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; Larsen, Eric; Liang, Der-Cherng; Möricke, Anja; Rasmussen, Kirsten K.; Samarasinghe, Sujith; Silverman, Lewis B.; van der Sluis, Inge M.; Stanulla, Martin; Tulstrup, Morten; Yadav, Rachita; Yang, Wenjian; Zapotocka, Ester; Gupta, Ramneek; Schmiegelow, Kjeld

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

by Benjamin O. Wolthers, Thomas L. Frandsen, Chirag J. Patel, Rachid Abaji, Andishe Attarbaschi, Shlomit Barzilai, Antonella Colombini, Gabriele Escherich, Marie Grosjean, Maja Krajinovic, Eric Larsen, Der-Cherng Liang, Anja Möricke, Kirsten K. Rasmussen, Sujith Samarasinghe, Lewis B. Silverman, Inge M. van der Sluis, Martin Stanulla, Morten Tulstrup, Rachita Yadav, Wenjian Yang, Ester Zapotocka, Ramneek Gupta, and Kjeld Schmiegelow

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

Running title: Pancreatitis in childhood ALL

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.

Affiliations

[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, Montreal, Quebec, Canada;

[4]Department of Pediatric Hematology and Oncology, St Anna Children's Hospital and Department of Pediatric and Adolescent Medicine, Medical University of Vienna, 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, Tel Aviv, 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, Montreal, Quebec, 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, Netherlands; Erasmus Medical Center, Sophia Children's Hospital, Department of Pediatric Hematology-Oncology, Rotterdam, Netherlands; Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands;

[16]Department of Pediatric Hematology and Oncology, Hannover Medical School, Hannover, 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;

[20]Institute of Clinical Medicine, University of Copenhagen, Denmark.

Corresponding author: Kjeld Schmiegelow, Department of Pediatrics and Adolescent Medicine, The University Hospital Rigshospitalet, JMC-4072, Copenhagen, Denmark. Telephone: +45 34351357; Email: Kjeld.Schmiegelow@regionh.dk.

Abstract word count: 221

Text word count: 2783

Figures: 3; tables: 2; Appendix: 21 pages

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 acute lymphoblastic leukemia trial groups contributed remission samples from patients aged 1.0–17.9 years and treated from 2000–2016. Cases were defined (n=244) by at least two of the following criteria: i) abdominal pain, ii) pancreatic enzymes $\geq$3 x upper normal limit, iii) imaging compatible with asparaginase-associated pancreatitis. Controls (n=1320) completed intended asparaginase therapy, 78% receiving ≥8 pegylated-asparaginase injections, without developing aparaginase-associated pancreatitis. rs62228256 on 20q13.2 showed the strongest association (OR=3.75; P=5.2x10^{-8}). Moreover, rs13228878 (OR=0.61; P=7.1x10^{-6}) and rs10273639 (OR=0.62; P=1.1x10^{-5}) on 7q34 showed significant association. A Dana Farber Cancer Institute ALL Consortium cohort consisting of patients treated protocols from 1987–2004 (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 (P=0.86), 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 trypsinogen encoding *PRSS1* gene and are known risk variants for alcohol-associated and sporadic pancreatitis in adults. Intra-pancreatic trypsinogen cleavage to proteolytic trypsin induces autodigestion and pancreatitis. Asparaginase-associated pancreatitis and non-asparaginase associated pancreatitis shares genetic predisposition and targeting the trypsinogen activation pathway may enable identification of effective interventions towards asparaginase-associated pancreatitis.

Introduction

Intensification of chemotherapy has generated 5-year survival rates of childhood acute lymphoblastic leukemia (ALL) above 90%, but therapy-related toxicity has similarly increased[1]. Asparaginase is a key drug in childhood ALL therapy and is gaining interest as an anti-metastatic agent in breast cancer[2]. It 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 due to impaired asparagine synthesizing ability[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 AAP, and thus AAP often entail truncation of asparaginase therapy, thereby decreasing the chance of ALL survival[4,5,9]. The pancreatitis-causing mechanism of asparaginase and identification of patients with an altered risk of AAP remains elusive[10]. To explore the AAP phenotype and investigate host-genome variants associated with risk of AAP, the Ponte di Legno toxicity working group (PTWG) initiated a three-step AAP study: I) Defining diagnostic consensus criteria for AAP[11]; II) Describing the AAP phenotype in patients across multiple ALL trial groups[9]; III) Exploring genotype-phenotype associations in 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 hypothesis and often including a speculative mechanism. Replication of GWA study results are 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 (supplemental table 1) 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 (supplemental figure 1). 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 was stored at the Technical University of Denmark's server Computerome[12].

Children (age 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 asparaginase (native *E. coli* asparaginase or polyethylene glycolated *E. coli* asparaginase (PEG-asparaginase)) injection and cases fulfilled the PTWG consensus definition for AAP: At least two of i) amylase, pancreatic amylase, or pancreatic lipase $\geq$ 3 x upper normal limit (UNL), ii) abdominal pain, iii) imaging compatible with AAP. 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 has previously been included in a Nordic Society of Pediatric Haematology and Oncology (NOPHO) GWA study[13]. These samples were genotyped on identical genotyping arrays as the remaining cohort, raw genotyping data on these patients were pooled with 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 an 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] (supplemental figures 2–4), 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; 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 $P < 1.10^{-5}$); 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 de novo genotyped in a cohort of patients from the DFCI ALL Consortium protocols 87–01, 91–01, 95–01 and 00–01 protocols (1987-2004). The AALL0232 ALL cohort included 76 cases diagnosed using National Cancer Institute's Common Terminology Criteria for Adverse Events (CTCAE criteria) (supplemental page 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 reference population as reported (ref: PMID 26265699). Time-dependent analysis (Cox proportional hazards regression) was performed adjusting for age and ancestry. The DFCI ALL Consortium protocols 87–01, 91–01, 95–01 and 00–01 cohort received 20–30 weeks of post-induction asparaginase therapy (supplemental page 9–11). Thirty-three cases diagnosed according to the CTCAE criteria and 285 controls were included in this cohort, and genotyped by allele-specific oligonucleotides (ASOs) hybridization as described elsewhere[18,19]. Pearson's correlation coefficient was used investigation association 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 (supplemental figure 3) and QQ plots showed no sign of population substructure (λ = 1.02, supplemental figure 4). Statistical analysis of pheno- and genotype associations was performed in the statistical program R version 3.3.3[20]. Linear and logistic regression analysis were used for association between genotypes and continuous or categorical clinical variables adjusting for age and ancestry. And 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 (pEFS) according to genotypes, and differences were compared with the 2-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 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 association to genes and pathways previously associated 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 (eQTLs)) by GTeX[22], regulatory effect by RegulomeDB[23], and regional association plots were produced by the LocusZoom tool[24]. Genes and SNPs previously associated to pancreatitis were investigated by searching PubMed for reports published in English within the last ten 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 244 (84%) cases and 1185/1320 (90%) of controls were of European

(CEU) ancestry (supplemental figure 4). Median age was 8.1 [interquartile range {IQR} 4.3 to 13.1] and 5.0

[IQR: 3.0 to 9.0] years in cases and controls, respectively. Fifty-five percent of both cases (133/244) and

controls (724/1320) were of male gender.

Figure 1 shows the significance of SNPs associated to AAP. The variant rs62228256 (reference allele=C,

minor allele=T {C>T}) on 20q13.2 showed the strongest association to AAP (odds ratio {OR}, 3.75; 95% CI,

2.33 to 6.04; P=5.2 x 10^{-8}). Rs62228256 is located 274 kilobase pairs upstream from Nuclear factor of

activated T cells (*NFATC2)* and has been documented to be an eQTL for this gene in pancreatic tissue

(supplemental figure 5).

Among the 30 SNPs most associated to AAP with a p-value of 5 x 10^{-5} or lower were 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}) (supplemental figures 6–8 and

supplemental table 2) remained strongly associated with AAP. Further investigation of previously validated

SNPs within genes known to regulate trypsin activation[26] (supplemental table 3) found AAP 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 direction and

effect very similar to what has previously been 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).

Validation of results

In validating our top SNP in the *NFATC2* locus (rs62228256) and top SNPs 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] (supplemental page 6–12). Whereas the association between AAP and rs6228256 genotype was not replicated in DFCI ALL Consortium cohort (*P*=0.77, supplemental 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 risk of developing pancreatitis in AALL0232 cohort (supplemental page 8).

Two previous studies have investigated SNPs associated to pancreatitis in children with ALL.

Using different diagnostic criteria for pancreatitis and including controls with less than five 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 were in LD with rs199695765. Of 32 SNPs 10kb up- and downstream from *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 associated the *ULK2* variant rs281366 (*P*=5.8 x 10^{-7}) and *RGS6* variant rs17179470 (*P*=1.3 x 10^{-6}) to be most associated to 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 to pancreatitis, and risk of a second 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 to *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 (supplemental figure 9; *P*=0.4). Out of 46 children who were re-exposed to asparaginase, 17 (37%) developed a second AAP, although not statistically significant, 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 to the risk of AAP in children with ALL. In doing so, we find 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 pancreatitis pathogenesis was long suspected, but 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 risk of alcohol-related and sporadic pancreatitis[30]. This association has recently been validated in larger European and Asian cohorts, and the haplotype has been studied in detail[31–33]. Rs10273639 is located 408 base pairs upstream from the translation initiation codon of cationic trypsinogen. A recent functional study has documented the proximal rs4726576 (C>A) variant (204 kb upstream from the translation initiation codon) to be driving the association[34]. The rs4726576 and the 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 pinpointing the association driving the signal is thus of importance in the latter population. Sequencing the PRSS1-PRSS2 risk allele has not found disease-associated coding variants accounting for the pancreatitis association 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 has been shown to lead to trypsin activation, and with higher expression of trypsin as seen in patients with the [rs4726576; rs10273639] risk alleles, 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 (ER) 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 independent of the anti-neoplastic effect of asparaginase, and drugs inhibiting CRAC channels could thus be used for AAP prevention during asparaginase therapy, not least in patients who after having had AAP are re-exposed to asparaginase, since these patients have a ~50% risk or developing a second episode of AAP[9,13].

Most significantly associated to AAP was the NFATC2 associated variant rs62228256. However, with no association found in the replication study and no association to 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 results. 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 to asparaginase exposure, and it is a prerequisite that

included controls received a significant amount of asparaginase to reduce the risk of false negative controls. III) The validation cohorts were relatively small[36,37].

Our results however, need to be judged in light of their limitations. The association analyses were strongly influenced by individuals with CEU ethnicity, and we cannot determine the effects in non-CEU populations. Moreover, our results in the *PRSS1-PRSS2* locus did not reach genome-wide significance with a p-value <5 x 10^{-8}. This highlights a challenge when doing GWAS 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 sample size challenge 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 find 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.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

We declare no competing interests.

Author contributions

Conception and design: All authors.

Provision of study materials and assembly of data: Benjamin O. Wolthers, Thomas L. Frandsen, Rachid Abaji, Andishe Attarbaschi, Shlomit Barzilai, Antonella Colombini, Gabriele Escherich, Maja Krajinovic, Eric Larsen, Der-Cherng Liang, Anja Möricke, Kirsten K. Rasmussen, Sujith Samarasinghe, Lewis B. Silverman, Inge M. van der Sluis, Martin Stanulla, Wenjian Yang, Ester Zapotocka and Kjeld Schmiegelow

Data analysis and interpretation: Benjamin O. Wolthers, Thomas L. Frandsen, Rachid Abaji, Chirag J. Patel, Marie Grosjean, Morten Tulstrup, Rachita Yadav, Wenjian Yang, Ramneek Gupta and Kjeld Schmiegelow

Manuscript writing: Benjamin O. Wolthers, Thomas L. Frandsen and Kjeld Schmiegelow

Final approval of manuscript: All authors

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.

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)	5.18×10^{-8}	
rs7270119	20	50436587	A>G	0.07	0.02	3.64 (2.28-5.8)	5.52×10^{-8}	
rs16996276	20	50455925	A>C	0.07	0.02	3.64 (2.27-5.85)	8.64×10^{-8}	
rs62228230	20	50445082	G>A	0.07	0.02	3.54 (2.23-5.63)	9.19×10^{-8}	
rs934350	13	103589776	A>G	0.32	0.22	1.84 (1.47-2.3)	1.16×10^{-7}	
rs170623	9	101984936	C>G	0.38	0.28	1.78 (1.42-2.22)	3.93×10^{-7}	*ALG2 (+0.69kb) & SEC61B (0)*
rs75245362	14	95990645	C>T	0.11	0.05	2.46 (1.73-3.51)	6.33×10^{-7}	*SCARNA13 (-9.05kb) & SNHG10 (-8.6kb)*
rs368819120	12	71747240	G>-	0.48	0.37	1.67 (1.35-2.06)	2.04×10^{-6}	
rs4769201	13	22698015	A>G	0.04	0.01	4.66 (2.45-8.86)	2.67×10^{-6}	
rs7851954	9	6796167	C>T	0.28	0.40	0.59 (0.47-0.74)	3.02×10^{-6}	*KDM4C (0)*
rs61734424	19	50747533	T>C	0.08	0.04	2.75 (1.8-4.2)	3.04×10^{-6}	*MYH14 (0)*
rs9912225	17	4680732	A>G	0.10	0.05	2.42 (1.67-3.5)	3.10×10^{-6}	*TM4SF5 (0) & VMO1 (-7.8kb)*
rs7155612	14	95976755	T>G	0.12	0.06	2.24 (1.6-3.15)	3.15×10^{-6}	
rs2167730	8	78103417	T>C	0.24	0.34	0.57 (0.45-0.73)	3.69×10^{-6}	
rs80170196	19	50747159	C>T	0.08	0.04	2.73 (1.78-4.18)	3.77×10^{-6}	*MYH14 (0)*
rs62228228	20	50443845	G>A	0.06	0.02	3.15 (1.94-5.13)	3.80×10^{-6}	
rs5010616	12	71748290	C>T	0.48	0.37	1.63 (1.32-2)	4.76×10^{-6}	
rs12494164	3	164967758	A>C	0.23	0.14	1.77 (1.39-2.27)	4.83×10^{-6}	
rs16848986	3	164979570	T>C	0.22	0.14	1.77 (1.38-2.26)	5.19×10^{-6}	
rs34375180	12	71779640	G>A	0.49	0.38	1.61 (1.31-1.98)	5.53×10^{-6}	
rs7139808	13	22693228	T>C	0.04	0.01	4.59 (2.38-8.87)	5.59×10^{-6}	
rs12582343	12	71766297	A>G	0.48	0.38	1.62 (1.31-1.99)	5.90×10^{-6}	
rs13228878	7	142473466	A>G	0.35	0.44	0.61 (0.5-0.76)	7.06×10^{-6}	*PRSS2 (-6.44kb) & PRSS3P2 (-5.29kb)*
rs6477109	9	6794938	C>T	0.29	0.40	0.61 (0.49-0.75)	7.64×10^{-6}	*KDM4C (0)*
rs74109922	13	103582300	A>G	0.12	0.06	2.13 (1.52-2.97)	1.03×10^{-5}	
rs1505495	4	172973580	A>C	0.10	0.17	0.48 (0.34-0.66)	1.06×10^{-5}	*GALNTL6 (0)*
rs1791520	18	22118315	T>C	0.32	0.23	1.65 (1.32-2.07)	1.12×10^{-5}	
rs10273639	7	142456928	C>T	0.35	0.44	0.62 (0.5-0.77)	1.13×10^{-5}	*PRSS1 (-0.4kb)*
rs4655107	1	23094454	G>A	0.13	0.24	0.53 (0.39-0.7)	1.16×10^{-5}	*EPHB2 (0)*
rs55634345	4	19846813	G>A	0.46	0.34	1.58 (1.29-1.94)	1.19×10^{-5}	

Legend

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

Table 2 Baseline characteristics, pancreatic enzyme levels and complications to 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 (median; IQR)	7.75 (4–12.06)	9.21 (4.46–13.43)	6.22 (4.9–12.1)	0.27
Sex (% males)	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)	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)	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

Legend

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

Figure 1

Manhattanplot

Legend

Manhattanplot showing 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's association calculated using logistic regression adjusting for age and ancestry. Genes previously associated to pancreatitis are marked in color. SNPs are annotated to genes based on genomic location (10 kb up- and downstream from transcription start site and transcription terminator, respectively. The human assembly *GRCh37* was used for reference.

Figure 2

Regional association plot of the *PRSS1-2* locus on chromosome 7

Legend

Regional asssociation plot showing 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's association calculated using logistic regression adjusting for age and ancestry. Rs13228878 ($P=7.1$ x 10^{-6}) is marked in purple and rs10273639 ($P=1.1$ x 10^{-5}) in red. The color of the dots reflect 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.

Figure 3

Schematic diagram illustrating the effects of asparaginase on pancreatic acinar cells

Legend

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 Ca^{++} release activated (CRAC) channels, further increasing intracellular calcium levels, reducing ATP levels and allowing activation of inactive trypsinogen to active trypsin. The diagram is heavily influenced by diagrams from Peng *et al*. Phil Trans Royal Soc, 2015 and Whitcomb *et al*. Nature Gen, 2012.
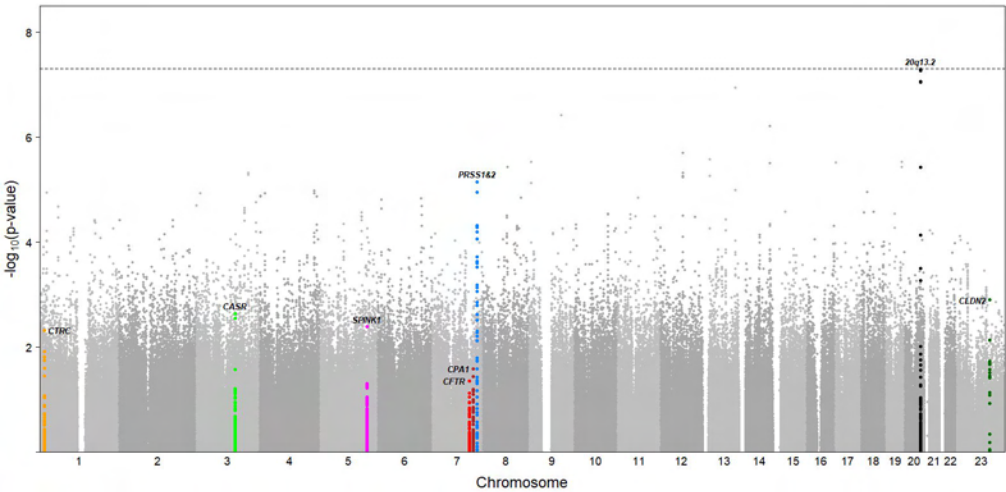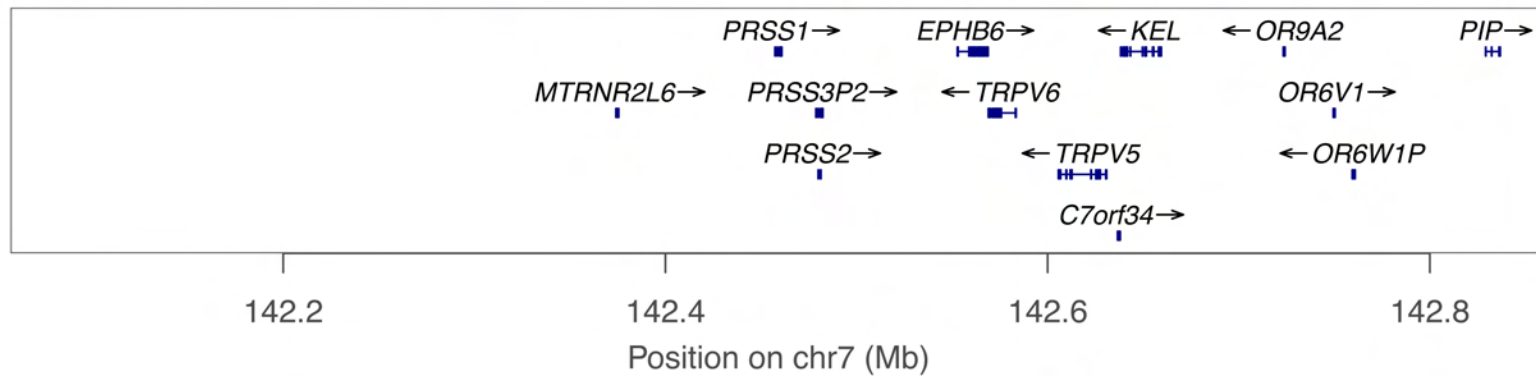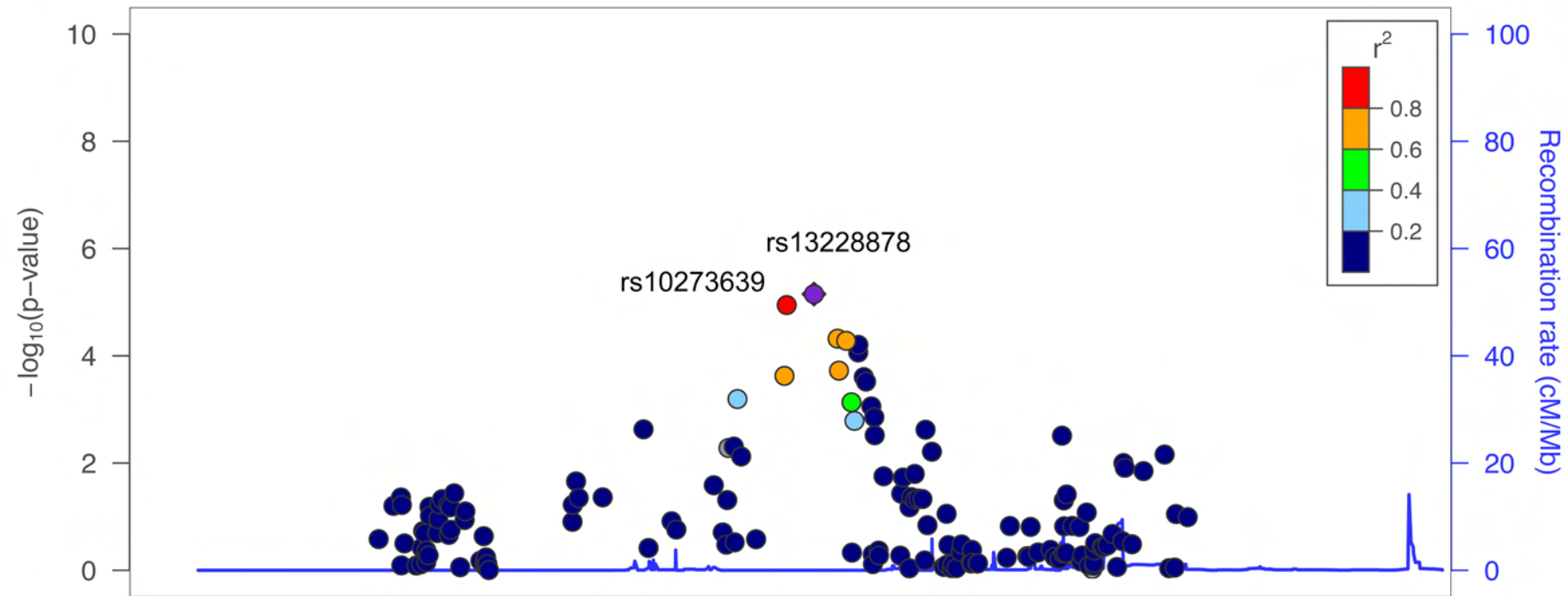
Pancreas

Exocrine pancreas

Acinar cell

Acinar cell

Zymogens including trypsinogens (*PRSS1-2*), chymotrypsin (*CTRC*) & serine peptidase inhibitor kazal type 1 (*SPINK1*)

Asparaginase

Protease-activated receptor 2

Ca^{++} release activated Ca^{++} channels

$\uparrow Ca^{++}$

Endoplasmic reticulum

$\blacktriangle Ca^{++}$

Trypsinogen $\rightarrow$ Trypsin
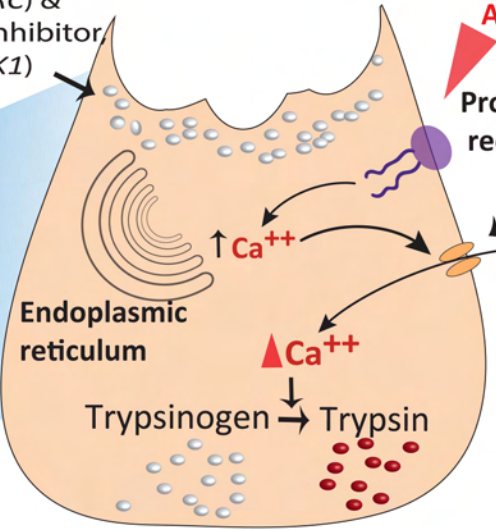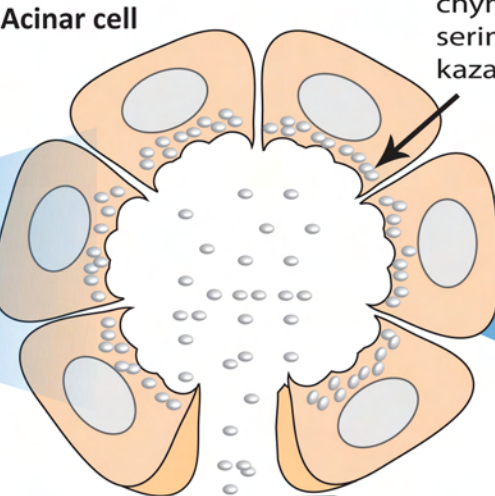
Duct cell

Cystic fibrosis transmembrane conductance regulator (*CFTR*)

Duodenum

Appendix

Supplemental table 1
Baseline data, acute complications and risk of re-exposure in 10 participating trial groups.

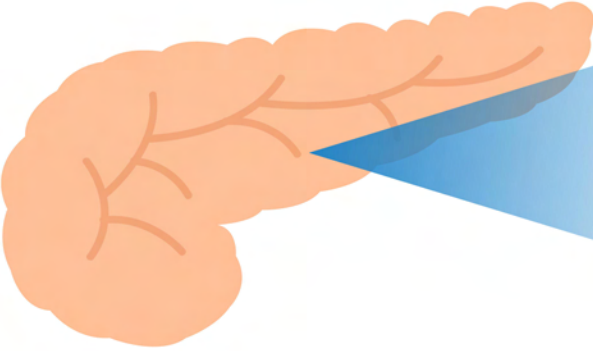| Trial group name	Trials	N	Male sex	Age (years)	B-cell precursor ALL	WBC at ALL diagnosis (10^9/L)	Time from ALL diagnosis to AAP (days)	Assisted ventilation (yes [n]/available data [N])	Acute insulin therapy (yes [n]/available data [N])	Pseudocysts (yes [n]/available data [N])	Second AAP if re-exposed (yes [n]/re-exposed patients [N])	Death due to AAP (n)
Associazione Italiana Ematologia Oncologia Pediatrica	AIEOP-BFM ALL 2009	18	61% (11/18)	9.8 (4.3–13.1)	94% (16/17)	10.7 (4.3–28) (n=17)	117 (43–215) (n=16)	13% (2/15)	13% (2/15)	21% (3/14)	0% (0/1)	2
Berlin-Frankfurt-Münster Austria	AIEOP-BFM ALL 2009	7	86% (6/7)	15.1 (10.4–16.5)	71% (5/7)	11.6 (6.5–15) (n=7)	55 (45–123) (n=7)	0% (0/6)	0% (0/7)	14% (1/7)	33% (1/3)	0
Berlin-Frankfurt-Münster	ALL-BFM 2000 AIEOP-BFM ALL 2009	49	55% (27/49)	9.7 (5.4–14.7)	76% (37/49)	12.1 (3.4–23.8) (n=49)	37 (30–161) (n=49)	8% (4/49)	17% (8/48)	20% (9/44)	25% (2/8)	1
The Cooperative ALL Study Group	CoALL 97/08-09	4	50% (2/4)	10.1 (8–12)	100% (4/4)	4.1 (3.1–5.7) (n=4)	231 (159–273) (n=3)	0% (0/1)	0% (0/1)	0% (0/1)	0% (0/2)	0
Czech Republic	ALL-BFM 2000 ALL-IC-BFM 2002 AIEOP-BFM ALL 2009	4	0% (0/4)	7.1 (2.5–12.3)	100% (4/4)	25.5 (4.8–98.8) (n=4)	149.5 (72–228) (n=4)	0% (0/4)	0% (0/4)	25% (1/4)	0% (0/1)	0
Dutch Childhood Oncology Group	DCOG-ALL10/11	25	44% (11/25)	8.2 (5.2–13.6)	80% (20/25)	8.1 (3.8–31.2) (n=24)	203 (46–295) (n=25)	12% (3/25)	30% (7/23)	28% (7/25)	50% (2/4)	1
Israel	INS 2010 AIEOP-BFM ALL 2009	10	70% (7/10)	11.2 (8.1–15.3)	56% (5/9)	41 (6.3–104) (n=9)	38 (31–113) (n=9)	0% (0/9)	22% (2/9)	33% (3/9)	0% (0/1)	0
Nordic Society of Pediatric Haematology and Oncology	NOPHO ALL 2008	92	55% (51/92)	6.3 (3.4–10.9)	86% (65/76)	20.6 (6.5–50) (n=76)	110 (75–144) (n=76)	4% (3/76)	22% (17/76)	28% (21/76)	53% (10/19)	0
Taiwan Pediatric Oncology Group	TPOG-ALL-97/2002/2013	5	60% (3/5)	6.9 (6.6–8.2)	100% (5/5)	26 (1.2–79.8) (n=5)	34 (31–150) (n=5)	0% (0/5)	50% (2/4)	40% (2/5)	50% (1/2)	0
United Kingdom ALL Working Party	UKALL2003	30	50% (15/30)	9.4 (4.7–13.3)	83% (24/29)	7.3 (3–41.1) (n=30)	146 (82–225) (n=30)	0% (0/30)	100% (3/3)	28% (8/29)	20% (1/5)	0
10 Groups	**14 Trials**	**244**	**55% (133/244)**	**8.1 years (4.3—13.1)**	**82% (185/225)**	**15.3 (4–41) (n=225)**	**99.5 (41–182) (n=224)**	**5% (12/220)**	**22% (41/190)**	**26% (55/214)**	**37% (17/46)**	**1·8% (4/224)**

Legend

Baseline data, acute complications and risk of re-exposure in 10 participating trial groups. Data are n, median (IQR), %(n/N), unless indicated otherwise.

*Number of re-exposed patients with data on second AAP.

Note: BCP refers to percentage with B-cell precursor acute lymphoblastic leukemia, the remaining group had T-cell leukemia.

Abbreviations: AAP, asparaginase-associated pancreatitis; IQR, interquartile range (25th/75th centiles); BCP ALL, B-cell precursor acute lymphoblastic leukemia; WBC, white blood cell count.

1

Supplemental figure 1

Pie charts showing the distribtution of cases and controls according to trialgroup
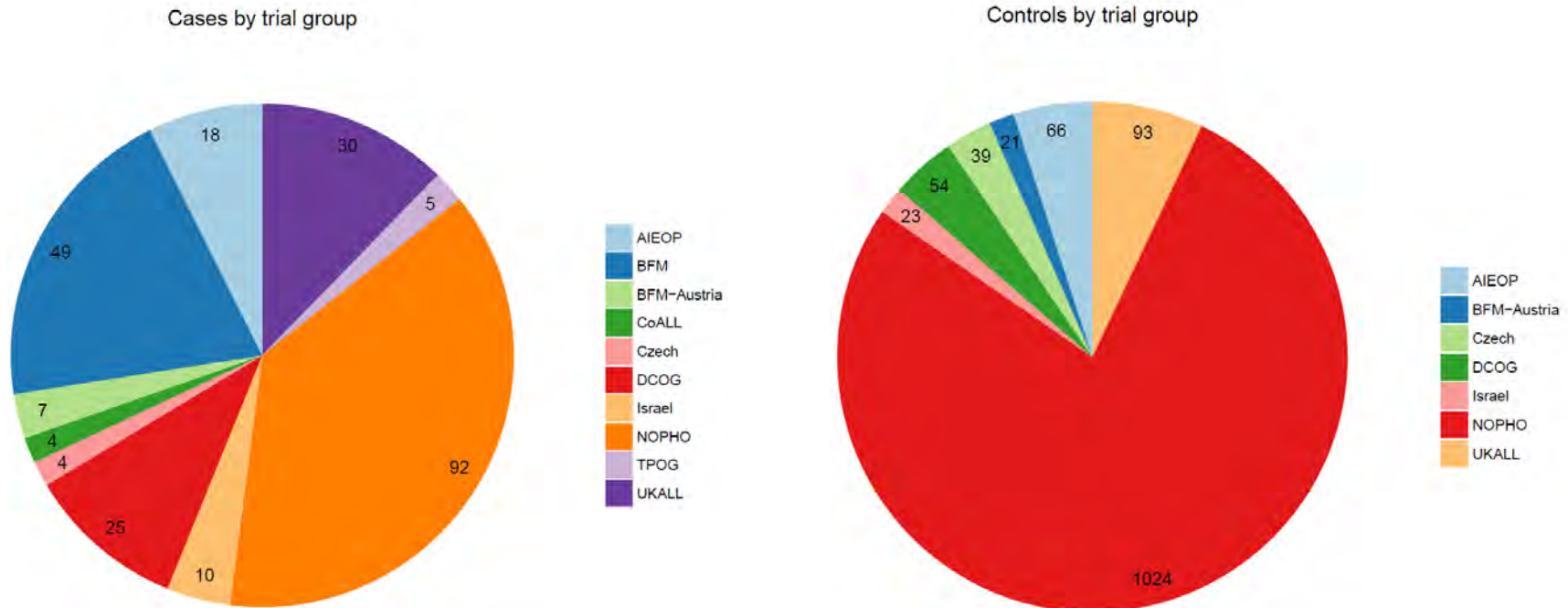
Legend

Pie charts showing the distribution of 244 AAP cases and 1320 controls according to childhood acute lymphoblastic leukemia trial group.

Abbreviations: AIEOP, Associazione Italiana Ematologia Oncologia Pediatrica; BFM, Berlin-Frankfurt-Münster (Germany) ; BFM-Austria, Berlin-Frankfurt-Münster Austria; CoALL, The Cooperative ALL Study Group (Germany); Czech, Czech Working Group for Paediatric Haematology; DCOG, Dutch Childhood Oncology Group; DFCI, Dana Farber Cancer Institute (US); Israel, Israeli childhood cancer group; NOPHO, Nordic Society of Pediatric Haematology and Oncology (Denmark, Estonia, Finland, Iceland, Lithuania, Norway, Sweden); TOPG, Taiwan Pediatric Oncology Group; UKALL, United Kingdom ALL Working Party.

Supplemental figure 2

Quality control flowchart

Assessed for eligibility
2631699 variants
1895 individuals (311 cases 1584 controls)

Discordant sex in geno and phenotypes
(n_{total} =113; n_{cases} = 34; $n_{controls}$ =79)
Individuals missing >10% of variants
(n_{total} =81; n_{cases} =15; $n_{controls}$=66)
Excess heterozygosity
(n_{total} =20; n_{cases} = 5; $n_{controls}$ =15)
Relatedness between individuals
(n_{total} =97; n_{cases}= 7; $n_{controls}$ =90)

2631699 variants
1584 individuals (250 cases 1334 controls)

SNPs missing in >2% of individuals
($n_{variants}$ =302644)
Not in HW equlibrium
($n_{variants}$ =1909)
Minor allele frequency
($n_{variants}$ =898665)
Difference in call rate between cases and controls
($n_{variants}$ =6)
Duplicated variants
($n_{variants}$=26567)

- **Individuals missing >3% of variants**
(n_{total}=20; n_{cases}=6; $n_{controls}$=14)

Study cohort
1401908 variants
1564 individuals (244 cases 1320 controls)

Legend

Flow diagram illustrating the quality control of individual and single nucleotide polymorphism data.

Supplemental figure 3

Multidimensional scaling plots of genetic ancestry in study population

Legend
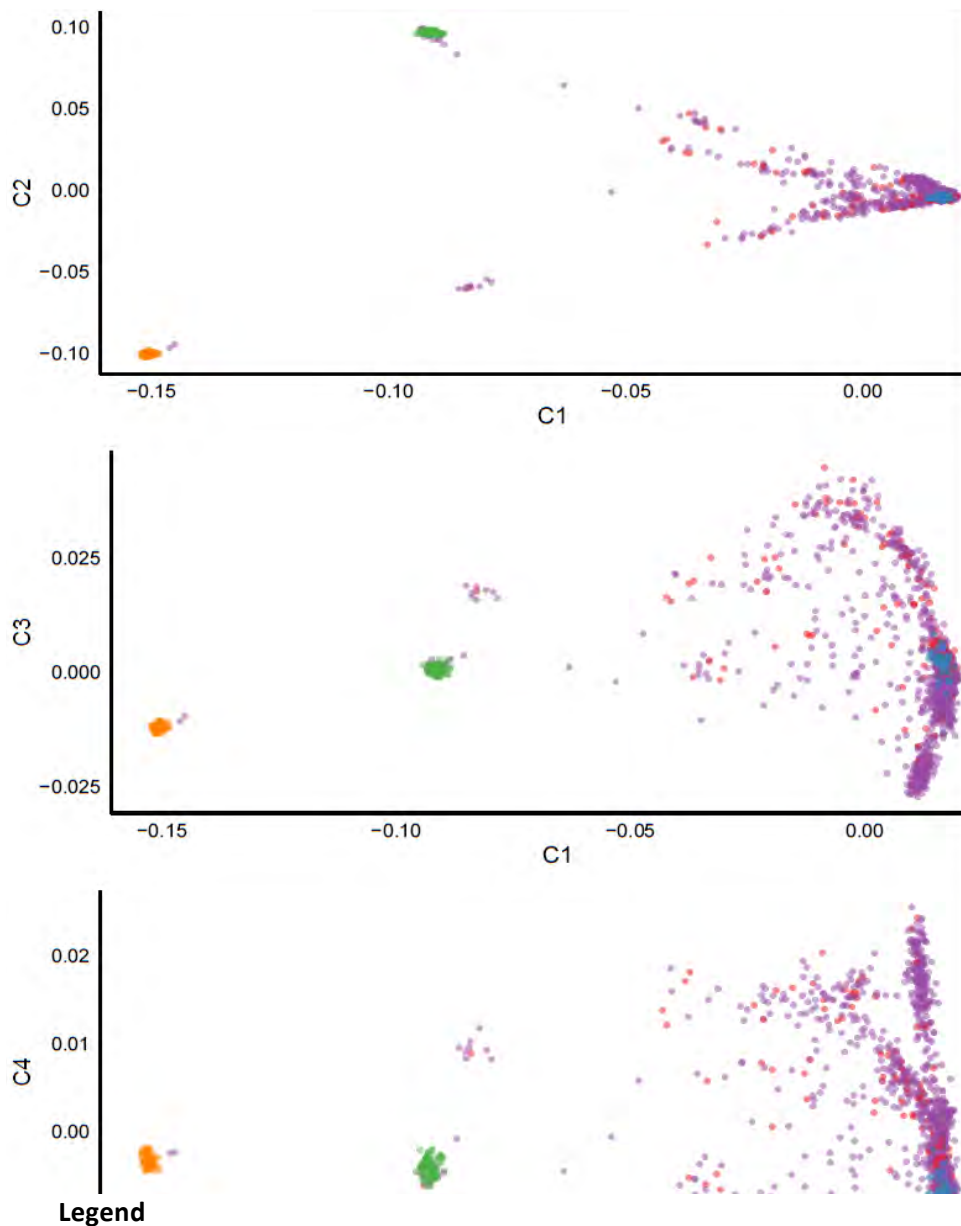
Multidimensional scaling plots plot of genetic ancestry of AAP cases, controls and HAPMAP controls for reference. Abbreviations: CEU, Utah residents with Northern and Western European ancestry; CHB, Han Chinese in Beijing, China; JPT, Japanese in Tokyo, Japan; YRI, Yoruba in Ibadan, Nigeria.

Supplemental figure 4

Quartile-quartile plot
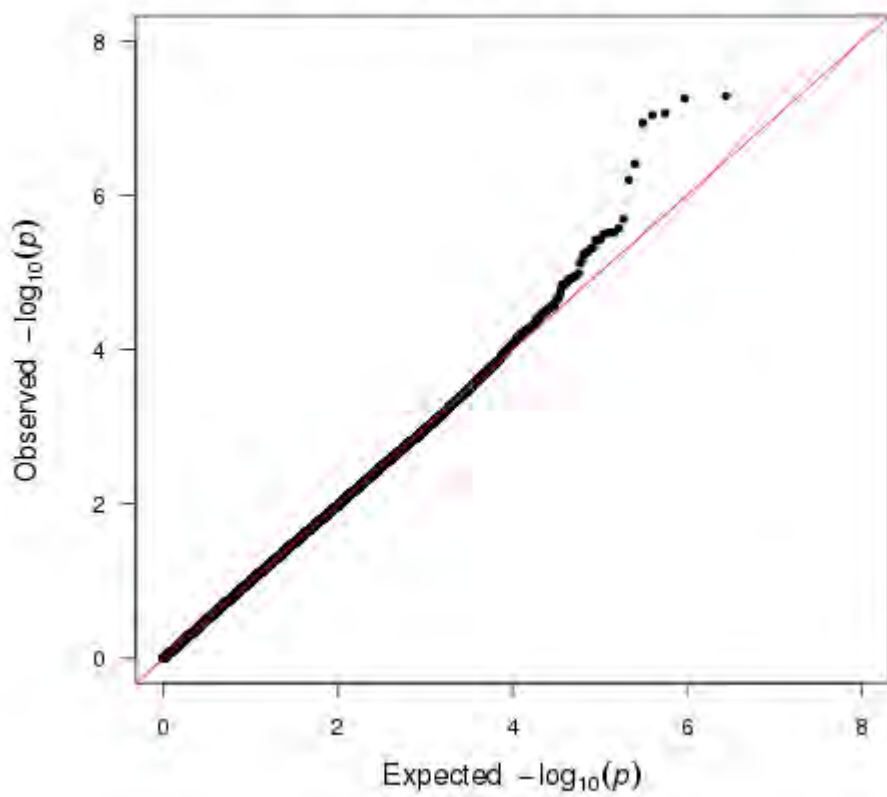
Legend

Quartile-quartile plot showing the expected *vs* the observed p-values in the cohort of 244 cases and 1320 controls. The logistic regression analysis was adjusted for age and genetic ancestry, with a lambda of 1.02 there was no evidence of population substructure.

Description of validation cohort

Replication of results in the Children's Oncology Group AALL0232 cohort

All data is provided from Liu *et al. "Clinical and Genetic Risk Factors for Acute Pancreatitis in Patients With Acute Lymphoblastic Leukemia*"; Journal of Clinical Oncology 2016**.**

The AALL0232 trial was used for validation. This cohort constitutes the largest cohort, with significant duration of asparaginase treatment (~16 weeks) previously published (Liu et al. Journal of Clinical Oncology, 2016). The trial included 3058 patients, out of which genotyping data was available in 76 cases diagnosed using National Cancer Institute's Common Terminology Criteria for Adverse Events, and 2577 controls. The cohort was included, in the only other GWAS published on pancreatitis in children with ALL[1]. The AALL0232 cohort was genotyped on the Affymetrix Genome-Wide Human SNP 6.0 (or GeneChip Human Mapping 500K) Array, and linkage disequilibrium (LD) between SNPs in the investigation and validation study was assessed by the National Cancer Institute LDassoc tool.

Diagnostic criteria for pancreatitis

National Cancer Institute's Common Terminology Criteria for Adverse Events for Pancreatitis (https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/ctcaev3.pdf) grade 2-4 were included as cases in the analysis:

Adverse Event	Short Name	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
Pancreatitis	Pancreatitis	Asymptomatic, enzyme elevation and/or radiographic findings	Symptomatic, medical intervention indicated	Interventional radiology or operative intervention indicated	Life-threatening consequences (e.g., circulatory failure, hemorrhage, sepsis)	Death

Asparaginase exposure in protocol AALL0232

Induction	Extended Induction[a]	Consolidation	Interim Maintenance and Delayed Intensification	Total ASP dose[b] U/m^2 (excluding Extended induction)	Total ASP weeks[c] (excluding Extended induction)	N (%) of patients developing pancreatitis
PEG 2500 U/m^2 x1	PEG 2500 U/m^2 x1	PEG 2500 U/m^2 × 2	DH/PH: PEG 2500 U/m^2 × 2 (single IM/DI) or PEG 2500 U/m^2 × 6 (double IM/DI)	250.000–4500000	10–18	32/1359 (2.4%)
			DC/PC: PEG 2500 U/m^2 × 4 (single IM/DI) or × 6 (double IM/DI)	350.000–550,000	14–22	44/1294 (3.4%)

All asparaginase was given intramuscularly. [a]Depends on patients' bone marrow blast percent during induction or minimal residual disease (MRD) status at the end of Induction. [b]Total dose and [c]total weeks of *PEG*-asparaginase during therapy, excluded the extended Induction. Abbreviations: ASP, asparaginase; PEG, pegylated; DH/PH, dexamethasone/prednisone during induction therapy and high-dose methotrexate/leucovorin during the first interim maintenance; DC/PC, dexamethasone/prednisone during induction therapy and escalating methotrexate/leucovorin during the first interim maintenance.

Baseline data in 2653 patients included in the AALL0232 genome-wide association study cohort

	ALL0232 (n=2653)
Age	
1–10 years old	907 (34.2%)
10–30 years old	1746 (65.8%)
Gender	
Male	1468 (55.3%)
Female	1185 (44.7%)
Race	
White	1455 (54.8%)
Black	134 (5.1%)
Hispanic	710 (26.8%)
Asian	60 (2.3%)
Other	294 (11.1%)
Immunophenotype	
B-lineage	2653 (100%)
T-lineage	0 (0%)

Association analysis

Time-dependent analysis (Cox proportional hazards regression) was performed adjusting for age and ancestry.

AALL0232 (high asparaginase dose) All patients. N=2653, 76 cases			
	Major>minor allele	HR (95% CI)	*P*
rs13228878	A>G	0.68 (0.48–0.96)	0.03
rs10273639	C>T	0.69 (0.49–0.98)	0.04

Legend

The table shows results from the replication study.

Genotypes

rs13228878			
	CC	CT	TT
Controls	585	1227	765
Cases with pancreatitis	11	35	30
% pan	0·02	0·03	0·04

Legend

Genotype of rs13228878. The table shows the rs13228878 genotype in cases with pancreatitis and controls without pancreatitis.

Genotypes

rs10273639			
	CC	CT	TT
Controls	561	1253	763
Cases with pancreatitis	11	35	30
% pan	0·02	0·03	0·04

Legend

Genotype of rs10273639. The table shows the rs10273639 genotype in cases with pancreatitis and controls without pancreatitis.

Replication of results in the Dana Farber Cancer Institute consortium 87−01, 91−01, 95−01 and 00−01 ALL cohort

A cohort of ALL children recruited from the Dana Farber Cancer Institute (DFCI) were used for the validation of the rs62228256 polymorphism on 20q13.2. Patients were treated according to DFCI ALL Consortium protocols DFCI 87-01, 91-01, 95-01, or 00-01 between January, 1987 and July, 2005 receiving variable asparaginase exposure[2–4]. This cohort is further referred to as DFCI. Genotyping was performed on a total of 318 patients, out of which 33 were cases of pancreatitis diagnosed according to the CTCAE criteria and 285 were controls. Cases were identified by retrospective evaluation of medical charts. This cohort was genotyped by allele-specific oligonucleotide hybridization as described elsewhere.[5] Briefly, while in DFCI 95-01 and DFCI 00-01, one dose of asparaginase was administered during remission induction, asparaginase was administered for 20 to 30 consecutive weeks during consolidation phase in all protocols.

Asparaginase exposure in Dana Farber Cancer Institute protocols

Asparaginase according to protocol	
Induction (4 weeks)	
Protocol 87-01	*E. coli*, *Erwinia* or PEG ASP × 1 dose (randomized; investigational window; 5 days pre-day 0)
Protocol 91-01	None
Protocol 95-01	*E. coli* or *Erwinia* ASP 25 000 IU/m^2 × 1 dose (randomized; day 4)
Protocol 00-01	E. coli ASP 25,000 IU/m2 IM × 1 dose
Intensification (20–30 weeks) every 3-week cycle	
Protocol 87-01	*E. coli* ASP 25 000 IU/m^2 weekly
Protocol 91-01	Randomized to *E. coli* ASP 25 000 IU/m^2 weekly or PEG ASP 2500 IU/m^2 every 2 weeks
Protocol 95-01	Randomized to *E. coli* ASP 25 000 IU/m^2 weekly or *Erwinia* ASP 25 000 IU/m^2 weekly
Protocol 00-01	Randomized to fixed dosing of *E. coli* ASP (based upon BSA) and individualized dosing (based upon NSAA every 3 weeks)

Abbreviations: ASP, asparaginase; PEG, pegylated; BSA, body surface area; NSAA, nadir serum asparaginase activity.

Baseline data in 318 patients included in the Dana Farber Cancer Institute cohort

	Dana Farber Cancer Institute Cohort (n=318)
Age	
1–10 years old	254 (79,9%)
10–18 years old	64 (20,1%)
Gender	
Male	167 (52,5%)
Female	151 (47,5%)
Immunophenotype	
B-lineage	294 (92.5%)
T-lineage	24 (7.5%)
Source of ASP	
Native *E. Coli*	289 (90,9%)
Erwinia	29 (9,1%)
DFCI protocol	
00-01	125 (39,3%)
95-01	122 (38,4%)
91-01	55 (17,3%)
87-01	16 (5%)

Association analysis

Binary logistic regression analysis was performed adjusting for age.

DFCI ALL cohort. N=318, 33 cases			
	Major>minor allele	OR (95% CI)	P
rs62228256	C>T	1.19 (0.35–4.03)	0.77

Legend

The table shows results from the replication study of rs62228256.

Genotypes

	rs62228256		
	CC	CT	TT
Controls	264	20	1
Cases with pancreatitis	30	3	0
% pan	0·10	0·13	0

Legend

Genotype of rs62228256. The table shows the rs62228256 genotype in cases with pancreatitis and controls without pancreatitis.

Regional association plot of the *PRSS1* and *PRSS2* loci on chromosome 7 in AALL0232
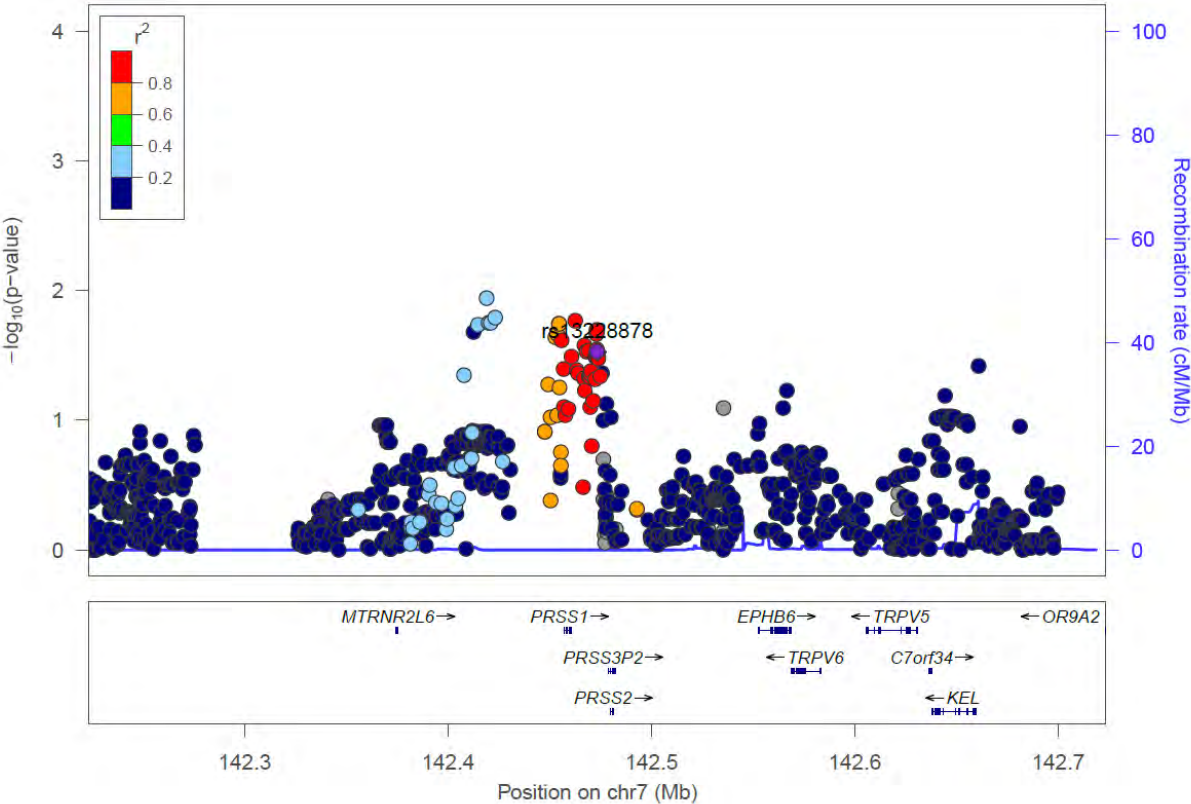
Legend

Regional asssociation plot showing SNPs associated with asparaginase-associated pancreatitis in 76 cases and 2570 controls. The *x* axis represents genomic location, and the *y* axis represents the *P* value for the SNP's association calculated using logistic regression adjusting for age and ancestry. The color of the dots reflects the linkage disequilibrium of the genotyped SNPs and rs13228878. LD is based on 1000 genomes European CEU samples, November 2014. The human assembly *GRCh37* was used for reference.

Supplemental figure 5

Rs62228256 and Nuclear factor of activated T cells (*NFATC2)*

Pancreas eQTL 20_50454447_C_T_b37 ENSG00000101096.15
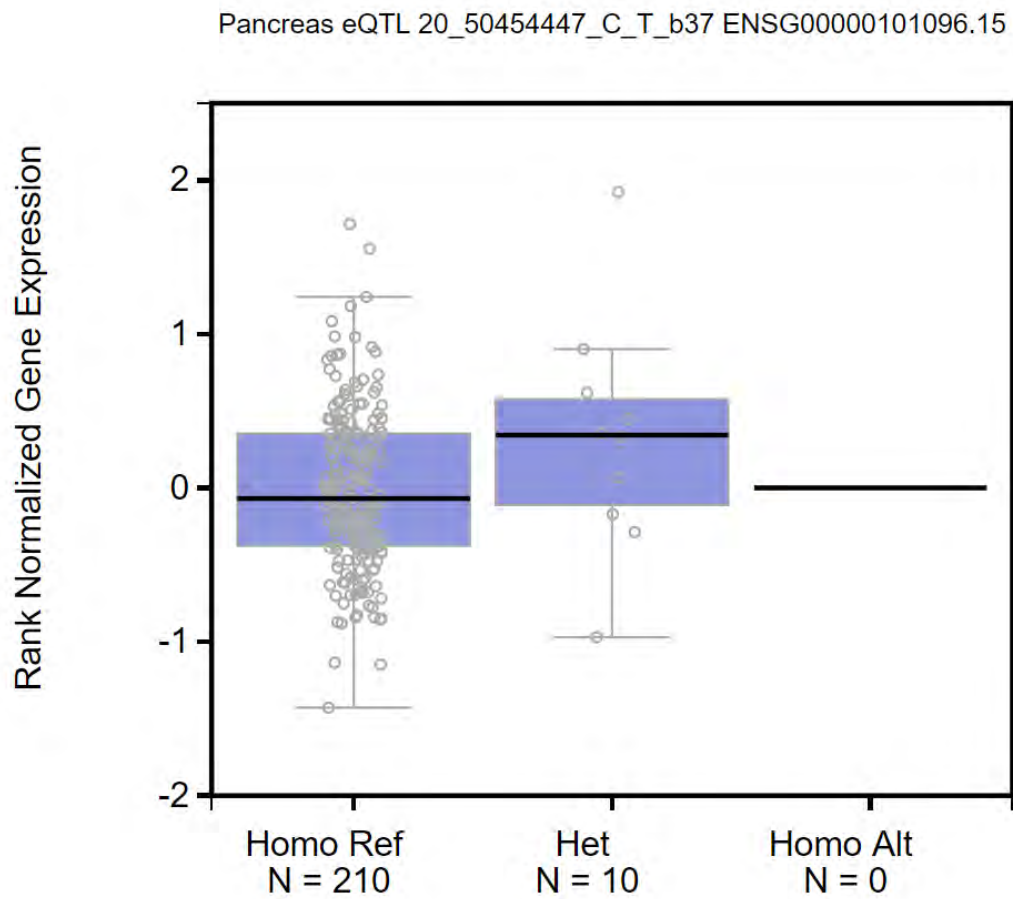
Legend

NFATC2 gene expression according to rs62228256 genotype. Homo Ref refers to the major C allele, Homo Alt refers to the minor T allele. The difference between groups is statistically significant, *P*=0.045. Plot is downloaded from www.gtexportal.org/home/.

Supplemental figure 6

Multidimensional scaling plots of genetic ancestry from CEU cohort
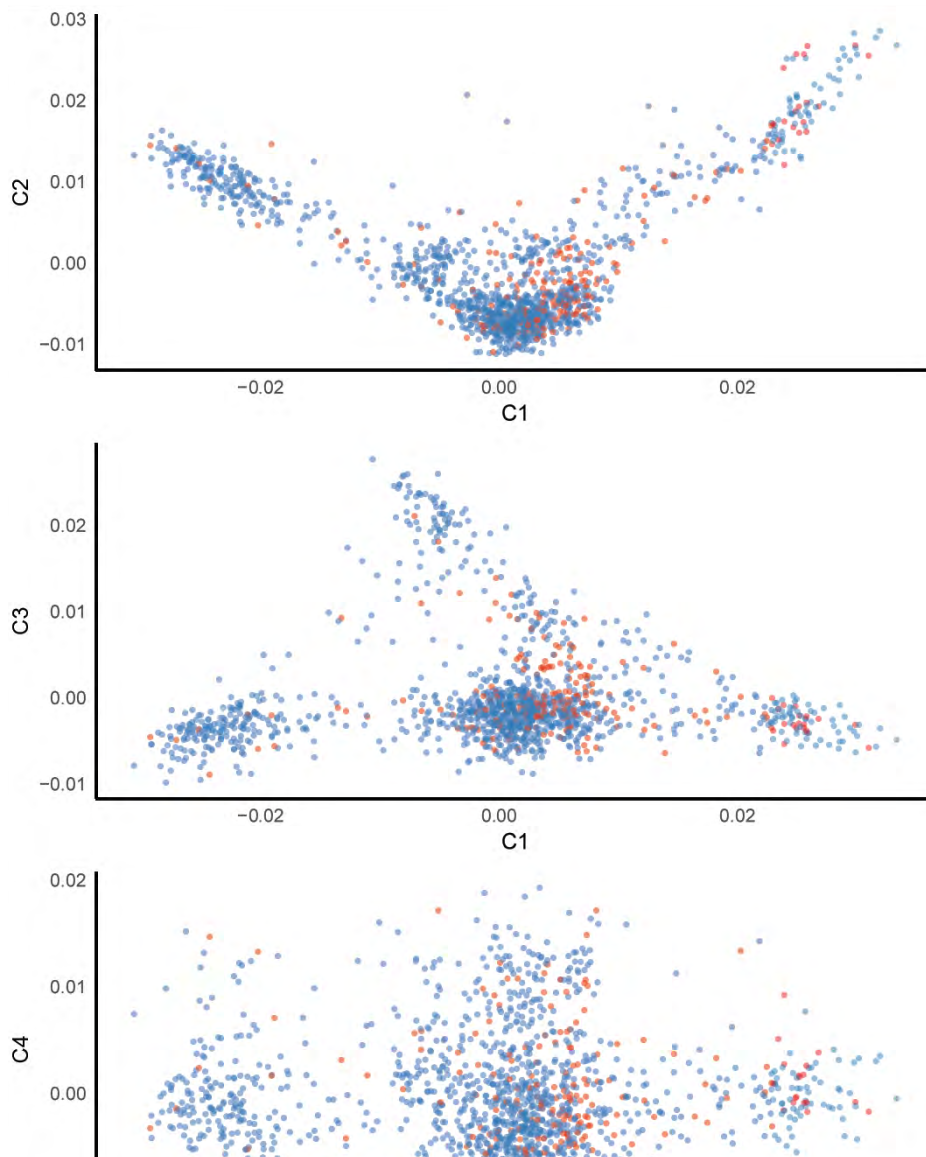
Legend

Multidimensional scaling plots plot of genetic ancestry in AAP cases and controls from the CEU population. Defined as individuals >16 standard deviations away from the HapMap-defined CEU (Northern and Western European) centroid mean.

Supplemental figure 7

Quartile-quartile plot of CEU cohort
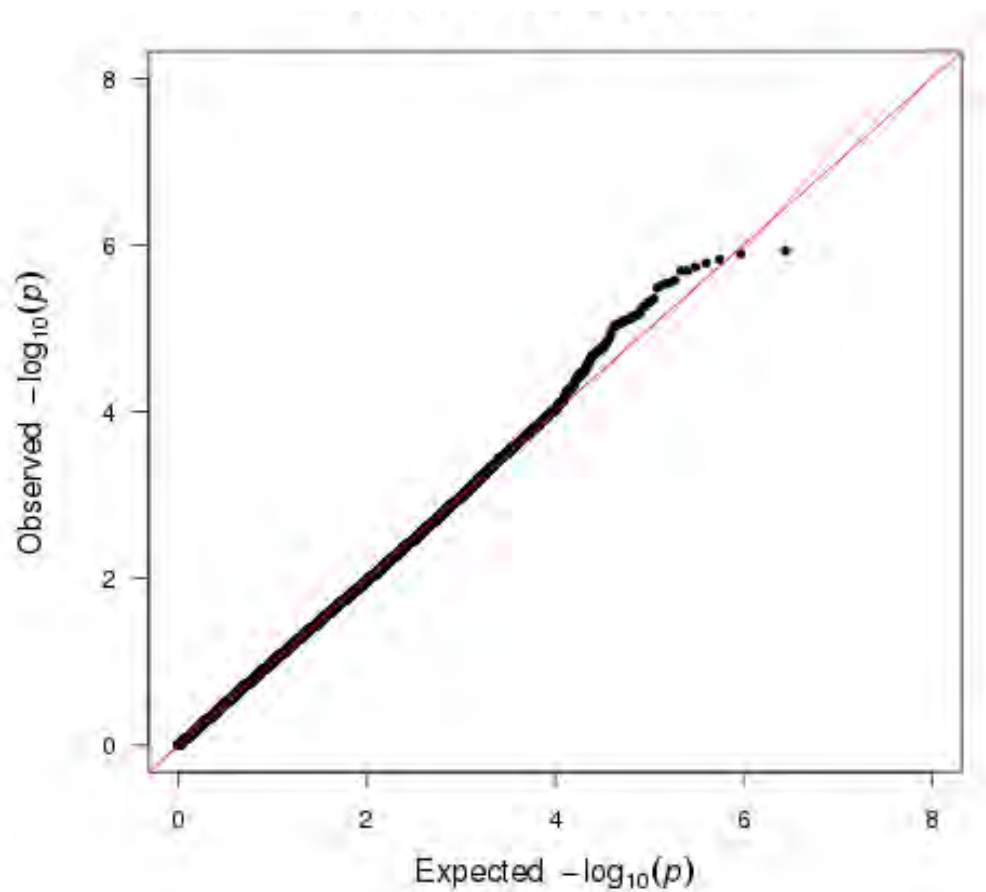
Legend

Quartile-quartile plot showing the expected *vs* the observed p-values in the CEU cohort of 205 cases and 1185 controls. The logistic regression analysis was adjusted for age and genetic ancestry, with a lambda of 1.02 there was no evidence of population substructure.

Supplemental figure 8

Manhattanplot of CEU cohort
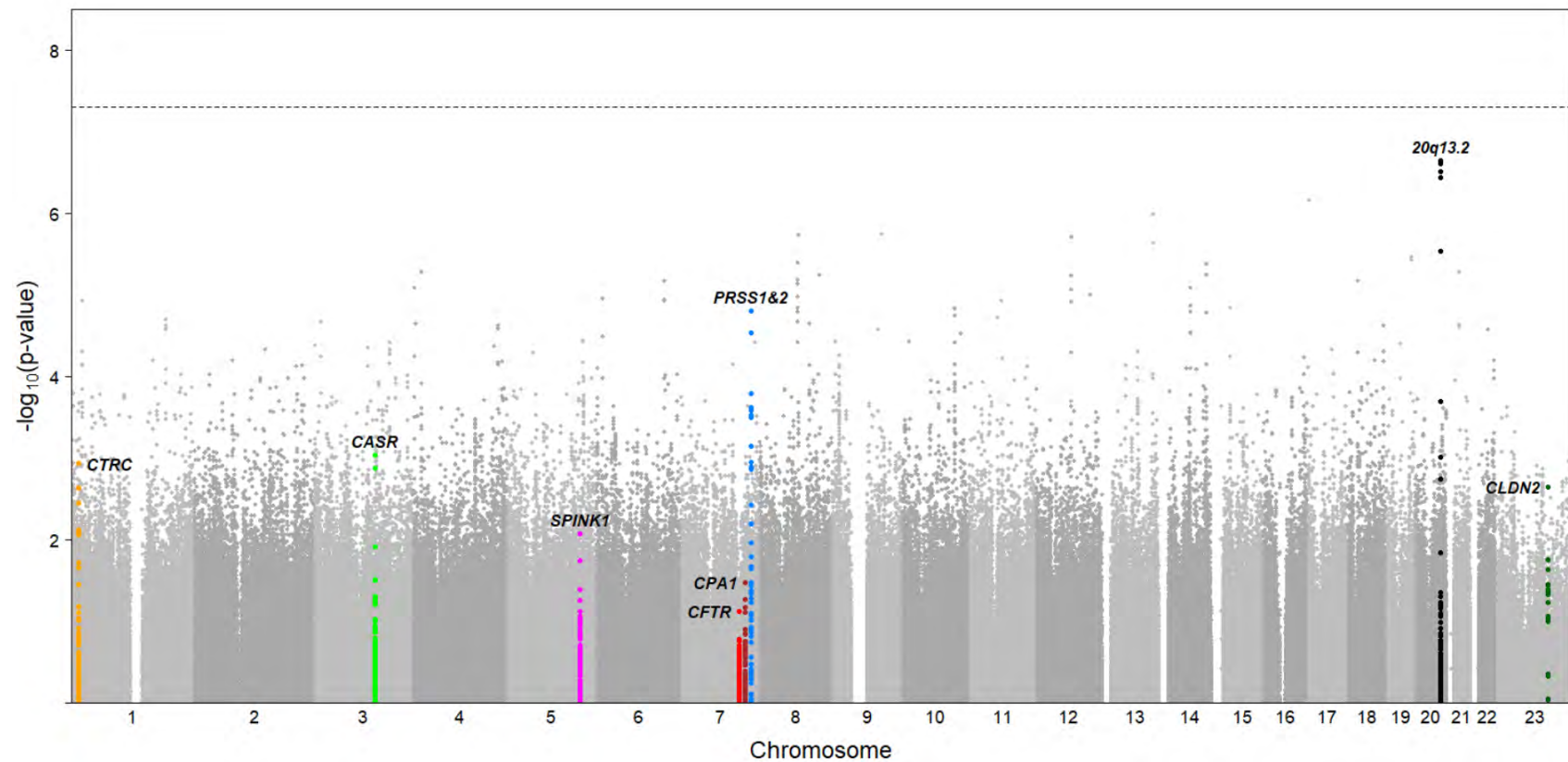
Legend

Manhattanplot showing SNPs associated with asparaginase-associated pancreatitis in 205 cases and 1185 controls in the CEU cohort. The *x* axis represents genomic location, and the *y* axis represents the *P* value for the SNP's association calculated using logistic regression adjusting for age and ancestry. Genes previously associated to pancreatitis are marked in color. SNPs are annotated to genes based on genomic location (10 kb

Supplemental table 2

Top SNPs in genes previously associated with risk of pancreatitis in CEU cohort

Gene	CHR	BP	SNP	Major>minor allele	MAF Cases	MAF Controls	OR (95%CI)	P
CASR	3	121913370	rs937627	G>T	0.17	0.23	0.73 (0.55–0.97)	0.03
CASR	3	121908434	rs4678029	T>C	0.17	0.23	0.73 (0.55–0.97)	0.03
CASR	3	121919740	rs16832787	G>A	0.17	0.23	0.73 (0.55–0.97)	0.03
CASR	3	121936323	rs13320637	G>A	0.22	0.27	0.76 (0.59–0.98)	0.04
CASR	3	121936370	rs13327652	A>G	0.22	0.27	0.76 (0.59–0.99)	0.04
CFTR	7	117145102	rs56296320	T>C	0.007	0.02	0.34 (0.11–1.11)	0.07
CFTR	7	117178754	rs17449197	A>G	0.11	0.14	0.79 (0.57–1.09)	0.15
CFTR	7	117119183	rs4148682	T>G	0.1	0.08	1.26 (0.88–1.82)	0.21
CFTR	7	117181509	rs4148703	G>A	0.1	0.09	1.26 (0.87–1.81)	0.22
CFTR	7	117219835	rs1469486	C>T	0.13	0.12	1.22 (0.88–1.68)	0.23
CLDN2	23	106134938	rs12853674	C>T	0.12	0.08	1.67 (1.11–2.53)	0.01
CLDN2	23	106140325	rs4409525	G>A	0.34	0.28	1.4 (1.06–1.83)	0.02
CLDN2	23	106160702	rs12008279	A>G	0.51	0.44	1.31 (1.02–1.69)	0.03
CLDN2	23	106183670	rs12014762	C>T	0.22	0.17	1.32 (0.97–1.8)	0.07
CLDN2	23	106136910	rs5962770	T>C	0.28	0.3	0.98 (0.74–1.3)	0.87
CPA1	7	130037805	rs73152870	A>G	0.039	0.02	1.67 (0.92–3.01)	0.09
CPA1	7	130018863	rs10954269	C>T	0.09	0.1	0.77 (0.53–1.12)	0.18
CPA1	7	130033556	rs17389898	T>C	0.07	0.05	1.36 (0.87–2.12)	0.18
CPA1	7	130019491	rs13226219	T>C	0.09	0.1	0.79 (0.54–1.15)	0.21
CPA1	7	130028089	rs17330508	C>T	0.037	0.03	1.3 (0.71–2.36)	0.39
CTRB1-2	16	75254970	rs57833904	C>T	0.015	0.005	2.56 (0.92–7.16)	0.07
CTRB1-2	16	75230638	rs1019537	C>T	0.13	0.15	0.74 (0.54–1.03)	0.07
CTRB1-2	16	75230739	rs1019539	C>T	0.12	0.15	0.74 (0.54–1.03)	0.08
CTRB1-2	16	75230230	rs1559362	T>C	0.32	0.34	0.85 (0.67–1.07)	0.16
CTRB1-2	16	75263661	rs7190458	G>A	0.06	0.05	1.39 (0.87–2.24)	0.17
CTRC	1	15758963	rs35994710	C>T	0.16	0.23	0.63 (0.47–0.84)	0.002
CTRC	1	15768304	rs10436957	G>A	0.17	0.24	0.65 (0.49–0.86)	0.003
CTRC	1	15757666	rs10754889	G>A	0.30	0.36	0.71 (0.56–0.9)	0.004
CTRC	1	15760092	rs7541863	C>T	0.31	0.37	0.71 (0.56–0.9)	0.005
CTRC	1	15763340	rs6693417	A>G	0.33	0.4	0.72 (0.57–0.91)	0.005
PRSS1-2	7	142473466	rs13228878	A>G	0.31	0.42	0.6 (0.48–0.76)	2.1×10^{-5}
PRSS1-2	7	142456928	rs10273639	C>T	0.32	0.43	0.62 (0.49–0.78)	3.8×10^{-5}
PRSS1-2	7	142487836	rs2734222	C>T	0.38	0.47	0.66 (0.53–0.82)	0.0002
PRSS1-2	7	142455538	rs3757377	G>A	0.3	0.39	0.65 (0.52–0.83)	0.0004
PRSS1-2	7	142488688	rs2734224	G>A	0.39	0.47	0.7 (0.56–0.87)	0.001
SPINK1	5	147207678	rs17107315	A>G	0.02	0.01	2.78 (1.29–5.97)	0.009
SPINK1	5	147220041	rs4705045	T>G	0.11	0.14	0.73 (0.52–1.02)	0.07
SPINK1	5	147211393	rs4705203	A>G	0.12	0.14	0.78 (0.56–1.08)	0.13
SPINK1	5	147204192	rs11319	C>T	0.05	0.05	1.35 (0.82–2.2)	0.23
SPINK1	5	147205839	rs17703305	G>T	0.46	0.43	1.13 (0.91–1.41)	0.27

Legend

Top-five SNPs associated with asparaginase-associated pancreatitis in the CEU cohort of 205 cases and 1185 controls. SNPs were annotated to genes if ≤10kb upstream or downstream from transcription start site or transcription terminator, respectively. Genefunctions below are defined by Genecards (www.genecards.org) and UniPort (www.uniprot.org).

Abbreviations: SNP, single nucleotide polymorphism; Chr, chromosome; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval; *CASR*, calcium-sensing receptor (G protein-coupled receptor, small changes in circulating calcium concentration are monitored affecting intracellular signaling pathways); CFTR, cystic fibrosis transmembrane conductance regulator (cAMP-regulated cell channel, conducting chloride and thiocyanate ions across epithelial membranes); CLDN2, claudin-2 (Major integral membrane protein localized exclusively at tight junctions); CPA1: Carboxypeptidase A1 (Member of the serine protease family, A1 form of the pancreatic procarboxypeptidase produced in pancreatic acinar cells, preferentially cleaves C-terminal branched-chain and aromatic amino acids from dietary proteins); CTRB1-2m, chymotrypsin B1-B2 (member of the serine protease family of enzymes, regulating activation and degradation of trypsinogens and procarboxypeptidases by targeting specific cleavage sites within their zymogen precursors); CTRC: chymotrypsin C (serum calcium-decreasing factor with chymotrypsin-like protease activity, regulating activation and degradation of trypsinogens and procarboxypeptidases by targeting specific cleavage sites within their zymogen precursors); PRSS1-2, protease, serine, 2 (encoding for cationic and anionic trypsinogen); SPINK1, serine peptidase inhibitor, kazal type 1 (Prevention of trypsin-catalyzed premature activation of zymogens within the pancreas and the pancreatic duct).

Supplemental table 3

Top SNPs in genes previously associated with risk of pancreatitis in the total cohort

Gene	CHR	BP	SNP	Major>minor allele	MAF Cases	MAF Controls	OR (95% CI)	P
CASR	3	121965199	rs9859381	G>T	0.18	0.13	1.3 (0.99-1.71)	0.06
CASR	3	121913370	rs937627	G>T	0.19	0.23	0.78 (0.61-1.01)	0.063
CASR	3	121919740	rs16832787	G>A	0.19	0.23	0.78 (0.61-1.01)	0.064
CASR	3	121908434	rs4678029	T>C	0.19	0.23	0.79 (0.61-1.02)	0.067
CASR	3	121971512	rs937625	T>G	0.11	0.08	1.37 (0.98-1.92)	0.068
CFTR	7	117145102	rs56296320	T>C	0.01	0.02	0.3 (0.09-0.97)	0.044
CFTR	7	117178754	rs17449197	A>G	0.11	0.14	0.77 (0.56-1.04)	0.09
CFTR	7	117181509	rs4148703	G>A	0.12	0.09	1.3 (0.94-1.79)	0.11
CFTR	7	117119183	rs4148682	T>G	0.12	0.09	1.28 (0.92-1.76)	0.14
CFTR	7	117279147	rs6466615	G>A	0.01	0.01	0.4 (0.12-1.39)	0.15
CLDN2	X	106134938	rs12853674	C>T	0.11	0.08	1.71 (1.16-2.52)	0.007
CLDN2	X	106140325	rs4409525	G>A	0.34	0.28	1.35 (1.05-1.73)	0.02
CLDN2	X	106160702	rs12008279	A>G	0.53	0.45	1.28 (1.01-1.62)	0.04
CLDN2	X	106183670	rs12014762	C>T	0.21	0.17	1.3 (0.97-1.74)	0.08
CLDN2	X	106136910	rs5962770	T>C	0.26	0.29	0.99 (0.76-1.3)	0.1
CPA1	7	130037805	rs73152870	A>G	0.04	0.02	1.88 (1.08-3.27)	0.03
CPA1	7	130028089	rs17330508	C>T	0.04	0.03	1.59 (0.92-2.74)	0.1
CPA1	7	130018863	rs10954269	C>T	0.09	0.10	0.75 (0.53-1.06)	0.1
CPA1	7	130019491	rs13226219	T>C	0.10	0.11	0.77 (0.55-1.09)	0.14
CPA1	7	130033556	rs17389898	T>C	0.07	0.05	1.36 (0.89-2.06)	0.16
CTRB1-2	16	75252306	rs8056797	G>T	0.03	0.01	2.63 (1.22-5.69)	0.01
CTRB1-2	16	75254970	rs57833904	C>T	0.03	0.01	2.46 (1.17-5.16)	0.02
CTRB1-2	16	75230638	rs1019537	C>T	0.12	0.14	0.74 (0.54-1)	0.05
CTRB1-2	16	75230739	rs1019539	C>T	0.12	0.14	0.75 (0.55-1.01)	0.06
CTRB1-2	16	75263661	rs7190458	G>A	0.06	0.04	1.45 (0.93-2.26)	0.1
CTRC	1	15768304	rs10436957	G>A	0.17	0.23	0.69 (0.53-0.89)	0.005
CTRC	1	15758963	rs35994710	C>T	0.16	0.22	0.68 (0.52-0.89)	0.005
CTRC	1	15760092	rs7541863	C>T	0.33	0.38	0.76 (0.61-0.94)	0.01
CTRC	1	15763340	rs6693417	A>G	0.35	0.41	0.77 (0.62-0.95)	0.01
CTRC	1	15757666	rs10754889	G>A	0.32	0.37	0.77 (0.62-0.96)	0.02
PRSS1-2	7	142473466	rs13228878	A>G	0.35	0.44	0.61 (0.5-0.76)	7.1×10^{-6}
PRSS1-2	7	142456928	rs10273639	C>T	0.35	0.44	0.62 (0.5-0.77)	1.1×10^{-5}
PRSS1-2	7	142487836	rs2734222	C>T	0.40	0.48	0.65 (0.53-0.8)	4.8×10^{-5}
PRSS1-2	7	142488688	rs2734224	G>A	0.41	0.49	0.68 (0.55-0.83)	0.0002
PRSS1-2	7	142455538	rs3757377	G>A	0.33	0.40	0.67 (0.54-0.83)	0.0002
SPINK1	5	147207678	rs17107315	A>G	0.02	0.01	2.86 (1.4-5.85)	0.004
SPINK1	5	147220041	rs4705045	T>G	0.12	0.14	0.73 (0.54-1)	0.05
SPINK1	5	147215120	rs4705204	T>G	0.23	0.18	1.23 (0.96-1.57)	0.1

| SPINK1 | 5 | 147211393 | rs4705203 | A>G | 0.13 | 0.15 | 0.78 (0.58-1.06) | 0.11 |
| SPINK1 | 5 | 147205839 | rs17703305 | G>T | 0.46 | 0.42 | 1.18 (0.96-1.44) | 0.12 |

Legend

Top-five SNPs associated with asparaginase-associated pancreatitis in 244 cases and 1320 controls. SNPs were annotated to genes if ≤10kb upstream or downstream from transcription start site or transcription terminator, respectively. SNPs and genes were identified from litterature search (see methods section), genefunctions are defined by Genecards (www.genecards.org) and UniPort (www.uniprot.org).

Abbreviations: SNP, single nucleotide polymorphism; Chr, chromosome; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval; *CASR*, calcium-sensing receptor (G protein-coupled receptor, small changes in circulating calcium concentration are monitored affecting intracellular signaling pathways); *CFTR*, cystic fibrosis transmembrane conductance regulator (cAMP-regulated cell channel, conducting chloride and thiocyanate ions across epithelial membranes); *CLDN2*, claudin-2 (Major integral membrane protein localized exclusively at tight junctions); *CPA1*: Carboxypeptidase A1 (Member of the serine protease family, A1 form of the pancreatic procarboxypeptidase produced in pancreatic acinar cells, preferentially cleaves C-terminal branched-chain and aromatic amino acids from dietary proteins); *CTRB1-2*, chymotrypsin B1-B2 (member of the serine protease family of enzymes, regulating activation and degradation of trypsinogens and procarboxypeptidases by targeting specific cleavage sites within their zymogen precursors); *CTRC*: chymotrypsin C (serum calcium-decreasing factor with chymotrypsin-like protease activity, regulating activation and degradation of trypsinogens and procarboxypeptidases by targeting specific cleavage sites within their zymogen precursors); *PRSS1-2*, protease, serine, 2 (encoding for cationic and anionic trypsinogen); *SPINK1*, serine peptidase inhibitor, kazal type 1 (Prevention of trypsin-catalyzed premature activation of zymogens within the pancreas and the pancreatic duct).

Supplemental figure 9

5-year event free survival according to *PRSS1-PRSS2* genotype
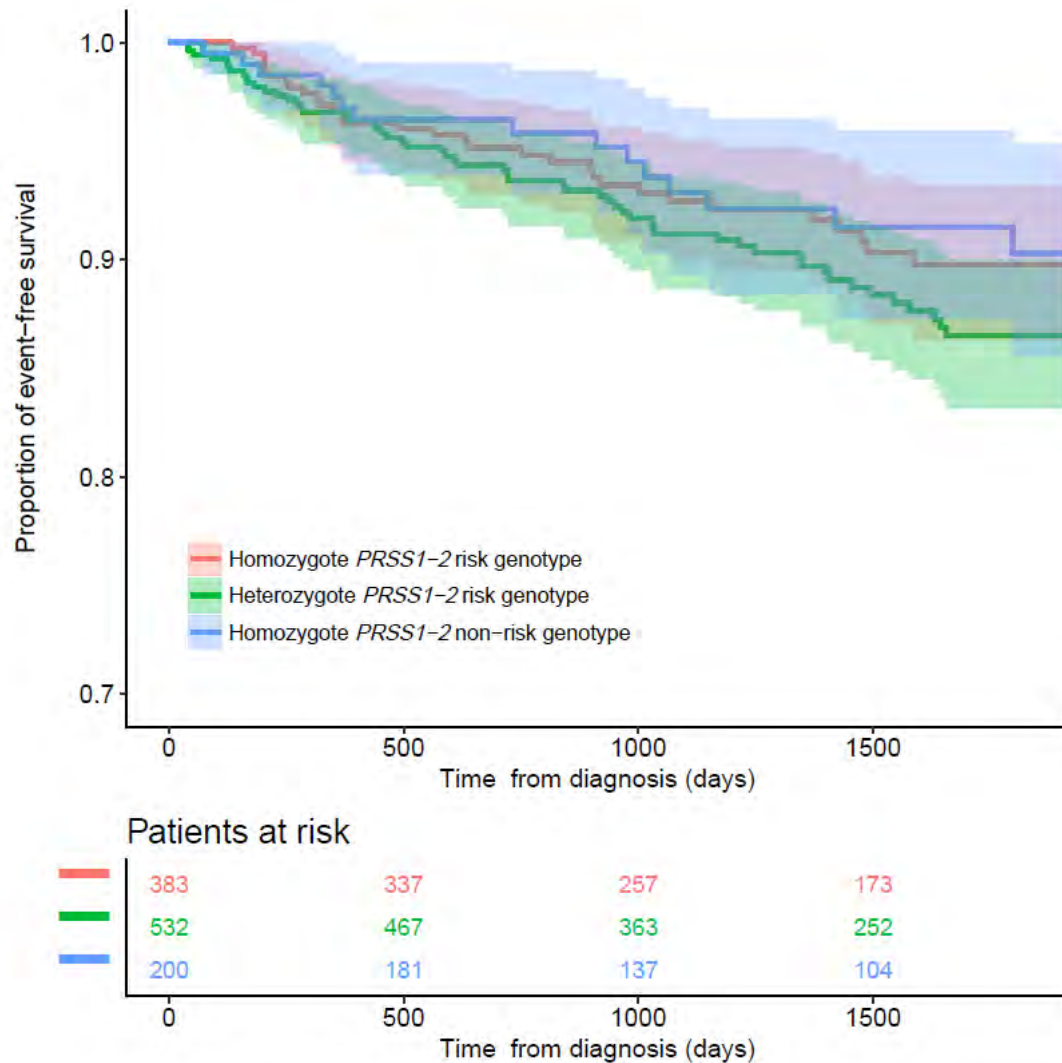
Legend

Five-year event free survival in the Nordic subset of cases (n=92) and controls (n=1024) grouped according to PRSS1-2 genotype (rs13228878). Event was defined as death, relapse or second malignant neoplasm. No difference in five-years event-free survival was found using the 2-sided log-rank test (*P*=0.4). See through colors denote 95% confidence intervals.