle University, UK
j.a.timms@ncl.ac.uk

Childhood ALL arises from genetic abnormalities that can occur *in utero.* Reported frequencies of certain genetic abnormalities at birth are significantly higher than the number of associated ALL cases arising, suggesting secondary 'hits' are required for disease development. Evidence suggests several environmental exposures influence childhood ALL risk, and with peak incidence of diagnosis between 2 - 5 years of age, early life exposures are likely to be key. DNA methylation is modifiable by environment and altered in childhood ALL. Global DNA methylation alterations seen in ALL are mostly seen across all subtypes, and therefore appear to be early events in ALL development. Here we explore the potential that DNA methylation may be involved in the causal pathway towards disease by acting as a mediator between established *in utero* environmental factors and childhood ALL development.

Environmental factors associated with childhood ALL (i.e. smoking, maternal alcohol, radiation, folic acid, iron, caffeine, soft drinks, herbicides, pesticides, household chemicals and paints) were identified via literature searches using the NCBI and ScienceDirect databases (1987 - 2016). The Illumina Infinium HumanMethylation450K platform was used to analyse genome-wide DNA methylation at birth using cord blood from a subset of children (n = 836) from the Avon Longitudinal Study of Parents and Children cohort (ALSPAC). ALSPAC also hold environmental data on mothers throughout pregnancy. The effect of environmental exposures associated with childhood ALL risk on DNA methylation at individual CpG sites were analysed using linear regression. DNA methylation was modelled as a continuous variable, in a multivariate regression model accounting for potential confounders (sex, parity, gestation, and batch).

Using a $P \leq 0.9 \times 10^{-5}$ DNA methylation changes were found in association with smoking (n = 398 genes), alcohol (n = 219 genes), radiation exposure (n = 161 genes), household chemicals (n = 74 genes), herbicides (n = 56 genes), Iron (n = 44 genes), caffeine (n = 39 genes), household paints (n = 38 genes), folic acid (n = 35 genes), soft drinks (n = 31 genes), and pesticides (n = 26 genes). Hypergeometric tests were carried out for each exposure and all observations were found to be statistically significant (except household paint use). Although previous literature provides a higher weighting of evidence to support the associations with smoking, alcohol, radiation exposure and folic acid. Aberrant methylation has also been reported in the literature in childhood ALL for 259 of the risk factor-associated genes we discovered. As environmental exposures can potentially drive changes in methylation which are also associated with the disease state, this supports our hypothesis that DNA methylation may mediate the effect of environmental risks factors associated with disease development and therefore may be involved in the causal disease pathway. Further studies to understand the importance of these methylation changes in inducing or contributing to disease development will be important as such findings may provide predictive disease biomarkers and offer insights into how preventative strategies may be introduced.

Investigation of DNA methylation as a mechanism in the delayed infection (or hygiene) hypothesis for childhood acute lymphoblastic leukaemia (ALL) risk

Jessica A Timms[1], Caroline L Relton[2], Gordon Strathdee[3], Judith Rankin[1], Jill A McKay[1]
[1]Institute of Health and Society, Newcastle University, UK, [2]MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, [3]Northern Institute for Cancer Research, Newcastle University, UK
j.a.timms@ncl.ac.uk

Childhood ALL incidence has increased steadily in westernised populations and is associated with affluence. Infection was one of the first suggested risk factors for ALL, and has led to the development of the hygiene hypothesis, which suggests that lack of exposure to infections in early childhood suppresses normal immune system development leading to increased susceptibility of certain diseases, including childhood ALL. Epidemiological studies investigating social contact as a proxy for infection support this hypothesis, however the mechanisms involved are not understood. Theoretically day care attendance at a young age should mean that a child is confronted with common infections at an early age, allowing them to build a more sophisticated immune system, and a reduced chance of increased proliferation and risk of mutation in confronted with common infections at a later date. DNA methylation is modifiable by environment and altered in childhood ALL, and therefore may play a mediating role in this pathway to disease.

Using day care attendance and breast feeding as proxy measures for social contact and infection respectively, we investigated the influence of infection exposure on DNA methylation, to explore a potential role in the causal pathway to childhood ALL. The Illumina Infinium Human Methylation 450K platform was used to analyse genome-wide DNA methylation at age 7 using blood from a subset of children from the Avon Longitudinal Study of Parents and Children cohort (ALSPAC). The effects of breast feeding and day care attendance were analysed using linear regression. DNA methylation was modelled as a continuous variable, in a multivariate regression model accounting for potential confounders (sex, parity, gestation, and batch).

Using a $P \leq 0.9 \times 10^{-5}$ for significance, 72 and 7 genes had altered methylation in association with day care attendance and breast feeding respectively. No overlap was found between genes aberrantly methylated in association with day care attendance and breastfeeding. Hypergeometric tests were carried out, revealing that the genes found to have altered methylation due to breastfeeding status were statistically significant. Aberrant methylation has also been reported in the literature in childhood ALL for 18 of the risk factor-associated genes we discovered. This data suggests DNA methylation may play a mechanistic role in the causal pathway to childhood ALL development proposed by the hygiene hypothesis. Further studies are needed to better understand the role DNA methylation may have in mediating and possibly acting as a driving mechanism in the ALL disease pathway.

Germ line mutations in DNAJC21 disrupt ribosome biogenesis and predispose to acute myeloid leukaemia

Hemanth Tummala[1], Amanda Walne[1], Inderjeet Dokal[1] and **Tom Vulliamy[1]**

[1]Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London

t.vulliamy@qmul.ac.uk

Myelodysplastic syndrome (MDS) and acute myeloid leukaemia (AML) are related haematological malignancies. They usually occur sporadically but in some cases, there is a family history, indicating an inherited cause or predisposition. There is also a strong link between MDS/AML and the inherited bone marrow failure (BMF) syndromes and these cases are often associated with an earlier age of onset. In addition, we have seen a number of sporadic cases who present with BMF or MDS/AML together with one or more somatic abnormality, often at a young age. In order to identify a potential genetic basis for this disease, we have performed exome sequencing in a cohort of 28 such patients.

From this dataset, we have identified three cases, all of whom were children and one of whom had AML, who have homozygous likely pathogenic variants in *DNAJC21* (DnaJ homolog subfamily C member 21). Targeted sequencing of *DNAJC21* in a second cohort of 23 similar cases identified one further child with a homozygous nonsense variant. Three of the variants (p.Arg173*, p.Glu265* and c.983+1G>A) are predicted to cause loss of function. In the fourth, a missense variant (p.Pro32Ala) disrupts the highly conserved histidine-proline-aspartic acid motif that lies at the heart of a J domain, which defines a family of proteins. This allelic series is very unlikely to have occurred by chance and so we conclude that biallelic mutations in *DNAJC21* are disease causing.

To date, little is known about human DNAJC21, although its yeast orthologue (Jjj1) has been studied in some detail. We show that the human protein associates with ribosomal RNA (rRNA) and plays a highly conserved role in the maturation of the 60S ribosomal subunit. Patient cells, which showed a lack of DNAJC21 immunoreactivity, exhibited increased sensitivity to the transcriptional inhibitor actinomycin D and reduced levels of both the 45S precursor and the mature 28S rRNA. DNAJC21 deficiency also resulted in the cytoplasmic accumulation of a 60S export factor (PA2G4), while RNAi studies revealed that acute loss of DNAJC21 lead to S and G2/M phase arrest and cell death.

We therefore believe that biallelic germline mutations in *DNAJC21* define a novel cancer prone BMF syndrome of childhood caused by defective maturation of the ribosome 60S subunit.

Biopanning Immunoglobuin Repertoires To Identify Infections Which May Play A Role In The Aetiology Of Childhood Acute Lymphoblastic Leukaemia (ALL)

Thomas Walker[1], Xiaotong He[1], Lynne Hampson[1], Ian Hampson[1]
[1]University of Manchester Viral Oncology Labs, Institute of Cancer Sciences, Manchester
thomas.walker@manchester.ac.uk

Introduction: Epidemiological evidence suggests that the aetiology of childhood ALL may be associated with as yet undefined infection(s) although no such pathogens have yet been identified. This is not surprising since it is difficult to test for an asymptomatic pathogen without prior knowledge of its identity and some infections may cause cellular damage which persists and contributes to the development of malignant disease after the infection is cleared ("hit-and-run" hypothesis). If it were possible to compare antibody repertoires between ALL patients and controls, this could provide a solution to these problems as it would capture current and historical infections. This is very challenging as each individual generates >10 billion specific antibodies. However the ability to discriminate between antibodies against common infections and those uniquely found in children with ALL could be used to identify candidate pathogens and potential biomarkers associated with ALL.

Methods: Elution from Thiophilic Magnetic beads was used to isolate total immunoglobulins (Igs) from pooled sera. In order to compare antibody repertoires between separate pools of sera from disease and normal individuals, these were analysed by a unique in-house method based on a combination of competitive bio-panning with next generation sequencing (NGS). Novel computational tools were developed and used to analyse the NGS output to predict the identity and abundance of disease-specific infection related epitopes. Confirmation of these *in silico* predictions was then carried out by immunoprobing synthetic peptide microarrays. This was done with the original pooled Ig's used for the biopanning in addition to separate pools prepared from new cohorts of disease associated and normal sera.

Results: Pooled Ig's have been isolated from the sera of 100 children with ALL and 100 age-matched disease-free children all obtained within the Manchester area. These are currently being analysed with the competitive bio-panning procedure. Optimisation of this process has been successful whereby NGS produced a combined total of >17,000,000 reads of disease and normal epitope-encoding DNA's. These date have been analysed with a series of unique in-house algorithms in order to generate lists of expressed epitope peptides (>700,000), cross-compare both abundance and overlap between the two pools of sera and map them to the total known protein sequence repository. Outputs were transformed, tiled, and the data mined to identify past or current infectious agents that were unique, or more common, in one group than the other. Confirmation of these *in silico* predictions was then performed by designing custom peptide microarrays which were immuno-probed with the original Ig pools from disease and normal sera used for biopanning. Very extensive and unique computational methods were subsequently developed to deconvolute the microarray output, which verified the presence of Ig's against the predicted candidate infections. New Ig pools were prepared from different cohorts of disease and normal sera and these also used to immunoprobe the peptide microarrays. Doconvolution of the output confirmed the presence of Ig's against already identified candidate organisms in these additional disease associated sera.

Conclusions: A novel, high-resolution method of comparing antibody repertoires has been developed and is currently being applied to children with ALL. To date, this procedure has been tested and robustly detected the differential presence of organisms which have been previously implicated in the development of malignancies other than ALL. Future work will be aimed at validating these findings in individual sera from ALL patients which will then form the basis of studies aimed at understanding their mode-of-action. Clearly identification of an aetiological role for unknown infection(s) in ALL would be highly advantageous since such knowledge could underpin both prevention and treatment strategies.