Identification of genetic variants associated with skeletal muscle function deficit in childhood acute lymphoblastic leukemia survivors

Geneviève Nadeau1
Erika Ouimet-Grennan1
Michelle Aaron1
Simon Drouin1
Laurence Bertout2
Albert Shalmie2
Patrick Beaulieu2
Pascal St-Onge2
Louis-Nicolas Veilleux3
Frank Rauch3
Kateryna Petrykey1,2
Caroline Laverdière1,2,4
Daniel Sinnett1,2,4
Nathalie Alos1,2,5,*
Maja Krajinovic1,2,4,*

1Department of Medicine, University of Montreal, Montreal, QC, Canada; 2Sainte-Justine University Hospital Research Centre, Montreal, QC, Canada; 3Division of paediatrics, Montreal Shriners Hospital for Children, Montreal, QC, Canada; 4Division of Hemato-Oncology, Sainte-Justine University Hospital Centre, Montreal, QC, Canada; 5Division of Endocrinology, Sainte-Justine University Hospital Centre, Montreal, QC, Canada

*These authors contributed equally to this work

Background: Although 80% of childhood acute lymphoblastic leukemia (ALL) cases are cured with current treatment protocols, exposure to chemotherapeutics or radiation therapy during a vulnerable period of child development has been associated with a high frequency of late adverse effects (LAE). Previous observations suggest important skeletal muscle size, density and function deficits in ALL survivors.

Purpose: Given that only a fraction of all patients will suffer from this particular complication, we investigated whether it could be predicted by genetic markers.

Patients and methods: We analysed associations between skeletal muscle force (Fmax) and power (Pmax) and germline genetic variants from 1039 genes derived through whole-exome sequencing. Top-ranking association signals retained after correction for multiple testing were confirmed through genotyping, and further analysed through stratified analyses and multivariate models.

Results: Our results show that skeletal muscle function deficit is associated with two common single nucleotide polymorphisms (SNPs) (rs2001616DUOX2, P=0.0002 (Pmax)) and rs41270041ADAMTS4, P=0.02 (Fmax)) and two rare ones located in the ALOX15 gene (P≤0.001 (Pmax)). These associations were further modulated by sex, body mass index and risk groups, which reflected glucocorticoid dose and radiation therapy (P≤0.02).

Conclusion: Occurrence of muscle function deficit in childhood ALL is thus strongly modulated by variations in the DUOX2, ADAMTS4 and ALOX15 genes, which could lead to personalized prevention strategies in childhood ALL survivors.

Keywords: acute lymphoblastic leukemia, late adverse effects, skeletal muscle deficit, genetic association study, whole exome sequencing

Introduction

Childhood acute lymphoblastic leukemia (ALL) accounts for nearly 25% of all paediatric malignancies.1,2 Optimization of multi-agent risk-adapted protocols has improved the 5-year survival rate, reaching by 2017 an estimated 90% in children less than 14 years old.3 This therapeutic success introduces new challenges, as 60% of childhood ALL survivors now experience a broad spectrum of late adverse effects (LAE) due to aggressive chemo- and radiation-therapy administered during critical child development stages.3,4 Childhood ALL survivors are more likely to develop a metabolic syndrome, cardiotoxicity, neuropsychological deficit and musculoskeletal morbidity — a cluster of interconnected factors that globally contribute to severe impairment of their quality of life.6–11
Clinical evaluation of childhood ALL survivors has revealed skeletal muscle size, density and function deficits as compared with healthy controls. Evidence suggests acute oxidative stress induced alteration of functional muscle mitochondria, motor neurons and satellite cells could contribute to these functional deficits. Muscle morbidity could lead to long-term mobility and autonomy deficits, such as impaired walking and gait or a diminished capacity to carry out daily activities. As lipid oxidation and glucose consumption is primarily performed in skeletal muscle, it is likely that treatment-induced muscle function deficit ultimately contributes to obesity, and cardiovascular or metabolic syndrome-related morbidity in ALL survivors.

While molecular mechanisms involved in skeletal muscle atrophy remain unclear, sarcopenia as a response to glucocorticoid (GC) signalling, irrespective of cancer treatment, has been associated with various genetic markers and molecular pathways in the adult population. Evidence shows that GCs act as up-regulators of atrophy-related genes such as the Notch signalling pathway and constituents of the E3 ubiquitin ligase cascade. Likewise, mTOR (mechanistic target of rapamycin kinase) and AMPK (AMP-activated protein kinase) pathways are involved in GC-induced delays of muscle development in an animal model, while GCs increase adiposity in muscle tissue by affecting IL-4 (Interleukin 4) regulation of fibro/adipogenic progenitors.

Genetic susceptibility to long-term skeletal muscle morbidity in childhood ALL survivors has yet to be documented. To address this issue, we present correlations between relevant clinical outcomes and genotypic profiling obtained through whole exome sequencing (WES) of ALL survivors.

Identification of genetic markers with predictive potential would allow further risk stratification of patients and personalisation of LAE prevention programs.

**Methods**

**Study participants**

The study group is composed of 242 survivors from the PETALE cohort, diagnosed with childhood ALL since 1987 at the Sainte-Justine University Hospital Center (SJUHC) in Montreal. Eligibility criteria include an age at diagnosis of 19 years or less, a 2-year adherence to a Dana-Farber Cancer Institute (DFCI)-ALL treatment protocol, Caucasian descent, and a post-remission delay of at least 5 years. Patients who relapsed, received hematopoietic stem cell transplants, suffered from concomitant syndromes, congenital bone disease, sickle-cell anemia or Legg-Calvé-Perthes disease, or were given osteotoxic drugs outside of their ALL treatment were considered ineligible for the study. All patients fulfilling eligibility criteria and who consented to participate were included in this study. A written informed consent was obtained in accordance with the principles laid down by the Declaration of Helsinki. The SJUHC Institutional Review Board approved the research proposal.

**Muscle force and power phenotypic data**

Study participants underwent comprehensive clinical, biological and psychosocial investigations carried at the SJUHC and the Montreal Shriners Hospital for Children (MSHC), in the course of a day, between 2013-02-12 and 2016-05-19.

We used the Leonardo Mechanograph® Ground Reaction Force Plate (Novotec Medical GmbH, Pforzheim, Germany) as the measurement device for peak muscle force and power data acquisition. Data were analysed using the Leonardo Mechanography GRFP Research Edition® software, which measures body mass during quiet stance, and calculates the instantaneous vertical acceleration upon jumping, by recording the force as a function of time. The acceleration is time integrated to obtain the instantaneous vertical velocity. Peak power (Pmax) designates the highest value of force times velocity during the initial acceleration phase, while peak force (Fmax) refers to the maximum vertical force applied to the platform. Each participant performed a single two-legged jump and multiple one-legged hopping. Weight-adjusted Pmax and Fmax data were collected for each participant during a single two-leg jump and transformed into age- and gender-specific z-scores by comparison with a normal population. A z-score of −2 was used as a threshold to discriminate between affected and unaffected individuals.

**Sequencing and candidate gene selection**

A list of candidate genes was generated through a review of the literature. Based on the functional bone-muscle local unit theory, genes involved in both muscle and bone molecular pathways were targeted. Studied pathways include Wnt-β-catenin signalling, osteoclastogenesis, osteoblast and osteocyte maturation, Vitamin D signalling, and pathways involved in glucocorticoid and methotrexate (MTX) action. Additional candidate genes were selected via a Gene Ontology Enrichment Analysis and Visualisation Tool (Gorilla). A total of 1039 candidate genes were chosen.

DNA was extracted from whole blood obtained during phenotyping evaluation of PETALE participants. Whole
exomes were captured in solution with Agilent’s SureSelect Human All Exon 50Mb kits and sequenced on either Life Technologies SOLiD System 4.0 (mean coverage = 40X) or Illumina HiSeq 2500 platform (mean coverage = 113.1X) at SJUHC integrated clinical genomic centre in paediatrics. Reads were aligned to the hg19 reference genome using SOLiD LifeScope software for the SOLiD samples and BWA-MEM for the Illumina samples. PICARD was used to mark PCR duplicates and collect sequencing quality control metrics. Germline variant calling was performed using the Haplotype Caller and quality score recalibration was performed using Variant Recalibrator, both implemented in the Genome Analysis Tool Kit (GATK). Variants were selected based on the variant quality score (VQSR = PASS) and minimum depth of coverage (DP ≥ 10). WES data were available for 222 participants. Common and rare variants were defined respectively by their minor allele frequency (MAF) in the reference population as ≥5% and <5%, as based on 1000genome and ESP6500 (European American population) datasets. A functional annotation from ANNOVAR was used to apply filters to genotype data so that only variants with missense (at least potentially damaging, Polyphen2 scores ≥ 0.85 and Sift scores ≤ 0.1) or nonsense variations, or variations in splicing sites, were further considered.

Of those, common variants either exceeding a missing rate of 20%, not in the Hardy-Weinberg Equilibrium (HWE, p < 0.001), or having a pairwise linkage disequilibrium (LD, r² ≥ 0.8) were excluded, leaving 253 common SNPs located in 188 genes. The rare variants were analysed in an exploratory fashion using a collapsing approach whereby minor alleles of at least 2 potentially functional variants within one gene, each in HWE and with missing rate <20%, were analysed together, resulting in an analysis of 655 genes.

**Statistical analysis**

We used a whole genome data analysis toolset (PLINK) to quantify associations between common SNPs and phenotypic datasets applying an allelic ratio test for dichotomous variables and a linear regression for continuous phenotypes. Quantitative datasets were tested for normality through an assessment of the distribution of residuals. For rare variant associations, the SKAT-O test (Optimal Sequence Kernel Association Test) in R Studio software, which collapses variants within the same gene, was used. Resulting P-values were adjusted for multiple testing using a Benjamini-Hochberg correction method. A maximum false discovery rate (FDR) of 5% was chosen as a cut-off value for significance. Selective genotyping of top-ranking SNPs (FDR<5%) following WES data analysis was carried out for all participants on the Sequenom platform at the McGill University and Génome Québec Innovation Centre. DNA samples from all 242 participants were available for genotyping. Depending on the studied SNP, genotyping was inconclusive for 3–14 participants. Concordance percentages between sequencing and genotyping results for top-ranking signals were recorded. Resulting genotypes and phenotypic data were analysed by IBM SPSS Statistics for Mac Version 24.0 (IBM Corp. Armonk, NY) with the chi-square or the Fisher exact tests, using appropriate genetic models. Rare SNPs were analysed through a collapsing approach with iterative exclusion of a single SNP per iteration, allowing weighting the SNP contribution to association signals.

Logistic regression models (in IBM SPSS) were used to assess an influence of genotypes on studied outcomes in the presence of covariables that significantly correlated with studied phenotypes and included sex, risk status either high (HR) or standard (SR), and body mass index (BMI). BMI was included as a cofactor as adipose tissue accumulation in obese individuals is thought to interfere with muscle function through aberrant adipokines production. A participant BMI was classified as normal (BMI ≤ 24.9 kg/m²), overweight (25 kg/m² ≤ BMI < 30 kg/m²), or obese (BMI ≥ 30 kg/m²), in accordance with the adult classification scale.

Due to the DUOX2 gene role in thyroid hormone synthesis, the polymorphism rs2001616 located in DUOX2, for which significant associations with musculoskeletal deficit were found, was analysed for a correlation with TSH and T4 levels available for 235 and 231 PETALE participants, respectively.

To allow better characterisation of the impact of genotypes in relation to Fmax and Pmax, stratified analyses were also performed according to sex, risk group, and BMI at visit (obese vs non-obese).

**Results**

**Cohort characteristics**

Fmax measurements were available for 215 participants (Table 1), of whom 40% had a z-score below −1 while 13% had a z-score ≤ −2. As for Pmax, results were available for 226 participants: 55% and 27% with z-scores ≤ −1 and ≤ −2, respectively. Thus, the PETALE cohort is more affected in terms of muscle morbidity than the normal population (P < 10⁻⁴).
Association study with WES data

Association analyses were performed between functionally predicted common and rare variants derived from selected candidate genes obtained from WES data. Among common SNPs, rs41270041 located in the ADAMTS4 gene, was positively associated with Fmax z-scores $\leq -2$ ($P=10^{-4}$), whereas rs2001616 in the DUOX2 gene was significantly associated with $P$-max, analysed both as a categorical (z-score $\leq -2$) and continuous variable ($P=10^{-5}$ and $P=4.6\times10^{-5}$, respectively. Table S1).

Rare-SNP analyses identified top-ranking ALOX15 ($P=5.7\times10^{-6}$) gene, associated with continuous Pmax z-scores. The signal originated from three collapsed rare variants in the ALOX15 gene (Table S2).

Confirmation by genotyping and association analyses of obtained genotypes

Participants were screened for top-ranking SNPs, including 2 common SNPs (rs41270041 in ADAMTS4 and rs2001616 in DUOX2) and 3 rare variants (rs113604586, rs34210653 and rs11568142) in ALOX15. Concordance between WES and genotyping data ranged from 92% to 95%. Given that similar results were obtained in association analyses for continuous Pmax z-scores and z-scores $\leq -2$, only the latter are presented. The rs2001616 in the DUOX2 gene was associated with Pmax according to a dominant model of inheritance ($P=2\times10^{-4}$, Table 2). Carriers of a minor G allele were more frequent among participants with a z-score $\leq -2$ as compared with those carrying the AA genotype (Odds ratio (OR)=3.4, 95%...
<table>
<thead>
<tr>
<th>PHENOTYPE</th>
<th>GENE_SNP</th>
<th>COMPLICATION</th>
<th>MODEL</th>
<th>COMPLICATION</th>
<th>OR (95% - CI)</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Unaffected</td>
<td>Affected</td>
<td>Unaffected</td>
<td>Affected</td>
<td></td>
</tr>
<tr>
<td>Pmax (Z ≤ -2)</td>
<td>rs2001616DUOX2 A &gt; G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td></td>
<td>135 (84.9%)</td>
<td>40 (62.5%)</td>
<td>AA</td>
<td>135 (84.9%)</td>
<td>1 0.0002</td>
</tr>
<tr>
<td>AG</td>
<td></td>
<td>20 (12.6%)</td>
<td>23 (35.9%)</td>
<td>AG + GG</td>
<td>24 (15.1%)</td>
<td>3.4 (1.7 - 6.6)</td>
</tr>
<tr>
<td>GG</td>
<td></td>
<td>4 (2.5%)</td>
<td>1 (1.6%)</td>
<td>GG</td>
<td>19 (65.5%)</td>
<td></td>
</tr>
<tr>
<td>Fmax (Z ≤ -2)</td>
<td>rs41270041ADAMTS4 G &gt; C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td></td>
<td>146 (84.4%)</td>
<td>19 (65.5%)</td>
<td>GG</td>
<td>146 (84.4%)</td>
<td>1 0.015</td>
</tr>
<tr>
<td>GC</td>
<td></td>
<td>26 (15.0%)</td>
<td>10 (34.5%)</td>
<td>GC + CC</td>
<td>27 (15.6%)</td>
<td>2.8 (1.2 - 6.8)</td>
</tr>
<tr>
<td>CC</td>
<td></td>
<td>1 (0.6%)</td>
<td>0</td>
<td>CC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: OR; odds ratio, CI; confidence interval.
confidence interval (CI)=1.7–6.6). Likewise, the minor C allele of the rs41270041 in ADAMTS4 was associated with Fmax (z-score ≤−2) according to a dominant model (OR=2.8, 95% CI, 1.2–6.8, \( P=0.02 \), Table 2). We next explored whether this association is influenced by clinical and treatment characteristics. Sex, prognostic risk groups, cranial radiation therapy (CRT), BMI, as well as cumulative GC and MTX dose were all significantly associated with Pmax Z score ≤−2 (\( P≤0.02 \), Table S3). Pmax Z score reduction was seen more frequently in female survivors, in obese subjects and participants who were treated according to the high risk protocol. A significant association was noted as well with cumulative MTX and GC doses and with CRT. This reflected more frequent CRT, higher GC doses, but lower MTX doses administered to patients assigned to the high risk group (\( P<0.00001 \), Table S4). No significant association with demographics or treatment characteristics was found for Fmax. A logistic regression model was thus derived only for DUOX2 rs2001616, which remained significant in the multivariate model (\( P=0.003 \), Table 3) that included, besides genotype, sex, BMI at visit, and risk status (as a surrogate for CRT and higher GC doses). Subsequent stratification analyses were performed and showed that the association of both the DUOX2 rs2001616 and ADAMTS4 rs41270041 polymorphisms with phenotypes was modulated by demographics and treatment characteristics. The association for DUOX2 rs2001616 was limited to females (\( P=0.001 \)), non-obese participants (\( P=0.0001 \)) and patients assigned to the HR group (\( P=0.004 \)) (Figure 1), whereas that of the rs41270041 in ADAMTS4 was more apparent among

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Co-variable</th>
<th>B</th>
<th>P-value</th>
<th>OR (95%-CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pmax (Z≤−2)</td>
<td>DUOX2</td>
<td>1.08</td>
<td>0.003</td>
<td>2.9 (1.4–6.1)</td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>0.87</td>
<td>0.009</td>
<td>2.4 (1.2–4.6)</td>
</tr>
<tr>
<td></td>
<td>BMI at visit</td>
<td>0.61</td>
<td>0.004</td>
<td>1.8 (1.2–2.8)</td>
</tr>
<tr>
<td></td>
<td>Risk</td>
<td>1.1</td>
<td>0.001</td>
<td>3.0 (1.5–5.9)</td>
</tr>
</tbody>
</table>

Notes: DUOX2 gene: dominant model (AG and GG vs AA). Sex: Females are compared to males; BMI: participants given normal (≤24.9 kg/m²), overweight (≥25 kg/m², <30 kg/m²), or obese status (≥30 kg/m²); risk: standard or high.

Figure 1 Stratification result of association for DUOX2 rs2001616. Frequencies of AA and AG+GG genotypes (gray bars) and unaffected (black bars) individuals, as defined by Pmax Z≤−2 SD and Z≥−2 SD, respectively. \( P \)-value with OR (odds ratio) and 95% CI (confidence interval) is presented for each comparison. Number of patients in each genotype group is indicated in brackets with percentages. BMI; non obese: <30 kg/m², obese: ≥30 kg/m².
males ($P=0.02$) and participants that received the HR treatment ($P=0.0001$) (Figure 2). Given the strong correlation between HR and exposure to CRT (Table S4), as well as with higher GC cumulative doses, the similar modulation of the effect is seen when CRT or GC doses are used as stratification variables (data not shown).

An analysis of the four different combinations of the rs34210653, rs113604586 and rs11568142 rare variants in the ALOX15 gene revealed that the association signal resulted from the rs113604586 and rs11568142 combination, with a lowest $P$-value of 0.001 (Table 4). The low frequency of rare variant carriers precluded a stratified analysis.

### Discussion

Studies exploring the role of a genetic component in LAE development in childhood ALL survivors are limited, and this study is the first to focus on skeletal muscle function deficit through a genomic approach. $^{40,41}$

Our analysis revealed an association of rs2001616 in the DUOX2 gene with skeletal muscle Pmax deficit. Survivors carrying a minor allele had a higher probability of Pmax reduction than other genotype groups. DUOX2 (Dual oxidase 2) is involved in muscle and bone metabolisms through its role in thyroid hormone synthesis. The encoded protein acts as a functional constituent of
a peroxide generating system located at the apical membrane of thyroid follicular cells. Several DUOX2 polymorphisms have been associated with congenital hypothyroidism, and inactivating mutations of this gene have recently been shown to play a causal role in very early onset inflammatory bowel diseases. Thyrotoxicosis is known to be a risk factor for hypercalcemia, osteoporosis and increased bone turnover. A study investigating correlations between thyroid hormone levels and skeletal muscle regeneration found that high serum T3 and T4 levels could interfere with the inflammatory process following muscle injuries through up-regulation of macrophage accumulation and cell infiltration. The detected effect of rs2001616 polymorphisms could therefore be due to up-regulation of thyroid hormone production in carriers of the variant allele.

We thus evaluated correlations between the DUOX2 rs2001616 variant and thyroid function in study participants. TSH and free thyroxine (T4) levels were available for 232 and 229 participants with genotype data. A significant association ($P=0.02$) was found between free T4 values and the rs2001616 polymorphism, carriers of a variant allele having a higher median T4 level (12.8 vs 12.2 $\mu$mol/L). Associations with TSH were statistically insignificant. The clinical significance of this result is limited, however, as none of the study participants was diagnosed with dysthyroidism as based on TSH and T4 levels following their ALL treatment. Furthermore, no significant association was found between thyroid hormone levels and skeletal muscle function phenotypes.

An association with altered skeletal muscle function (Fmax) was also demonstrated with rs41270041, located in the ADAMTS4 gene. ADAMTS4 (ADAM metalloproteinase with thrombospondin type 1 motif 4) is a protease-coding gene responsible for the degradation of aggregan, a cartilage-specific proteoglycan core protein. This gene has been shown to contribute to osteoarthritic disorders and tendon injuries. An increased susceptibility to tendinopathies or painful joints could lead to limb immobilization and thus alter muscle mass through lower muscle strain. The ADAMTS4 protein has also been shown to be a risk factor for atherosclerosis plaque vulnerability in mice models.

ALOX15 codes for arachidonate 15-lipoxygenase, a pro-inflammatory protein involved in the arachidonic acid metabolism pathway. It has been reported to play a role in inflammatory response and immunity, and in ischemic heart disease. Studies carried on mice models suggest that ALOX15 could also affect bone mass through its negative regulation of maximum bone mineral density. Associations have been detected between ALOX15 variants and bone micro-architecture in young women, while other variants in this gene seem to predispose to fractures in older women. We detected a significant association for two rare SNPs included in the ALOX15 locus, justifying further study of this gene in related contexts.

A likely explanation why top-ranking SNPs did not all correlate with both Fmax and Pmax phenotypes lies in the distinct skeletal muscle functions reflected by these two measurements. Muscle force represents the ability to accelerate a given mass (Force = Mass $\times$ Acceleration) whereas power is the ability to produce force at a given velocity (Power = Force $\times$ Velocity). Previous studies have demonstrated that peak power, but not peak force, depends on muscle fiber type. On the other hand, evidence suggests a strong correlation between peak muscle force and muscle mass as defined by cross-sectional area measurements. We can thus hypothesize that the ADAMTS4 gene could modulate molecular pathways involved with skeletal muscle size, while ALOX15 and DUOX2 would affect muscle fiber contractile properties.

Associations of skeletal muscle deficits with common variants were more significant in groups known to suffer more from treatment-related complications (e.g. CRT, HR, and high GC doses groups). In fact, various studies demonstrated that muscle function is substantially more degraded in patients exposed to a highly aggressive treatment regimen, including higher glucocorticoid doses and CRT, which may explain the interaction with risk status noted for associated SNPs of DUOX2 and ADAMTS4 genes. A study carried in the St Jude Lifetime Cohort demonstrated severe anterior pituitary deficit related to CRT exposure in childhood cancer survivors. However, none of the patients in the PETALE cohort were diagnosed with secondary clinical hypopituitarism, which may be due to lower overall CRT doses received. It remains possible that a higher proportion of participants in the CRT-positive group suffered from CRT-induced skeletal muscle deficit due to alternative mechanisms involved with neuroendocrine signalling. Indeed, low dose ionising radiation therapy has been demonstrated to rapidly alter synaptic and mitochondrial signalling pathways in murine models, which could account for poorer outcomes, particularly apparent in patients with a given genetic background. The impact of obesity seems to overcome genetic predisposition, as a higher risk of skeletal muscle function deficit in the case of the DUOX2 SNP was seen only among non-obese patients. Studies investigating muscle pathologies in obese individuals have found that
dysregulated adipokines production in hypertrophied adipose tissue decreased the expression of contractile proteins in myotubes, pointing towards an alternative mechanism for poorer skeletal muscle function in this group.39

A limitation of this study results from our representation of skeletal muscle function deficit through muscle power and force phenotypes. Additional measures of muscle quality, architecture, and function and myosteatosis such as muscle cross-sectional area, lean body mass, and metabolic indices could have been used to complement the study phenotypes and better grasp the issue of sarcopenia in childhood leukemia survivors.62

It is not excluded that variables not considered in this study, such as nutritional status, level of physical activity, as well as concomitant endocrine or chronic inflammatory conditions, could have been confounding factors in genetic associations. Some of these factors were not included due to the subjectivity of self-reported phenotype data (e.g. level of physical activity), and others because none of the PETALE cohort participants presented evidence of such conditions (e.g. hypopituitarism). Likewise, although we controlled the analyses as much as possible for potential major confounders, a number of covariables could have still influenced the analyses, such as wide age range, duration of follow-up, or different complications that occurred during the treatment.63,64

It is worth noting that the limited sample size might have reduced the accuracy of the stratified analyses, and may have affected the detection of lower risk effects (the lowest OR that can be detected for common variants after correction for multiple testing in WES dataset is 2.6). On the other hand, cohort homogeneity with respect to treatment and ethnic background reduced the impact of confounding variables on the study results.10 While the candidate gene approach is generally biased towards rational gene selection, the investigation of many such genes allowed the identification of novel variants associated with genetic predictability of skeletal muscle deficit in the targeted population. The analyses of WES data were corrected for multiple testing using a 5% FDR, which in this dataset corresponded well with Bonferroni-adjusted P-values for the number of common variants tested or the number of genes in rare variants testing.

A replication analysis should follow this study to validate findings, whereas future studies conducted at the exome or genome level may detect in an unbiased fashion new genetic predictors underlying skeletal muscle deficit in childhood ALL survivors. Together, they may explain the observed heterogeneity of these late-adverse effects in ALL survivors, and could eventually lead to personalised prevention strategies.

**Abbreviation list**

ALL, acute lymphoblastic leukemia; LAE, late adverse effects; GC, glucocorticoid; BMI, body mass index; FDR, false discovery rate; PETALE, Prévenir les effets tardifs des traitements de la leucémie aiguë lymphoblastique chez l’enfant; SNP, single nucleotide polymorphism; WES, whole exome sequencing; SJUHC, Sainte-Justine University Hospital Center; DFCI, Dana-Farber Cancer Institute; MSHC, Montreal Shriners Hospital for Children; Fmax, peak muscle force; Pmax, peak muscle power; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval; CRT, cranial radiation therapy.

**Acknowledgments**

The authors thank all childhood ALL survivors and their parents who consented to take part in the PETALE study, as well as all study collaborators for their precious involvement in the production of this manuscript. The PETALE study is funded by the Canadian Institutes of Health Research in collaboration with the Cancer Research Society Inc. (CRS), the Garron Family Cancer Center of the Hospital for Sick Children, the Paediatric Oncology Groups of Ontario (POGO), the Canadian Cancer Society Research Institute (CCSRI), the C17 Research Network (C17), the Sainte-Justine Hospital Foundation and the FRQS Applied Medical Genetics Network.

**Disclosure**

The abstract for this manuscript was presented at the 2017 International Workshop on Musculoskeletal and Neuronal Interactions, Montreal, October 6–8, and was published in the Journal of Musculoskeletal and Neuronal Interactions (J Musculoskeletal Neuronal Interact 2017; 17(4):347–357). Dr Nathalie Alos reports grants from CIHR, grants from Cancer Research Society Inc., grants from Pediatric Oncology Groups of Ontario (POGO), grants from Canadian Cancer Society Research Institute (CCSRI), grants from C17 Research Network (C17), grants from Sainte-Justine Hospital Foundation, nothing from FRQS Applied Medical Genetics Network, grants from Garron Family Cancer Center of the Hospital for Sick Children, during the conduct of the study; other from Novartis, other from Amgen, other from Alexion, other from novonordisk, outside the...
submitted work. The other authors report no conflict of interests in this work.

References


# Supplementary materials

## Table S1 Top-ranking common SNPs, WES data

<table>
<thead>
<tr>
<th>SNP</th>
<th>rs41270041</th>
<th>rs2001616</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene</td>
<td>ADAMTS4</td>
<td>DUOX2</td>
</tr>
<tr>
<td>Locus</td>
<td>chr1: 161161284</td>
<td>chr15: 45404066</td>
</tr>
<tr>
<td>Genomic Location</td>
<td>Exon</td>
<td>Exon</td>
</tr>
<tr>
<td>Major allele &gt; minor allele</td>
<td>G&gt;C</td>
<td>A&gt;G</td>
</tr>
<tr>
<td>MAF (%)</td>
<td>8.6</td>
<td>15.5</td>
</tr>
<tr>
<td>Role of protein connected with muscle metabolism</td>
<td>Degradation of aggrecan proteoglycan</td>
<td>Role in thyroid hormone synthesis</td>
</tr>
<tr>
<td>Associated Phenotype</td>
<td>Fmax (Z-score ≤ −2 SDs)</td>
<td>Pmax (Z-score ≤ −2 SDs)</td>
</tr>
<tr>
<td>P-value</td>
<td>1.19×10⁻⁴</td>
<td>1.02×10⁻⁵</td>
</tr>
<tr>
<td>Allelic Ratio</td>
<td>AFFECTED: 12/46</td>
<td>AFFECTED: 33/83</td>
</tr>
<tr>
<td>Allelic OR (95%-CI)</td>
<td>4.24 (1.93–9.31)</td>
<td>3.55 (1.98–6.36)</td>
</tr>
</tbody>
</table>

**Abbreviations:** MAF, minor-allele frequency; OR, Odds ratio; CI, confidence interval.

## Table S2 Top-ranking rare SNPs, WES data

<table>
<thead>
<tr>
<th>Gene</th>
<th>ALOX15</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNPs</td>
<td></td>
</tr>
<tr>
<td>rs34210653</td>
<td>rs11568142</td>
</tr>
<tr>
<td>MAF (%)</td>
<td>0.78</td>
</tr>
<tr>
<td>Major allele &gt; minor allele</td>
<td>G &gt; A</td>
</tr>
<tr>
<td>Genomic location</td>
<td>Exon</td>
</tr>
<tr>
<td>Gene locus</td>
<td>17p13.2</td>
</tr>
<tr>
<td>Role of protein</td>
<td>Arachidonic acid metabolism pathway</td>
</tr>
<tr>
<td>Associated phenotype</td>
<td>Pmax (Z-score)</td>
</tr>
<tr>
<td>P-value</td>
<td>5.7×10⁻⁶</td>
</tr>
<tr>
<td>Participants with &gt; 1 minor allele</td>
<td>9 (4.05%)</td>
</tr>
</tbody>
</table>

**Abbreviation:** MAF, minor-allele frequency.
Table S3 Co-variable association with Pmax (Z-score ≤ −2)

<table>
<thead>
<tr>
<th>Category</th>
<th>Group</th>
<th>Affected (%)</th>
<th>Unaffected (%)</th>
<th>P-value</th>
<th>OR (95%-CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>Male</td>
<td>22 (20.4)</td>
<td>86 (79.6)</td>
<td>0.008</td>
<td>2.2 (1.2–4.1)</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>43 (33.9)</td>
<td>75 (63.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk level</td>
<td>Standard</td>
<td>17 (16.2)</td>
<td>88 (83.8)</td>
<td>0.0001</td>
<td>3.4 (1.8–6.4)</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>48 (39.7)</td>
<td>73 (60.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTX cumulative dose (^a) (mg/m(^2))</td>
<td>≤6723</td>
<td>40 (35.1)</td>
<td>74 (64.9)</td>
<td>0.03</td>
<td>0.5 (0.3–0.9)</td>
</tr>
<tr>
<td></td>
<td>&gt;6723</td>
<td>24 (21.8)</td>
<td>86 (78.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC cumulative dose (pred. Eq. mg/m(^2))</td>
<td>≤9025</td>
<td>23 (20.4)</td>
<td>90 (79.6)</td>
<td>0.006</td>
<td>2.3 (1.3–4.2)</td>
</tr>
<tr>
<td></td>
<td>&gt;9025</td>
<td>41 (36.9)</td>
<td>70 (63.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRT</td>
<td>No</td>
<td>15 (16.3)</td>
<td>77 (83.7)</td>
<td>0.001</td>
<td>3.1 (1.6–5.9)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>50 (37.3)</td>
<td>84 (62.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI category at visit</td>
<td>Normal</td>
<td>31 (20.7)</td>
<td>119 (79.3)</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Overweight</td>
<td>21 (47.7)</td>
<td>23 (52.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Obese</td>
<td>13 (40.6)</td>
<td>19 (59.4)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: \(^a\)Cumulative MTX and GC doses are recoded as below and above median.

Abbreviations: CRT, cranial radiation therapy; MTX, methotrexate; GC, glucocorticoids; BMI, body mass index.

Table S4 Association of risk groups with CRT and cumulative MTX and GC dose

<table>
<thead>
<tr>
<th>Risk</th>
<th>Standard</th>
<th>High</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRT</td>
<td>No</td>
<td>90 (81.1%)</td>
<td>9 (6.9%)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>21 (18.9%)</td>
<td>122 (93.1%)</td>
</tr>
<tr>
<td>MTX Cumulative(^1) (mg/m(^2))</td>
<td>≤6723</td>
<td>34 (30.6%)</td>
<td>86(66.4%)</td>
</tr>
<tr>
<td></td>
<td>&gt;6723</td>
<td>77 (69.4%)</td>
<td>43 (33.6%)</td>
</tr>
<tr>
<td>GC Cumulative (mg/m(^2))</td>
<td>≤9025</td>
<td>106 (95.5%)</td>
<td>14 (10.9%)</td>
</tr>
<tr>
<td></td>
<td>&gt;9025</td>
<td>5 (4.5%)</td>
<td>115 (89.1%)</td>
</tr>
</tbody>
</table>

Note: \(^1\)Cumulative MTX and GC doses are recoded as above and below median.

Abbreviations: CRT, cranial radiation therapy; MTX, methotrexate; GC, glucocorticoids.