DTU Library

Trypsin encoding PRSS1-PRSS2 variation influence the risk of asparaginase-associated pancreatitis in children with acute lymphoblastic leukemia: a Ponte di Legno toxicity working group report

Wolthers, Benjamin O.; Frandsen, Thomas L.; Patel, Chirag J; Abaji, Rachid; Attarbaschi, Andishe; Barzilai, Shlomit; Colombini, Antonella; Escherich, Gabriele; Grosjean, Marie; Krajinovic, Maja

Ferrata Storti Foundation

Trypsin-encoding *PRSS1-PRSS2* variations influence the risk of asparaginase-associated pancreatitis in children with acute lymphoblastic leukemia: a Ponte di Legno toxicity working group report

Benjamin O. Wolthers,[1] Thomas L. Frandsen,[1] Chirag J. Patel,[2] Rachid Abaji,[3] Andishe Attarbaschi,[4] Shlomit Barzilai,[5] Antonella Colombini,[6] Gabriele Escherich,[7] Marie Grosjean,[8] Maja Krajinovic,[3,9] Eric Larsen,[10] Der-Cherng Liang,[11] Anja Möricke,[12] Kirsten K. Rasmussen,[1] Sujith Samarasinghe,[13] Lewis B. Silverman,[14] Inge M. van der Sluis,[15] Martin Stanulla,[16] Morten Tulstrup,[1] Rachita Yadav,[8,17] Wenjian Yang,[18] Ester Zapotocka,[19] Ramneek Gupta[8] and Kjeld Schmiegelow[1,20] on behalf of the Ponte di Legno toxicity working group

[1]Department of Pediatrics and Adolescent Medicine, University Hospital Rigshospitalet, Copenhagen, Denmark; [2]Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA; [3]CHU Sainte-Justine Research Center and Department of Pharmacology, University of Montreal, QC, Canada; [4]Department of Pediatric Hematology and Oncology, St Anna Children's Hospital and Department of Pediatric and Adolescent Medicine, Medical University of Vienna, Austria; [5]Pediatric Hematology and Oncology, Schneider Children's Medical Center of Israel, Petah-Tikva, Israel and Sackler Faculty of Medicine, Tel Aviv University, Israel; [6]Department of Pediatrics, Ospedale San Gerardo, University of Milano-Bicocca, Fondazione MBBM, Monza, Italy; [7]University Medical Center Eppendorf, Clinic of Pediatric Hematology and Oncology, Hamburg, Germany; [8]Center for Biological Sequence Analysis, Technical University of Denmark, Lyngby, Denmark; [9]Department of Pediatrics, University of Montreal, QC, Canada; [10]Maine Children's Cancer Program, Scarborough, ME, USA; [11]Division of Pediatric Hematology-Oncology, Mackay Memorial Hospital, Taipei, Taiwan; [12]Christian-Albrechts-University Kiel and University Medical Center Schleswig-Holstein, Department of Pediatrics, Kiel, Germany; [13]Great Ormond Street Hospital for Children, London, UK; [14]Department of Pediatric Oncology, Dana-Farber Cancer Institute and Division of Hematology/Oncology, Boston Children's Hospital, Boston, MA, USA; [15]Dutch Childhood Oncology Group, The Hague and Princess Máxima Center for Pediatric Oncology, Utrecht, the Netherlands; [16]Department of Pediatric Hematology and Oncology, Hannover Medical School, Germany; [17]Molecular Neurogenetics Unit, Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA; [18]St. Jude Children's Research Hospital, Department of Pharmaceutical Sciences, Memphis, TN, USA; [19]University Hospital Motol, Department of Pediatric Hematology/Oncology, Prague, Czech Republic and [20]Institute of Clinical Medicine, University of Copenhagen, Denmark.

ABSTRACT

Asparaginase-associated pancreatitis is a life-threatening toxicity to childhood acute lymphoblastic leukemia treatment. To elucidate genetic predisposition and asparaginase-associated pancreatitis pathogenesis, ten trial groups contributed remission samples from patients aged 1.0–17.9 years treated for acute lymphoblastic leukemia between 2000 and 2016. Cases (n=244) were defined by the presence of at least two of the following criteria: (i) abdominal pain; (ii) levels of pancreatic enzymes ≥3 x upper normal limit; and (iii) imaging compatible with pancreatitis. Controls (n=1320) completed intended asparaginase therapy, with 78% receiving ≥8 injections of pegylated-asparaginase, without developing asparaginase-associated pancreatitis. rs62228256 on 20q13.2 showed the strongest association with the development of asparaginase-associated pancreatitis (odds ratio=3.75; $P=5.2 \times 10^{-8}$). Moreover, rs13228878 (OR=0.61; $P=7.1 \times 10^{-6}$) and rs10273639 (OR=0.62; $P=1.1 \times 10^{-5}$) on 7q34 showed significant association with the risk of asparaginase-associated pancreatitis. A Dana Farber Cancer Institute ALL Consortium cohort consisting of

patients treated on protocols between 1987 and 2004 (controls=285, cases=33), and the Children's Oncology Group AALL0232 cohort (controls=2653, cases=76) were available as replication cohorts for the 20q13.2 and 7q34 variants, respectively. While rs62228256 was not validated as a risk factor (P=0.77), both rs13228878 (P=0.03) and rs10273639 (P=0.04) were. rs13228878 and rs10273639 are in high linkage disequilibrium (r^2=0.94) and associated with elevated expression of the *PRSS1* gene, which encodes for trypsinogen, and are known risk variants for alcohol-associated and sporadic pancreatitis in adults. Intra-pancreatic trypsinogen cleavage to proteolytic trypsin induces autodigestion and pancreatitis. In conclusion, this study finds a shared genetic predisposition between asparaginase-associated pancreatitis and non-asparaginase-associated pancreatitis, and targeting the trypsinogen activation pathway may enable identification of effective interventions for asparaginase-associated pancreatitis.

Introduction

Intensification of chemotherapy for childhood acute lymphoblastic leukemia (ALL) has generated 5-year survival rates greater than 90%, but has been associated with an increase in therapy-related toxicity.[1] Asparaginase is a key drug in the treatment of childhood ALL and there is growing interest in its use as an anti-metastatic agent in breast cancer.[2] Asparaginase depletes the body of the non-essential amino acid asparagine through deamidation of asparagine into aspartic acid and ammonia,[3] and targets protein synthesis in malignant lymphoblasts by impairing the ability to synthesize asparagine.[4,5] Pancreatitis associated with asparaginase therapy (AAP) is a frequent toxicity affecting 4–10% of children treated on contemporary ALL protocols, and is associated with severe complications.[6–9] In addition, re-exposure to asparaginase after AAP is associated with a high risk (~50%) of a second episode of AAP, and thus AAP often entails truncation of asparaginase therapy, thereby decreasing the patients' chance of survival.[4,5,9] The mechanism(s) by which asparaginase causes pancreatitis are elusive, thus hampering attempts to identify patients with an altered risk of AAP.[10] The Ponte di Legno toxicity working group, therefore, initiated a study with three main purposes: (i) to define diagnostic consensus criteria for AAP;[11] (ii) to describe the phenotype of AAP in patients across multiple ALL trial groups;[9] and (iii) to explore genotype-phenotype associations, using a genome-wide approach, to identify patients with altered risk of AAP.[9,11] Genome-wide association (GWA) studies are agnostic by design, reporting phenotype-genotype associations without prior hypotheses and often including speculative mechanisms. Replication of GWA study results are thus a requisite for credibility Accordingly, this study presents results from the largest AAP GWA study so far, with a strong focus on investigating previously validated variants associated with non-asparaginase-induced pancreatitis and replicating top results in similar childhood ALL cohorts.

Methods

Study design and participants

Ten international childhood ALL trial groups (*Online Supplementary Table S1*) contributed to the discovery cohort. Post-remission DNA was collected from June 2015 to January 2017, three groups collected DNA on AAP cases only while seven groups did so on both cases and controls (*Online Supplementary Figure S1*). The database containing phenotype data was approved by the regional ethical review board of The Capital Region of Denmark (H-2-2010-022), the Danish Data Protection Authorities (j.nr.: 2012-58-0004), and by relevant regulatory authorities in all participating countries. Genotype data were stored at the Technical University of Denmark's server Computerome.[12]

Children (aged 1.0–17.9 years) with newly diagnosed ALL between January 2000 and January 2016 were eligible, irrespective of ethnicity. Pancreatitis was defined as asparaginase-associated if diagnosed within 50 days of the last injection of native *E. coli* asparaginase or polyethylene glycolated *E. coli* asparaginase (PEG-asparaginase) and cases fulfilled the Ponte di Legno toxicity working group consensus definition for AAP: i.e., at least two of (i) amylase, pancreatic amylase, or pancreatic lipase ≥3 x upper normal limit; (ii) abdominal pain; and (iii) imaging compatible with imaging compatible with pancreatitis. All controls received the planned amount of asparaginase therapy in their respective protocols, with more than 78% (1024/1320) receiving at least eight injections of PEG-asparaginase without developing AAP. A subset of 62 AAP cases was previously included in a Nordic Society of Pediatric Hematology and Oncology (NOPHO) GWA study.[13] These samples were genotyped on identical genotyping arrays as the remaining cohort, and raw genotyping data on these patients were pooled with those of the remaining cohort prior to quality control, and association analyses were done in one cohort.

Genotyping

Post-remission DNA was genotyped by Aros Applied Biotechnology A/S (Aarhus, Denmark) on Illumina Omni2.5exome-8 BeadChip arrays using the human genome assembly (*GRCh37*) for reference. Quality control was performed using the PLINK tool,[14] and single nucleotide polymorphisms (SNPs) were annotated in Ensembl Variant Effect Predictor GCRCh37.[15] Alleles given are refSNP alleles according to dbSNP (not necessarily the alleles supplied by the Illumina map).[16]

Quality control

Quality control was performed according to previously published criteria[17] (*Online Supplementary Figures S2-4*), excluding individuals with: (i) a discordance in number of X chromosomes between geno- and phenotypes; (ii) missing data on >3% of SNPs; (iii) excess heterozygosity between autosomal SNPs; and (iv) high relatedness between samples. SNPs were excluded based on: (i) missing data on >2% of individuals (call rate); (ii) Hardy-Weinberg equilibrium; (iii) minor allele frequency <0.01; (iv) difference in call rate between cases and controls (Fisher exact test P<1.10^{-5}); and (v) duplicated genomic position.

Replication

Three top SNPs were tested for validation in two separate cohorts. The Children's Oncology Group AALL0232 ALL cohort included previously genotyped data on *PRSS1-2* variants (but not on the *NFATC2* variant). The *NFATC2* variant was genotyped *de novo* in a cohort of patients from the Dana Farber Cancer Institute (DFCI) ALL Consortium protocols 87–01, 91–01, 95–01 and 00–01 (1987-2004). The AALL0232 ALL cohort included 76 cases diagnosed using National Cancer Institute's Common Terminology Criteria for Adverse Events (CTCAE criteria) (*Online Supplement pages 6–9*) and 2577 controls.[8] The cohort is described in detail in the appendix.[8] The AALL0232 cohort was genotyped on the Affymetrix Genome-Wide Human SNP 6.0 Array, and imputed genotypes were generated using 1000 Genomes as the reference population as reported (ref: PMID 26265699). Time-dependent analysis (Cox proportional hazards regression) was performed adjusting for age and ancestry. The cohort of patients treated on the DFCI ALL Consortium protocols 87–01, 91–01, 95–01 and 00–01 received 20–30 weeks of post-induction asparaginase therapy (*Online Supplement pages 9-11*). Thirty-three cases diagnosed according to the CTCAE criteria and 285 controls were included in this cohort, and genotyped by allele-specific oligonucleotide hybridization as described elsewhere.[18,19] The Pearson correlation coefficient was used to investigate associations between genotype and AAP.

Statistical analysis

Association analysis was done in PLINK using logistic regression, assuming an additive genetic model, and adjusting for genetic ancestry and age. Genetic ancestry was determined by clustering analysis, and non-CEU ancestry was defined as individuals >16 standard deviations away from the HapMap-defined CEU (Northern and Western European) centroid mean. Using this model, multidimensional scaling plots showed an equal distribution of cases and controls according to ancestry (*Online Supplementary Figure S3*) and QQ plots showed no sign of population substructure (λ = 1.02, *Online Supplementary Figure S4*). Statistical analysis of phenotype and genotype associations was performed in the statistical program R version 3.3.3.[20] Linear and logistic regression analyses were used to identify associations between genotypes and continuous or categorical clinical variables adjusting for age and ancestry. Genotype was treated as a numerical value (0, 1 or 2 minor alleles) for additive effect. The Kaplan-Meier method was used to estimate probability of event-free survival according to genotypes, and differences were compared with the two-sided log-rank test. No prior sample size calculations were applied for pre-study power calculations. Two-sided *P*-values below 0.05 were regarded as statistically significant. SNPs were annotated to genes 10 kb up- or downstream from transcription start- and end-sites, respectively, and all SNPs with a *P*-value below 5 x 10[-5] were manually inspected for associations with genes and pathways previously related to pancreatitis. Investigated SNPs were explored using dbSNP[16] and Ensembl,[15] linkage disequilibrium (LD) between SNPs by the National Cancer Institute LDassoc tool,[21] tissue expression (expressive quantitative trait loci) by GTeX,[22] regulatory effect by RegulomeDB,[23] and regional association plots were produced by the LocusZoom tool.[24] Genes and SNPs previously associated with pancreatitis were investigated by searching PubMed for reports published in English within the last 10 years, using the search terms "pancreatitis" AND "genome" OR "genetic" OR "genotype" in the title. Gene functions were defined by Genecards[25] (www.genecards.org).

Results

After quality control filtering, 244 cases, 1320 controls and 1401908 SNPs were eligible for association analysis. Two hundred and five of the 244 (84%) cases and 1185/1320 (90%) of controls were of European (CEU) ancestry (*Online Supplementary Figure S4*). The median age (interquartile range, IQR) of the cases was 8.1 years (IQR, 4.3 to 13.1) while that of the controls was 5.0 years (IQR, 3.0 to 9.0). Fifty-five percent of both cases (133/244) and controls (724/1320) were male.

Figure 1 shows the significance of SNPs associated with
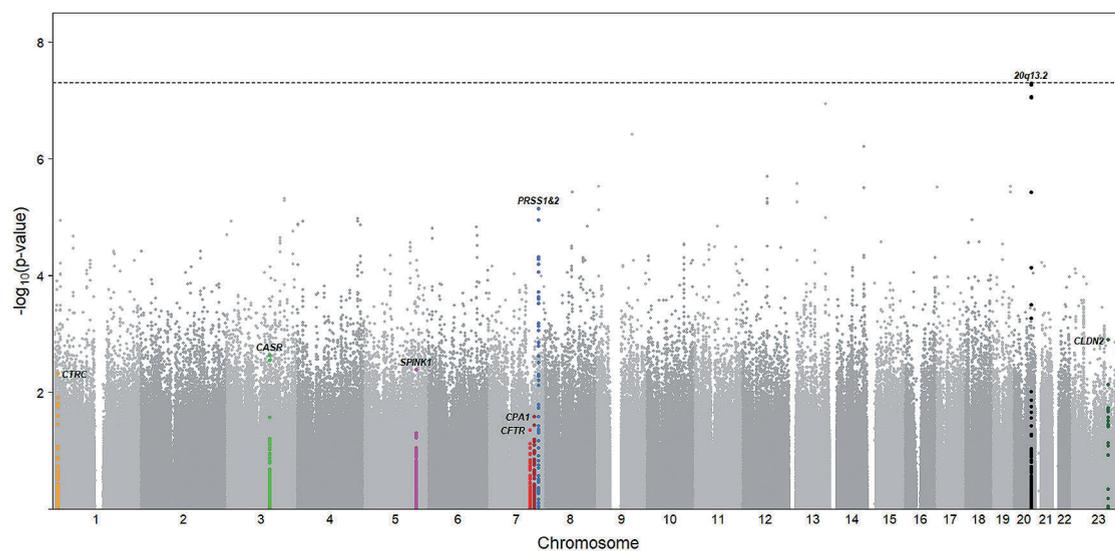
Figure 1.Manhattan plot. Manhattan plot showing single nucleotide polymorphisms (SNPs) associated with asparaginase-associated pancreatitis in 244 cases and 1320 controls. The x axis represents genomic location, and the y axis represents the *P* value for the SNP associations calculated using logistic regression adjusting for age and ancestry. Genes previously associated with pancreatitis are represented in color. SNPs are annotated to genes based on genomic location (10 kb upstream and downstream of the transcription start site and transcription terminator, respectively. The human assembly *GRCh37* was used for reference.

AAP. The variant rs62228256 [reference allele=C, minor allele=T (C>T)] on 20q13.2 showed the strongest association with AAP [odds ratio (OR), 3.75; 95% confidence interval (95% CI): 2.33 to 6.04; P=5.2 x 10^{-8}]. rs62228256 is located 274 kilobase pairs upstream of Nuclear factor of activated T cells (*NFATC2*) and has been documented to be an expressive quantitative trait locus for this gene in pancreatic tissue (*Online Supplementary Figure S5*).

The 30 SNPs most associated with AAP, with *P*-values of 5 x 10^{-5} or lower, included rs13228878 (A>G; OR, 0.61; 95% CI: 0.5 to 0.76; P=7.1 x 10^{-6}) and rs10273639 (C>T; OR, 0.62; 95% CI: 0.5 to 0.77; P=1.1 x 10^{-5}) (Table 1). These SNPs reside on the same haplotype and are in high LD (CEU population LD; r^2=0.94) in the *PRSS1-PRSS2* locus on chromosome 7 (Figure 2). *PRSS1* and *PRSS2* encode for the proteases cationic and anionic trypsinogen, respectively. Both minor alleles rs13228878_G and rs10273639_T reduce the risk of AAP. When performing association analysis in the CEU population of 205 AAP cases and 1185 controls, rs62228256 (OR, 3.75; 95% CI: 2.27 to 6.2;

P=2.47 x 10^{-7}), rs13228878 (OR, 0.60; 95% CI: 0.48 to 0.76; P=2.1 x 10^{-5}) and rs10273639 (OR, 0.62; 95% CI: 0.49 to 0.78; P=3.8 x 10^{-5}) (*Online Supplementary Figures S6-8* and *Online Supplementary Table S2*) remained strongly associated with AAP. Further investigation of previously validated SNPs within genes known to regulate trypsin activation[26] (*Online Supplementary Table S3*) showed that AAP was associated with rs17107315 in pancreatic secretory trypsin inhibitor (*SPINK1*; OR, 2.87; 95% CI: 1.36 to 5.8; P=4 x 10^{-3}), rs10436957 in chymotrypsin C (*CTRC*; OR, 0.69; 95% CI: 0.53 to 0.89; P=5 x 10^{-3}) and rs4409525 in Claudin-2 (*CLDN2*; OR, 1.41; 95% CI: 1.08 to 1.83; P=0.01) with all minor alleles altering AAP risk in the direction and with effects very similar to those previously reported.

In a logistic regression model, testing whether the effect of the most associated *PRSS1-PRSS2* variant (rs13228878) was modified by rs17107315 (*SPINK1*), rs10436957 (*CTRC*) and rs4409525 (*CLDN2*), no significant interactions were identified (P=0.48, P=0.95 and P=0.93, respectively).

Table 1. Top 30 single nucleotide polymorphisms.

SNP	Chr	Position	Major>minor allele	MAF cases	MAF controls	OR (95% CI)	*P*-value	Gene (distance from gene)
rs62228256	20	50454447	C>T	0.07	0.02	3.75 (2.33-6.04)	5.18x10^{-8}	
rs7270119	20	50436587	A>G	0.07	0.02	3.64 (2.28-5.8)	5.52x10^{-8}	
rs16996276	20	50455925	A>C	0.07	0.02	3.64 (2.27-5.85)	8.64x10^{-8}	
rs62228230	20	50445082	G>A	0.07	0.02	3.54 (2.23-5.63)	9.19x10^{-8}	
rs934350	13	103589776	A>G	0.32	0.22	1.84 (1.47-2.3)	1.16x10^{-7}	
rs170623	9	101984936	C>G	0.38	0.28	1.78 (1.42-2.22)	3.93x10^{-7}	*ALG2 (+0.69 kb) & SEC61B (0)*
rs75245362	14	95990645	C>T	0.11	0.05	2.46 (1.73-3.51)	6.33x10^{-7}	*SCARNA13 (-9.05 kb) & SNHG10 (-8.6 kb)*
rs368819120	12	71747240	G>-	0.48	0.37	1.67 (1.35-2.06)	2.04x10^{-6}	
rs4769201	13	22698015	A>G	0.04	0.01	4.66 (2.45-8.86)	2.67x10^{-6}	
rs7851954	9	6796167	C>T	0.28	0.40	0.59 (0.47-0.74)	3.02x10^{-6}	*KDM4C (0)*
rs61734424	19	50747533	T>C	0.08	0.04	2.75 (1.8-4.2)	3.04x10^{-6}	*MYH14 (0)*
rs9912225	17	4680732	A>G	0.10	0.05	2.42 (1.67-3.5)	3.10x10^{-6}	*TM4SF5 (0) & VMO1 (-7.8 kb)*
rs7155612	14	95976755	T>G	0.12	0.06	2.24 (1.6-3.15)	3.15x10^{-6}	
rs2167730	8	78103417	T>C	0.24	0.34	0.57 (0.45-0.73)	3.69 x10^{-6}	
rs80170196	19	50747159	C>T	0.08	0.04	2.73 (1.78-4.18)	3.77x10^{-6}	*MYH14 (0)*
rs62228228	20	50443845	G>A	0.06	0.02	3.15 (1.94-5.13)	3.80x10^{-6}	
rs5010616	12	71748290	C>T	0.48	0.37	1.63 (1.32-2)	4.76x10^{-6}	
rs12494164	3	164967758	A>C	0.23	0.14	1.77 (1.39-2.27)	4.83x10^{-6}	
rs16848986	3	164979570	T>C	0.22	0.14	1.77 (1.38-2.26)	5.19x10^{-6}	
rs34375180	12	71779640	G>A	0.49	0.38	1.61 (1.31-1.98)	5.53x10^{-6}	
rs7139808	13	22693228	T>C	0.04	0.01	4.59 (2.38-8.87)	5.59x10^{-6}	
rs12582343	12	71766297	A>G	0.48	0.38	1.62 (1.31-1.99)	5.90x10^{-6}	
rs13228878	7	142473466	A>G	0.35	0.44	0.61 (0.5-0.76)	7.06x10^{-6}	*PRSS2 (-6.44 kb) & PRSS3P2 (-5.29 kb)*
rs6477109	9	6794938	C>T	0.29	0.40	0.61 (0.49-0.75)	7.64x10^{-6}	*KDM4C (0)*
rs74109922	13	103582300	A>G	0.12	0.06	2.13 (1.52-2.97)	1.03x10^{-5}	
rs1505495	4	172973580	A>C	0.10	0.17	0.48 (0.34-0.66)	1.06x10^{-5}	*GALNTL6 (0)*
rs1791520	18	22118315	T>C	0.32	0.23	1.65 (1.32-2.07)	1.12x10^{-5}	
rs10273639	7	142456928	C>T	0.35	0.44	0.62 (0.5-0.77)	1.13x10^{-5}	*PRSS1 (-0.4 kb)*
rs4655107	1	23094454	G>A	0.13	0.24	0.53 (0.39-0.7)	1.16x10^{-5}	*EPHB2 (0)*
rs55634345	4	19846813	G>A	0.46	0.34	1.58 (1.29-1.94)	1.19x10^{-5}	

Top 30 SNPs most associated with AAP in 244 cases and 1320 controls. The model used here includes covariates for age and genetic ancestry. SNPs were annotated to genes if ≤10 kb upstream (-) or downstream (+) of the transcription start site or transcription terminator, respectively. SNP: single nucleotide polymorphism; Chr,: chromosome; MAF: minor allele frequency; OR: odds ratio; CI: confidence interval.

Validation of results

In validating our top SNP in the *NFATC2* locus (rs62228256) and top SNPs in the *PRSS1-PRSS2* locus (rs13228878 and rs10273639) we used two cohorts of children with ALL: one treated according to the DFCI ALL Consortium 87–01, 91–01, 95–01 and 00–01 protocols,[27,28] and the other according to Children's Oncology Group AALL0232 protocol[8] (*Online Supplement pages 6−12*). Whereas the association between AAP and rs6228256 genotype was not replicated in the DFCI ALL Consortium cohort (*P*=0.77, *Online Supplement page 10*), both rs13228878 (hazard ratio, 0.68; 95% CI: 0.48 to 0.96; *P*=0.03) and rs10273639 (hazard ratio, 0.69; 95% CI: 0.49 to 0.98; *P*=0.04) were significantly associated with the risk of developing pancreatitis in the AALL0232 cohort (*Online Supplement page 8*).

Two previous studies have investigated SNPs associated with pancreatitis in children with ALL. Using different diagnostic criteria for pancreatitis and including controls with less than 5 weeks of asparaginase therapy, Liu *et al.* associated the rare variant (general population minor allele frequency=0.009%) rs199695765 in carboxypeptidase A2 encoding *CPA2* (hazard ratio, 587; 95% CI: 66.8 to 5166; *P*=9 x 10^{-9}) with AAP.[8] We did not directly genotype this SNP, and none of the genotyped SNPs in the *CPA2* region was in LD with rs199695765. Of 32 SNPs 10 kb up- and down-stream of *CPA2*, rs66839817 (T>C, OR, 1.28; 95% CI: 1.03 to 1.57; *P*=0.02) showed the strongest association.

In a NOPHO GWA study we previously found the *ULK2* variant rs281366 (*P*=5.8 x 10^{-7}) and *RGS6* variant rs17179470 (*P*=1.3 x 10^{-6}) to be most associated with AAP.[13] Excluding cases and controls from the NOPHO study, we failed to validate these results in 184 cases and 712 controls. Both rs281366 and rs17179470 were directly genotyped in this cohort with non-significant *P*-values (*P*=0.84 and *P*=0.32, respectively).

Genotype-phenotype associations

Table 2 shows associations between *PRSS1-PRSS2* genotype and amount of PEG-asparaginase given prior to AAP, days from PEG-asparaginase injection to AAP, complications of pancreatitis, and risk of a second episode of AAP after re-exposure to asparaginase. The risk allele was not associated with number of PEG-asparaginase injections prior to AAP or time from injection of PEG-asparaginase to diagnosis of AAP. Furthermore, the risk of developing acute complications was not found to be associated with *PRSS1-PRSS2* genotype (Table 2). In the Nordic subset of cases (n=92) and controls (n=1024) we found no association between *PRSS1-PRSS2* genotype and 5-year event-free survival (*Online Supplementary Figure S9*; *P*=0.4). Out of 46 children who were re-exposed to asparaginase, 17 (37%) developed a second episode of AAP; although not in a statistically significant manner, the *PRSS1-PRSS2* minor allele indicated the same protective effect as in the risk of initial AAP (OR, 0.49; 95% CI: 0.15 to 1.41; *P*=0.20).

Discussion

This study is the first to find and validate variants in the *PRSS1-PRSS2* locus associated with the risk of AAP in children with ALL. In doing so, we found that activation of trypsin within pancreatic acinar cells is a key initiating event in the pathogenesis of pancreatitis, regardless of the exposure i.e. alcohol, hyperlipidemia, or asparaginase. The role of trypsin activation in the pathogenesis of pancreatitis had long been suspected, but an underlying genetic susceptibility was not documented until 1996 when Whitcomb *et al.* documented mutations in the *PRSS1* gene causing hereditary pancreatitis,[29] and later associated a common genetic variant in the *PRSS1-PRSS2* locus (rs10273639) with the risk of alcohol-related and sporadic pancreatitis.[30] This association was recently validated in larger European and Asian cohorts, and the haplotype has been studied in detail.[31–33] rs10273639 is located 408 base pairs upstream of the translation initiation codon of cationic trypsinogen. A recent functional study documented that the proximal rs4726576 (C>A) variant (204 kb upstream of the translation initiation codon) drives the association.[34] The rs4726576 and rs10273639 variants are in high LD (r^2 >99%) in European and Asian populations but have a r^2 = 0.8 in the African meta-population, and
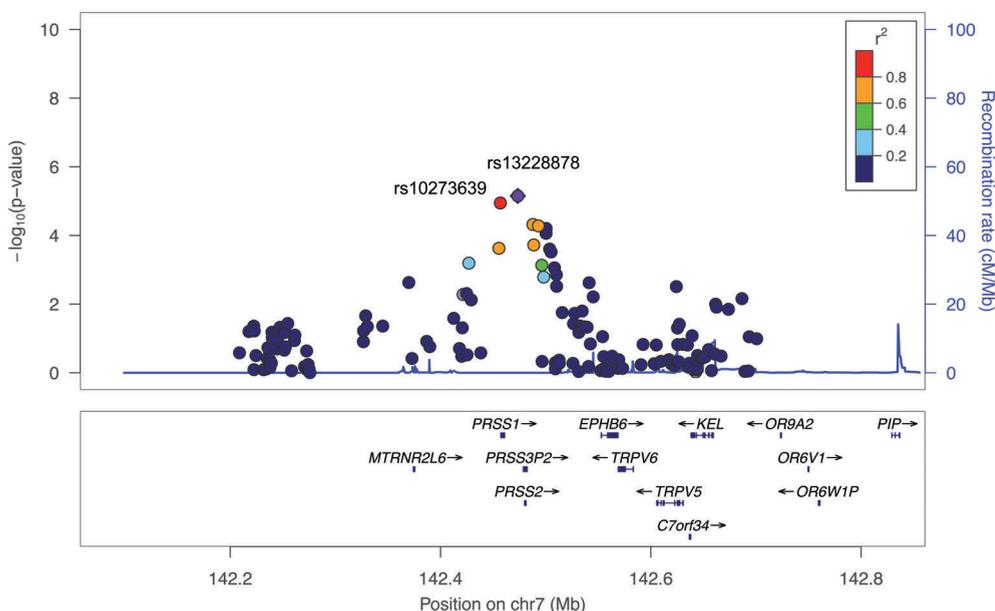

Figure 2. Regional association plot of the *PRSS1-2* locus on chromosome 7. Regional association plot showing single nucleotide polymorphisms (SNPs) associated with asparaginase-associated pancreatitis in 244 cases and 1320 controls. The *x* axis represents genomic location, and the *y* axis represents the *P* values for the SNP associations calculated using logistic regression adjusting for age and ancestry. rs13228878 (*P*=7.1 x 10^{-6}) is represented in purple and rs10273639 (*P*=1.1 x 10^{-5}) in red. The color of the dots reflects linkage disequilibrium (LD) of the genotyped SNPs. LD is based on 1000 genomes European samples, November 2014. The human assembly *GRCh37* was used for reference.

pinpointing the association driving the signal is thus of importance in the latter population. Sequencing the *PRSS1-PRSS2* risk allele has not revealed disease-associated coding variants accounting for the association with pancreatitis in GWA studies; however, the risk allele (rs4726576; rs10273639) is an expressive quantitative trait

locus for *PRSS1* shown to elevate gene expression.[30,34]

The activation of trypsinogen is mediated by cleavage of the N-terminus extension of cationic trypsinogen (a calcium binding site) to active trypsin. Higher levels of calcium have been shown to lead to trypsin activation, and with higher expression of trypsin as seen in patients with
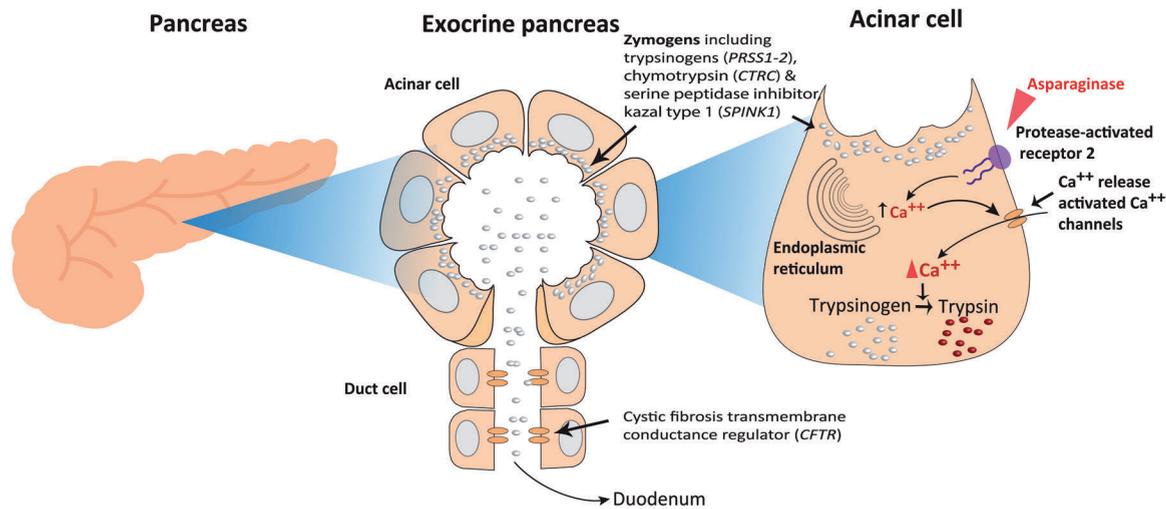
Figure 3. Schematic drawing illustrating the effects of asparaginase on pancreatic acinar cells. Schematic drawing illustrating the likely effect of asparaginase (red triangle) on the protease-activated receptor 2 (PAR2) receptor, and how this leads to increased calcium (Ca^{++}) efflux from the endoplasmic reticulum. This in turn leads to opening of Ca^{++} release activated (CRAC) channels, further increasing intracellular calcium levels, reducing ATP levels and allowing activation of inactive trypsinogen to active trypsin. The drawing is heavily influenced by diagrams by Peng *et al.* (Phil Trans Royal Soc, 2015) and Whitcomb et al. (Nature Gen, 2012).

Table 2. Baseline characteristics, pancreatic enzyme levels and complications of pancreatitis in children according to the *PRSS1-2* (rs13228878) genotype.

	Homozygote for rs13228878_A risk allele n=104	Heterozygote for rs13228878_A risk allele n=109	Homozygote for rs13228878_G non- risk allele n=30	*P*
Age (years), median (median; IQR)	7.75 (4–12.06)	9.21 (4.46–13.43)	6.22 (4.9–12.1)	0.27
Sex, % males n. males/total))	53% (55/104)	55% (60/109)	57% (17/30)	0.92
Total amylase at diagnosis of AAP (U/L) median (IQR; n. available data)	397 (263–673; n= 43)	382 (206–671; n=51)	222 (151–617; n=14)	0.33
Lipase at diagnosis of AAP (U/L) median (IQR; n. available data)	1255 (758–2140; n=39)	1096 (415–1754; n=48)	867 (193–2862; n=11)	0.24
Days from last PEG- asparaginase exposure to AAP, median (IQR; n. available data)	10.5 (6–16; n=70)	10.5 (6–14; n=82)	11 (8–13.25; n=16)	0.96
Number of PEG-asparaginase administrations prior to AAP, median (IQR; n. available data)	4 (2–7; n=74)	3 (2–6; n=85)	3.5 (1.25–6.75; n=18)	0.34
Assisted mechanical ventilation, % (yes (n.)/available data (n.))	5% (5/93)	7% (7/101)	0% 0/26	0.38
Acute insulin therapy, % (yes (n.)/available data (n.))	21% (17/80)	23% (20/87)	18% (4/22)	0.88
Pseudocysts, % (yes (n.)/available data (n.))	25% (22/89)	31% (31/100)	8% (2/24)	0.07
Death, % (yes (n.)/available data (n.))	3% (3/93)	1% (1/101)	0% (0/27)	0.38

Baseline characteristics, pancreatic enzyme levels and complications of pancreatitis according to rs13228878 genotype. Differences among groups were analyzed with the Kruskal-Wallis rank sum test (continuous variables) and chi-square test (categorical variables).

the (rs4726576; rs10273639) risk allele, these patients are at higher risk of trypsin activation and pancreatitis.[26] In a study investigating the effect of asparaginase on pancreatic acinar cells, asparaginase evoked intracellular calcium release from the endoplasmic reticulum mediated by the protease-activated receptor 2 (Figure 3). This elevation in calcium levels in turn activated calcium release activated calcium (CRAC) channels further increasing intracellular calcium levels, leading to decreased ATP levels, trypsin activation and necrosis.[35] The pancreatitis-causing mechanism seems to be independent of the anti-neoplastic effect of asparaginase, and drugs inhibiting CRAC channels could thus be used to prevent AAP during asparaginase therapy, not least in patients who are re-exposed to asparaginase after having had AAP, since these patients have a ~50% risk or developing a second episode of AAP.[9,13]

The *NFATC2* associated variant rs62228256 was most significantly associated with AAP. However, with no association found in the replication study and no association with pancreatitis found in adult studies on non-asparaginase associated pancreatitis,[26] the association between *NFATC2* and AAP seems to have low credibility. We were not able to associate the *PRSS1-PRSS2* genotype to risk of AAP-related complications, indicating that this allele only alters the risk of AAP, while the complications are a result of other factors. The validation of our *PRSS1-PRSS2* top SNPs in the AALL0232 cohort strengthens the credibility of this result. The association was of similar effect size, but of borderline statistical significance, which may reflect three key issues: (i) diagnostic criteria differ between the cohorts, and cases are not completely comparable; (ii) pancreatitis is strongly associated with asparaginase exposure, and it was a prerequisite that included controls received a significant amount of asparaginase to reduce the risk of false negative controls; and (iii) the validation cohort was relatively small.[36,37]

Our results need to be judged in the light of their limitations. The association analyses were strongly influenced by individuals of CEU ethnicity, and we cannot determine the effects in non-CEU populations. Moreover, our results for the *PRSS1-PRSS2* locus did not reach genome-wide significance with a *P*-value $<5 \times 10^{-8}$. This highlights a challenge when doing GWA studies in cohorts of limited size, as will often be the case in childhood ALL, and requires strategies for validation in independent and similar cohorts. In this present study, we attempted to limit the problem of limited sample size by improving the quality of phenotyping, collecting individual clinical data on AAP cases and including controls with documented completion of extensive asparaginase therapy.

In conclusion, we found that children who develop AAP possess identical genetic risk variants as adults with non-asparaginase-associated pancreatitis. This may allow future preventive measures for reduction of AAP.

References

1. Schmiegelow K, Müller K, Mogensen SS, et al. Non-infectious chemotherapy-associated acute toxicities during childhood acute lymphoblastic leukemia therapy. F1000Res. 2017;6444.
2. Knott SRV, Wagenblast E, Khan S, et al. Asparagine bioavailability governs metastasis in a model of breast cancer. Nature. 2018;554(7692):378-381.
3. Müller HJ, Boos J. Use of L-asparaginase in childhood ALL. Crit Rev Oncol Hematol. 1998;28(2):97–113.
4. Pession A, Valsecchi MG, Masera G, et al. Long-term results of a randomized trial on extended use of high dose L-asparaginase for standard risk childhood acute lymphoblastic leukemia. J Clin Oncol. 2005;23(28):7161–7167.
5. Silverman LB. Improved outcome for children with acute lymphoblastic leukemia: results of Dana-Farber Consortium Protocol 91-01. Blood. 2001;97(5):1211–1218.
6. Haskell CM, Canellos GP, Leventhal BG, et al. L-Asparaginase. N Engl J Med. 1969;281 (19):1028–1034.
7. Raja RA, Schmiegelow K, Frandsen TL. Asparaginase-associated pancreatitis in children. Br J Haematol. 2012;159(August):18–27.
8. Liu C, Yang W, Devidas M, et al. Clinical and genetic risk factors for acute pancreatitis in patients with acute lymphoblastic leukemia. J Clin Oncol. 2016;34(18):2133-2140.
9. Wolthers BO, Frandsen TL, Baruchel A, et al. Asparaginase-associated pancreatitis in childhood acute lymphoblastic leukaemia: an observational Ponte di Legno Toxicity Working Group study. Lancet Oncol. 2017;18(9):1238–1248.
10. Pemmaraju N, Rytting ME. Questions on asparaginase-associated pancreatitis. Lancet Oncol. 2017;18(9):1148–1149.
11. Schmiegelow K, Attarbaschi A, Barzilai S, et al. Consensus definitions of 14 severe acute toxic effects for childhood lymphoblastic leukaemia treatment: a Delphi consensus. Lancet Oncol. 2016;17(6):e231–e239.
12. Computerome. http://www.computerome.dtu.dk/.
13. Wolthers BO, Frandsen TL, Abrahamsson J, et al. Asparaginase-associated pancreatitis. A study on pheno-and genotype in the NOPHO ALL2008 protocol. Leukemia. 2017;31(2):325-332.
14. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–575.
15. Yates A, Akanni W, Amode MR, et al. Ensembl 2016. Nucleic Acids Res. 2016;44(D1):D710–D716.
16. Sherry ST, Ward MH, Kholodov M, et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 2001;29(1):308–311.
17. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris P, Zondervan KT. Data quality control in genetic case-control association studies. Nat Protoc. 2011;5(9):1564–1573.
18. Bourgeois S, Labuda D. Dynamic allele-specific oligonucleotide hybridization on solid support. Anal Biochem. 2004;324(2):309–311.
19. Labuda D, Krajinovic M, Richer C, et al. Rapid detection of CYP1A1, CYP2D6, and NAT variants by multiplex polymerase chain reaction and allele-specific oligonucleotide assay. anal biochem. 1999;275(1):84–92.
20. Team RC. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
21. Machiela MJ, Chanock SJ. LDassoc: an online tool for interactively exploring genome-wide association study results and prioritizing variants for functional investigation. Bioinformatics. 2018;34(5):887-889.
22. Lonsdale J, Thomas J, Salvatore M, et al. The genotype-tissue expression (GTEx) project. Nat Genet. 2013;45(6):580–585.
23. Boyle AP, Hong EL, Hariharan M, et al.

Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 2012;22(9):1790–1797.

24. Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010;26(18):2336–2337.

25. Rebhan M, Chalifa-Caspi V, Prilusky J, Lancet D. GeneCards: integrating information about genes, proteins and diseases. Trends Genet. 1997;13(4):163.

26. Zator Z, Whitcomb DC. Insights into the genetic risk factors for the development of pancreatic disease. Therap Adv Gastroenterol. 2017;10(3):323-336.

27. Vrooman LM, Supko JG, Neuberg DS, et al. Erwinia asparaginase after allergy to E. coli asparaginase in children with acute lymphoblastic leukemia. Pediatr Blood Cancer. 2010;54(2):199-205.

28. Silverman LB, Stevenson KE, Brien JEO, et al. Long-term results of Dana-Farber Cancer Institute ALL Consortium protocols for children with newly diagnosed acute lymphoblastic leukemia (1985–2000). Leukemia. 2010;24(2):617–632.

29. Whitcomb DC, Gorry MC, Preston RA, et al. Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. Nat Genet. 1996;14(2):141–145.

30. Whitcomb DC, LaRusch J, Krasinskas AM, et al. Common genetic variants in the CLDN2 and PRSS1-PRSS2 loci alter risk for alcohol-related and sporadic pancreatitis. Nat Genet. 2012;44(12):1349–1354.

31. Rosendahl J, Kirsten H, Hegyi E, et al. Genome-wide association study identifies inversion in the CTRB1-CTRB2 locus to modify risk for alcoholic and non-alcoholic chronic pancreatitis. Gut. 2018;67(10):1855-1863.

32. Masamune A, Nakano E, Hamada S, Kakuta Y, Kume K, Shimosegawa T. Common variants at PRSS1–PRSS2 and CLDN2–MORC4 loci associate with chronic pancreatitis in Japan. Gut. 2015;64(8):1345–1346.

33. Paliwal S, Bhaskar S, Reddy DN, et al. Association analysis of PRSS1-PRSS2 and CLDN2-MORC4 variants in nonalcoholic chronic pancreatitis using tropical calcific pancreatitis as model. Pancreas. 2016;45(8):1153–1157.

34. Boulling A, Sato M, Masson E, Génin E, Chen J-M, Férec C. Identification of a functional PRSS1 promoter variant in linkage disequilibrium with the chronic pancreatitis-protecting rs10273639. Gut. 2015;64(11):1837–1838.

35. Peng S, Gerasimenko JV, Tsugorka T, et al. Calcium and adenosine triphosphate control of cellular pathology: asparaginase-induced pancreatitis elicited via protease-activated receptor 2. Philos Trans R Soc Lond B Biol Sci. 2016;371(1700).

36. Sud A, Kinnersley B, Houlston RS. Genome-wide association studies of cancer: current insights and future perspectives. Nat Rev Cancer. 2017;17(11):692–704.

37. Ioannidis JPA, Thomas G, Daly MJ. Validating, augmenting and refining genome-wide association signals. Nat Rev Genet. 2009;10(5):318–329.