Analysis of transcriptional changes associated with pubertal development

Justyna Resztak*[1], Jane Choe*[1], Julong Wei[1], Rachel Bruinsma[2], Russell Houpt[2], Adnan Alazizi[1], Henriette E. Mair-Meijers[1], Richard B. Slatcher[3], Samuele Zilioli[2,4], Roger Pique-Regi[1,5+], Francesca Luca[1,5+].

1 Center for Molecular Medicine and Genetics, Wayne State University, Detroit, Michigan, USA
2 Department of Psychology, Wayne State University, Detroit, Michigan, USA
3 Department of Psychology, University of Georgia, Athens, Georgia, USA
4 Department of Family Medicine and Public Health Sciences, Wayne State University, Detroit, Michigan, USA
5 Department of Obstetrics and Gynecology, Wayne State University, Detroit, Michigan 48201, USA
* Equal contribution
+ To whom correspondence should be addressed: fluca@wayne.edu, rpique@wayne.edu

Abstract

Puberty is an important developmental period marked by hormonal, metabolic and immune changes, which have been implicated in predisposition to immune diseases later in life. Yet, little is known about the gene expression changes in immune cells that occur during pubertal development. In a longitudinal cohort of 251 children we assessed pubertal development and leukocyte gene expression. We identified largely overlapping gene expression changes in boys and girls for thousands of genes over a short time period. For 108 genes we detected gene expression changes associated with pubertal development. Substantial transcriptional changes were detected between pre- and post-menarcheal girls, which suggested a shift from predominantly innate to adaptive immunity as girls sexually matured.

We identified genetic effects on gene expression that dynamically change during pubertal development for *IGKV1-27* and *PGAP1* in males, *DSC1* and *TRBV30* in females. We discovered that menarche is associated with a change in sign of a genetic effect on the expression of the asthma-associated gene FLOT2. Using a Transcriptome-Wide Association Study approach, we discovered that in girls, gene expression changes during puberty are associated with age at menarche. For example, expression of the endometrial cancer risk gene *EEFSEC* increases as girls become older, suggesting a link between age at menarche, transcriptional changes happening during puberty and poor health outcomes. These findings shed light on immune gene expression processes accompanying puberty and can advance the understanding of the molecular mechanisms through which pubertal development relates to immune diseases later in life.

Introduction

Puberty is an important developmental period marked by hormonal, metabolic and immune changes, which have been implicated in disease predisposition later in life. Starting at puberty the physiology of men and women is influenced by different hormones - primarily testosterone in males and estrogen and progesterone in females - with broad effects on all body systems[1]. Physiological changes during this period that are common to both sexes include growth spurts, increase in BMI[2], growth of body hair, skin changes (increased oil production and acne), and the maturation of gonads[3]. Girls experience breast growth and menarche, while boys undergo deepening of the voice and develop facial hair[3]. Timing of puberty also differs between girls and boys - the average age range for puberty is approximately 12-16 years in males and 10-14 years in females[4].

In addition to these physiological changes, certain pathological conditions also manifest during puberty or develop sex-specific symptom profiles[5–8]. In particular, incidence of many autoimmune diseases increases in the peri-pubertal period, with females bearing a higher disease risk than males[9–12]. Puberty also marks a shift in sex differences in susceptibility to asthma - boys have higher asthma prevalence in childhood, but in adulthood females are more prone to asthma[13]. Age at menarche is also known to have substantial effects on growth and disease[14]. Earlier menarche has been associated with increased risk of asthma[15], cardiovascular disease[16], type 2 diabetes[17] and higher BMI[14]. Despite sex differences in susceptibility to asthma before and after puberty onset[13], earlier onset of puberty has been found to increase the risk of asthma in both sexes[15]. Sex hormones are hypothesized to play a role in the switch of sex skew, as asthma symptoms are also known to worsen around menstruation[18,19]. Recently, a Mendelian randomization study showed a protective effect of increased sex hormone-binding globulin (SHBG) levels on asthma onset, with a larger effect in females than males[20]. A thorough understanding of the longitudinal changes in the immune system that accompany puberty is needed to understand their impacts on asthma and other disease susceptibility.

The molecular mechanisms underlying the physiological and pathological changes in the immune system during puberty remain understudied and largely uncharacterized. Previous studies have explored the effects of puberty on DNA methylation[21–26] in peripheral blood immune cells. DNA methylation is an important epigenetic mark regulating gene expression during development and aging[27]. Hormonal changes in puberty have been associated with changes in DNA methylation[21]. For example, differentially methylated regions between pre- and post-pubertal females were enriched for estrogen response element, reproductive hormone signaling and immune and inflammatory responses, while differentially methylated regions in males were enriched for genes involved in adrenaline and noradrenaline biosynthesis[22]. Prior studies found a larger number of genomic regions changing methylation status during puberty in females than in males[22,28]. Furthermore, regions experiencing changes in methylation status in both sexes show changes with higher magnitude in females[24]. In contrast, other studies on pre- and post-pubertal subjects found limited differences in DNA methylation patterns between sexes[23,25]. These inconsistencies may be due to age differences between studies and could be further investigated by employing a study design that focuses on all pubertal developmental stages rather than comparing pre- and post-pubertal stages only.

Although DNA methylation findings support the hypothesis that blood gene expression during puberty undergoes reprogramming, little is known about the transcriptional changes accompanying human pubertal development, particularly from a genome-wide perspective. Most studies to date explored epigenetic differences between pre- and post-puberty, yet we don't have a complete picture of the gene expression changes during pubertal development. To fill these gaps, we analyzed a longitudinal cohort of 103 girls and 148 boys aged 10-17 years old and diagnosed with childhood onset asthma. We investigated the associations between pubertal factors (i.e., age, pubertal stage, and menarche status in females) and genome-wide gene expression in blood. To assess the contribution of genetic factors, we also explored the effects of genetic variation on gene expression changes associated with puberty using an expression quantitative trait locus (eQTL) mapping approach.

Results

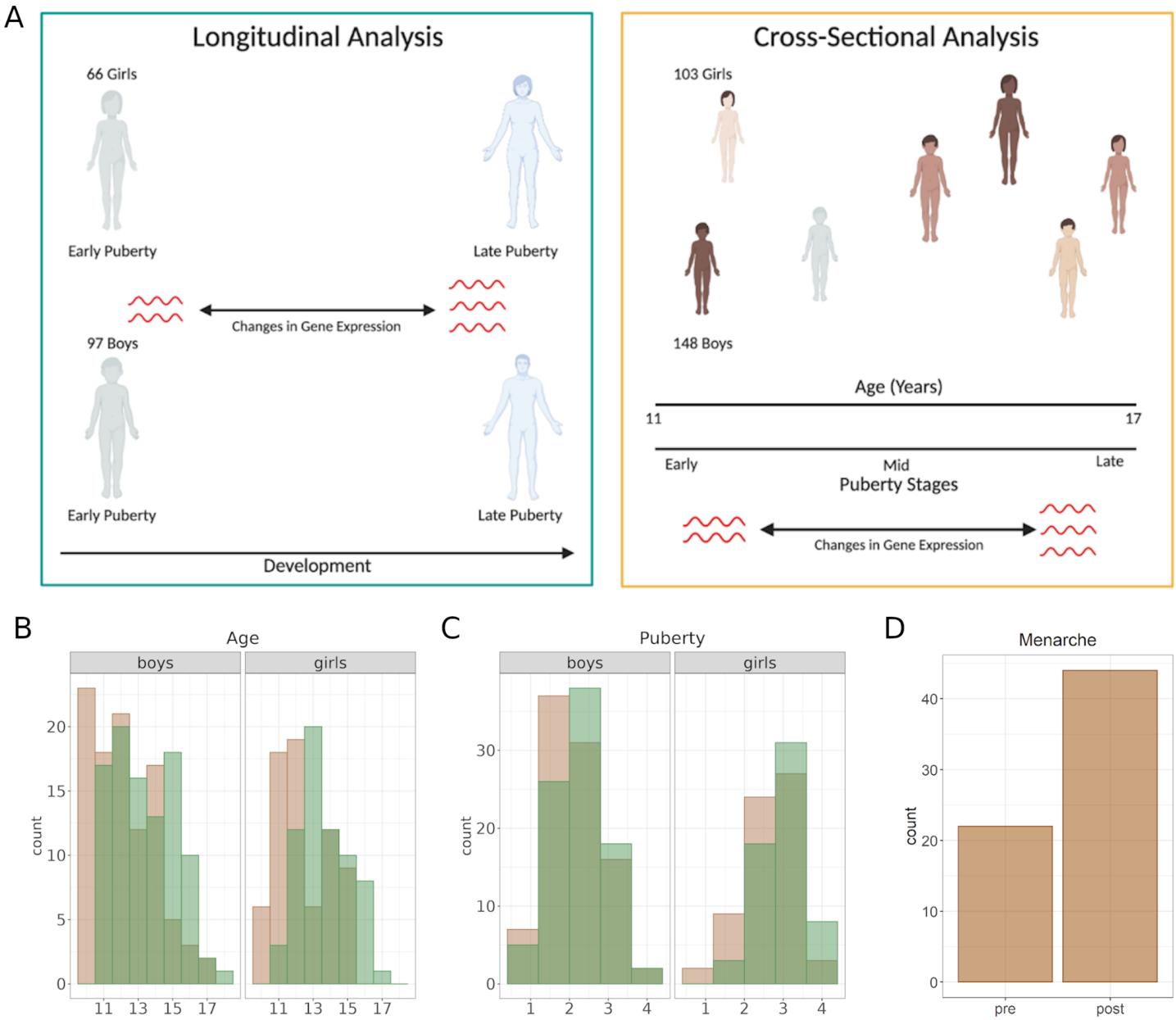

Fig. 1. Analysis of gene expression in peri-pubertal boys and girls. A – Study design. We employed longitudinal and cross-sectional study designs to investigate the patterns of gene expression associated with pubertal development, B – Histograms of age distribution in boys (left panel) and girls (right panel) at baseline (tan) and at follow-up (green), C – Histograms of Pubertal Development Score[29] distribution in boys (left panel) and girls (right panel) at baseline (tan) and at follow-up (green), D - Bar graph of menarcheal status in the cross-sectional sample.

Gene expression changes over time in peri-pubertal boys and girls

Participants in this study were recruited as part of the longitudinal Asthma in the Lives of Families Today (ALOFT) study. The ALOFT study recruited boys and girls with childhood onset asthma and between 10 and 17 years old from Metropolitan Detroit. Children were followed annually for up to three years.

We collected gene expression, genotype, puberty and menarche data for a total of 251 participants. For 163 participants we collected data for two timepoints: baseline and follow-up (Fig. 1, Table S1). Average time to

follow-up was 1.27 years (SD=0.46). The entire range of puberty stages, measured using a scale equivalent to Tanner stages[29], were represented for both sexes.

We employed a longitudinal study design to investigate leukocyte gene expression changes in peri-pubertal boys and girls. We modeled gene expression with DESeq2[30] in each sex separately and identified 160 genes in females and 535 genes in males (Fig. 2A, Fig. S2, Files S1, S2), whose expression changed between the two time points. These changes are likely to be associated with age, but also with pubertal development and other developmental changes that happen during this period of life in children. The difference in the number of differentially expressed genes in each sex likely reflected higher statistical power in boys than girls (due to larger sample size). In agreement with this explanation is the observation that genes differentially expressed as girls grew older were enriched among genes differentially expressed in boys (OR=9, p-value=$6.5*10^{-16}$, Fig. 2B), suggesting a similar effect of age on gene expression in both sexes. To further validate these results, we considered all 251 individuals with at least one observation and used a cross-sectional study design to identify gene expression differences associated with the age of the donor. We observed a significant correlation between the gene expression changes measured longitudinally and cross-sectionally (Spearman rho=0.43 in girls, rho=0.3 in boys, p-value<$2.2*10^{-16}$), thus strengthening the evidence of gene expression changes associated with age in peri-pubertal children (Fig. 2C and D, Files S3, S4).

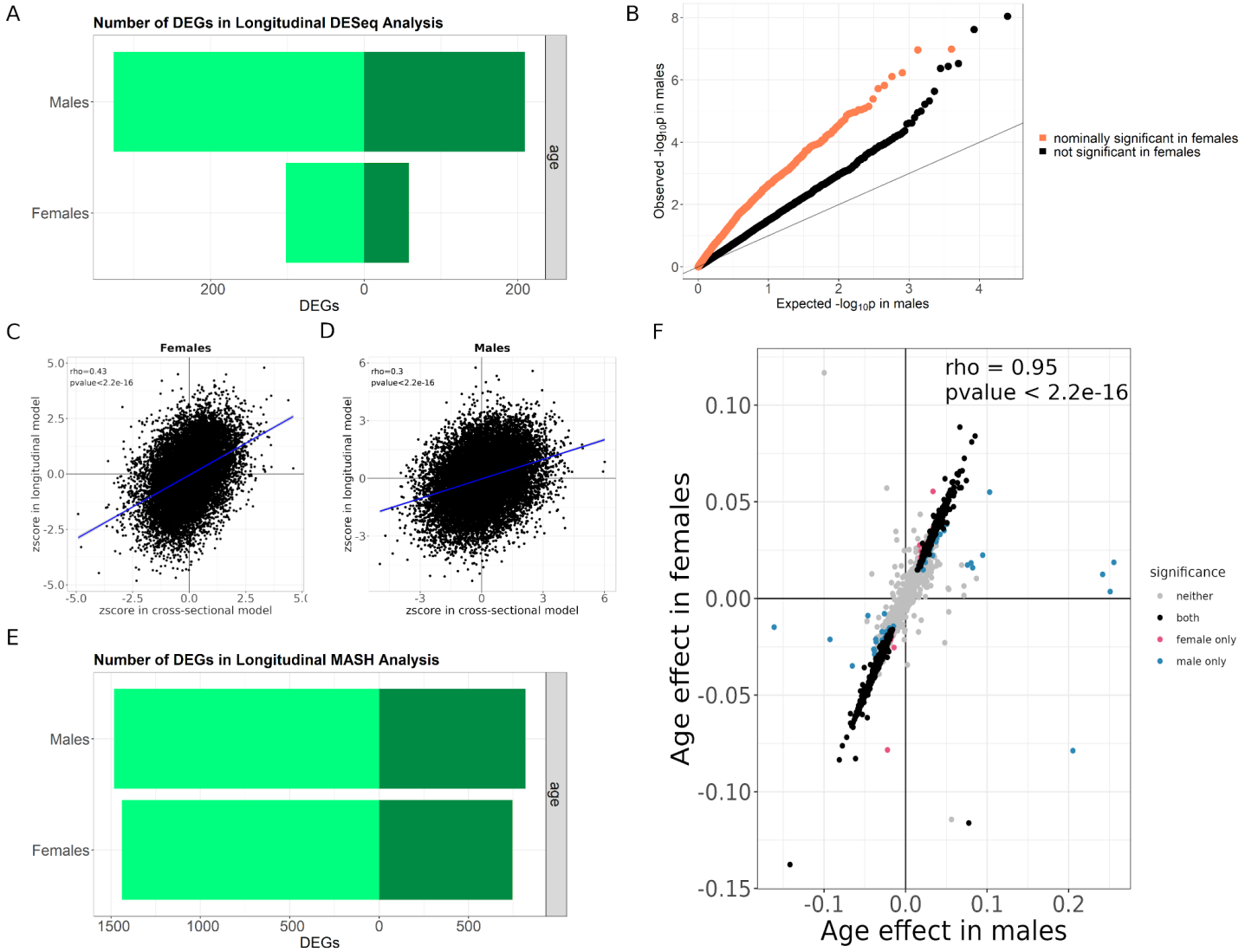

Fig. 2. Gene expression changes over time in peri-pubertal boys and girls. A – Barplot with the number of genes differentially expressed with age in the two sexes (10% FDR). Light green denotes downregulated genes, dark green denotes upregulated genes, B – QQplot of p-values of age effect on gene expression in males, split by whether the gene was also nominally differentially expressed in females (p-

value<0.05, coral) or not significant (p-value>0.05, black), C and D - Scatterplot of z-scores of age effects on gene expression in females (C) and males (D) analyzed with the cross-sectional model (x axis) or longitudinal model (y axis). Rho - Spearman correlation and its p-value. Blue line represents the linear regression trendline, E - Barplot with the number of genes differentially expressed with age in the two sexes analyzed together with multivariate adaptive shrinkage (mash, 10% LFSR). Light green denotes downregulated genes, dark green denotes upregulated genes, F - Scatterplot of age effects on gene expression in males (x axis) and females (y axis) from mash analysis in both sexes jointly. Color denotes significance (10% LFSR): black – significant in both sexes, pink – significant in females only, blue – significant in males only, grey – not significant in either sex.

To directly investigate whether the observed longitudinal changes in gene expression were similar in boys and girls, we used multivariate adaptive shrinkage (mash)[31]. This method uses correlations between effects measured in different conditions (e.g., males and females) to improve effect size estimates, increasing the power of discovery and enabling direct comparison between effects measured in different individuals. Applying this method to the results of the longitudinal analysis, we discovered over two thousand differentially expressed genes (Fig. 2E, Files S5, S6). Specifically, we found 818 genes in boys and 746 genes in girls being upregulated when children grew older, and 1482 genes in boys and 1439 genes in girls being downregulated when children grew older. We found that these gene expression changes were largely shared between sexes (Fig. 2F, Fig. S3), this could be because we have limited power to detect sex-specific effects of age and puberty with the current sample size. One of the genes whose expression increased dramatically as a function of age was *PLXND1* (Fig. 3A top), which encodes a semaphorin receptor important for thymocyte development[32]. One of the genes most strongly downregulated as a function of age was *MACROD2* (Fig. 3A bottom), which encodes a deacetylase involved in removing ADP-ribose from mono-ADP-ribosylated proteins. Mutations near this gene have been implicated in a form of congenital hypogonadism[33–35].

There are limited data available on immune gene expression changes throughout the lifespan. A recent study investigated gene expression changes between the eighth and ninth decade of life[36], using a longitudinal study design. Here, we asked to what extent age effects on gene expression were similar at different life stages, i.e. during puberty and ageing. Genes differentially expressed in our longitudinal peri-pubertal sample were enriched within genes differentially expressed in the elderly (OR = 1.2, p-value = 0.03). Among the genes differentially expressed in both youth and elderly, 44 genes increased expression and 76 decreased expression as a function of age in both groups, for a total of 126 genes. These genes may be unrelated to ageing and puberty, specifically; instead they may be associated with age-related biological processes shared across the lifespan. For 68 genes, age was associated with opposite effects on gene expression in peri-pubertal children and the elderly.

To understand what biological processes were represented by the observed gene expression changes in peri-pubertal children, we performed Gene Ontology and KEGG Pathways enrichment analyses on the longitudinal results. Genes downregulated with age in both sexes were enriched for genes in the Herpes simplex virus infection pathway. Genes upregulated when the children grew older were most strongly enriched for genes involved in biological processes related to carbohydrate metabolism, neutrophil activation and vacuole organization (Fig. 3B)[37]. We also found significant enrichments for genes in the following biological processes and pathways: lysosome, carbon metabolism, and glycan degradation (Fig. 3C).

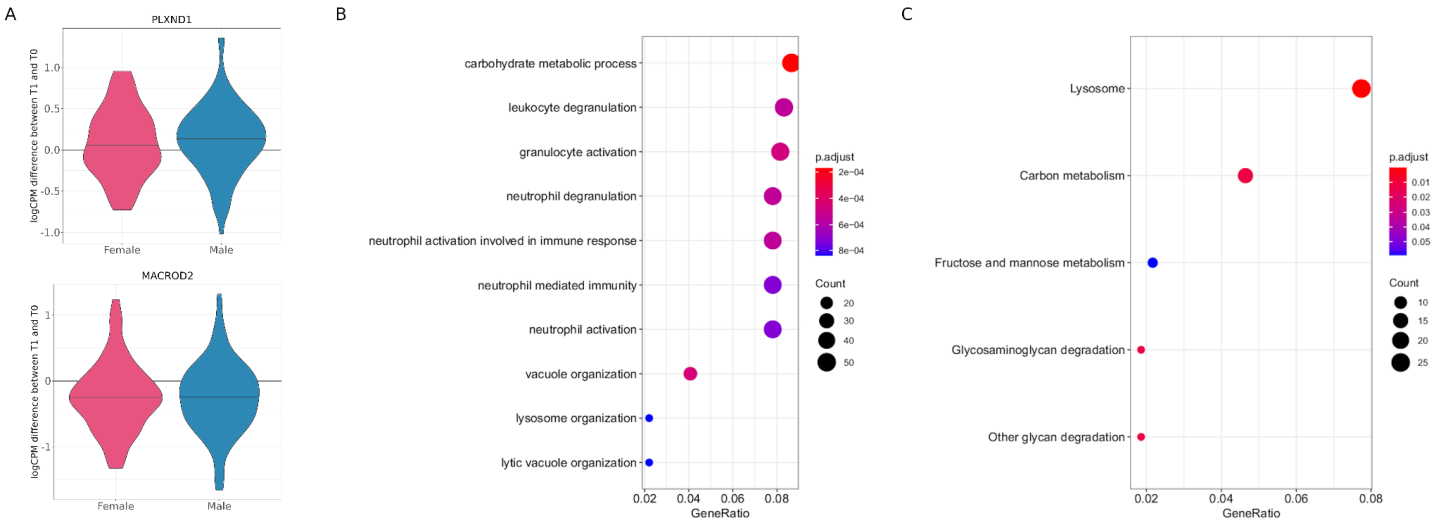

Fig. 3. Biological processes changing with time in peri-pubertal boys and girls. A – Violin plots of the distribution of normalized difference in gene expression between the two timepoints for each individual, plotted separately for each sex, B and C - dotplots represent enrichment of Gene Ontology biological processes (B) or KEGG pathways (C) within genes whose expression is higher in older children.

Gene expression changes across pubertal stages

Despite the association between age and puberty, children of the same age might be at different stages of pubertal development[38]. To directly investigate transcriptional changes in immune cells during pubertal development, we used the same longitudinal study design described above and considered a quantitative pubertal development index analogous to Tanner stages[39,40] that ranged from 1 (no development) to 4 (completed development)[29] (Table S2). We found 108 genes differentially expressed with advancement through puberty in females and none in males (10% FDR) (Fig. 4A, Files S7, S8). These results may reflect the different timing of puberty between girls and boys, with girls in our study being more advanced in their pubertal development, compared to boys. Among the 108 genes differentially expressed as girls progressed through puberty, 72 of them were also differentially expressed as children became older: most genes (66) were downregulated, while only six genes were upregulated.

Our study design allowed us to ask whether similar gene expression changes were observed across individuals experiencing different stages of pubertal development at the time of sampling using the cross-sectional sample. To this end, we considered all individuals with at least one observation in a cross-sectional study design in each sex. We observed significant correlations (Spearman rho=0.13 in girls and rho=0.31 in boys, p-value<$2.2*10^{-16}$) between the gene expression log fold changes from the longitudinal and cross-sectional models (Files S9, S10). Differences in the results between the two models could be due to additional sources of variation which are usually present in cross-sectional datasets, including genetic and environmental effects that vary across individuals. Nevertheless, cross-sectional results support the findings of gene expression changes across pubertal stages from the longitudinal analysis.

To compare the transcriptional effects we discovered with previously-known epigenetic changes associated with pubertal development, we considered the overlap between genes differentially expressed as girls advanced through puberty and 338 genes previously associated with regions differentially methylated in pre- and post-pubertal girls[22]. Three of the 108 genes changing expression as girls advance through puberty had also been found to be differentially methylated between pre- and post-pubertal stages. Of these, *ITGB3* was downregulated in children in later pubertal stages and also hypermethylated, while two others, *TOX* and *HMG20A*, were downregulated but hypomethylated in children that were in a later pubertal stage (Fig. 4B-D). Genes upregulated as girls advanced through puberty were enriched for biological processes related to negative regulation of protein modification. Additionally, several genes were involved in processes important for pubertal

development in females. For example, *MGAT1* and *RPN1* are involved in N-glycan biosynthesis, reflecting previously-reported changes in serum concentration of N-glycans with age[41,42], *TSPO* enables cholesterol translocation into the mitochondria to allow steroid hormone synthesis, and *ACAA1* plays a role in fatty acid metabolism/degradation, potentially reflecting the changes in female body fat distribution that appear during puberty and characterize the adult body[43].

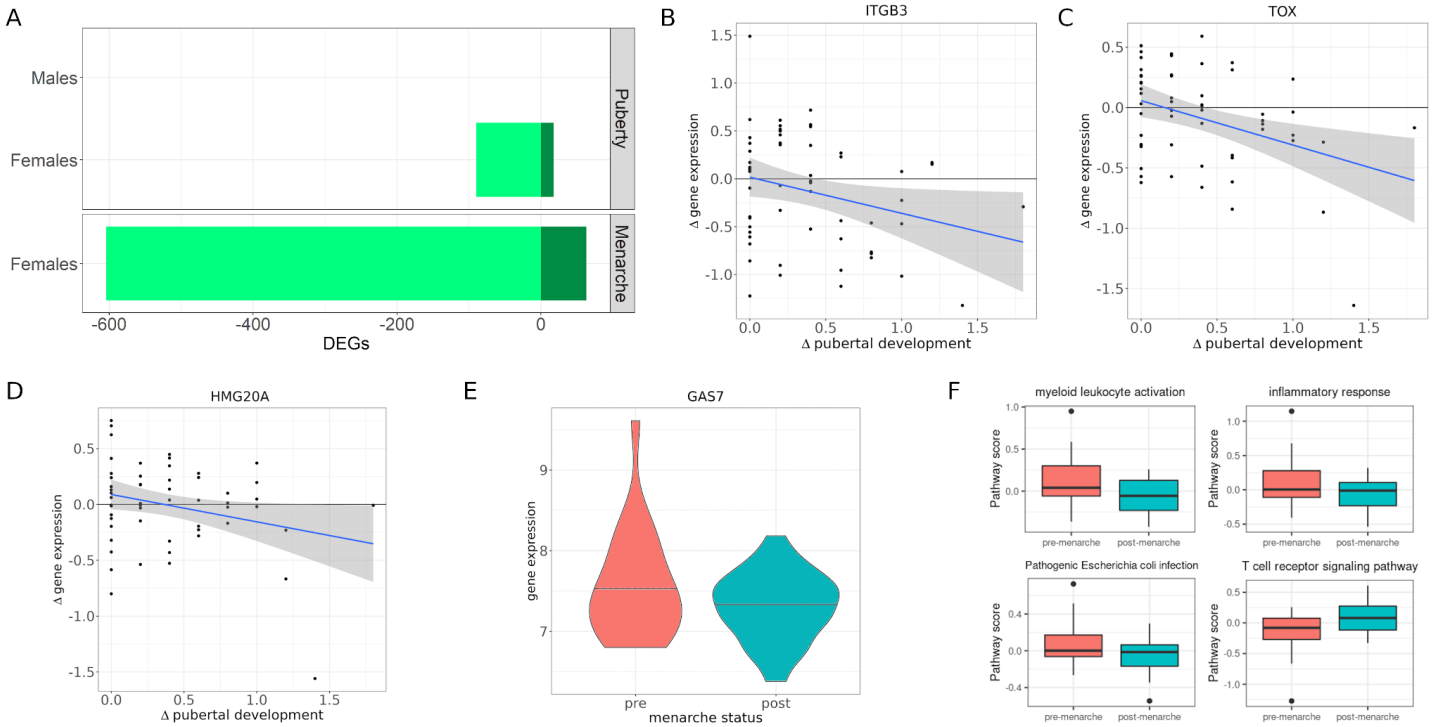

Fig. 4. Gene expression changes across puberty. A - Barplot with the number of genes differentially expressed with puberty and menarche (10% FDR). Light green denotes downregulated genes, dark green denotes upregulated genes. B – D – Scatterplots represent longitudinal change in pubertal development (x axis) against log2-fold change in gene expression (y axis). Blue line represents best linear fit, and shaded grey area its 95% confidence interval. Rho denotes Spearman correlation and its p-value, E – Violin plot represents log2CPM of GAS7 gene expression in pre- and post-menarcheal girls. F – Gene expression across biological processes and pathways differentially expressed between pre- and post-menarcheal girls.

Gene expression changes post-menarche

Menarche is an important milestone of pubertal transition marking sexual maturity in females. Because only four individuals reported to have experienced menarche while enrolled in the study, we utilized a cross-sectional study design to analyze whether gene expression differed between pre-menarche and post-menarche girls (22 and 44 individuals, respectively). We identified 667 differentially expressed genes, 604 genes had higher expression in pre-menarche girls while 63 genes had higher expression in post-menarche girls (File S11).

Genes downregulated after menarche were enriched for GO biological processes involved with immune responses, such as myeloid leukocyte activation, and inflammatory response and KEGG pathways related to bacterial infection. Genes with higher expression after menarche were enriched for the KEGG term T cell receptor signaling, suggesting a shift from predominantly innate to adaptive immunity as girls sexually matured (Fig. 4F, Fig. S4). Genes which differed between pre- and post-menarcheal girls were also more likely to be differentially expressed between boys and girls (OR=3.29, p-value<2.2e-16), suggesting that post-menarcheal gene expression profiles are more sex-specific. Nine of the genes that changed expression past menarche had also been found to have methylation changes between pre- and post-puberty in girls[22]. All of these were downregulated past menarche, and six were hypermethylated in girls post puberty (e.g., Fig. 4E). Genes changing methylation status between pre- and post-puberty had also been found to be enriched for genes harboring nearby estrogen response elements[22]. Estrogen and androgen are the main hormones driving the

developmental changes during puberty in girls and boys, respectively. These hormones bind to their specific hormone receptors, which in turn bind the DNA at response elements and regulate the expression of target genes. The androgen receptor was not expressed in leukocytes in our cohort, however, the estrogen receptor gene (*ESR1*) was expressed both in males and females. To investigate whether estrogen may directly regulate gene expression in immune cells during puberty, we performed a motif scan for estrogen receptor response elements in the cis-regulatory region of differentially expressed genes. We identified binding sites for the estrogen receptor in the regulatory region of 30 genes with transcriptional changes during puberty and 236 genes that were differentially expressed between pre- and post-menarcheal girls. However, we observed no enrichment for genes with nearby estrogen receptor binding sites.

Genetic variation interacts with puberty to affect gene expression

Timing of puberty is influenced by both environmental[44] and genetic[45,46] factors acting primarily through their influence on gene expression[46]. We sought to identify genetic variants whose effect on gene expression changed throughout puberty. We first investigated whether genetic effects on gene expression changed across age. We did not find age to modify the effect of genetic variation on gene expression in either boys or girls. To find genetic variants whose effect on gene expression changes during puberty, we tested for interaction between puberty stage or menarche and genotype on expression of the nearby gene. We found that genetic effects on gene expression varied across puberty stages for four genes: *IGKV1-27* and *PGAP1* in boys, and *TBV30* and *DSC1* in girls. Additionally, in girls, genetic effects on the expression of *FLOT2* varied pre- and post-menarche (Fig. 5, File S12). We observed the C allele at rs12938658 to increase the expression of the *FLOT2* gene in pre-menarcheal females, while its effect in post-menarcheal females trended in the opposite direction (Fig. 5A). *FLOT2* encodes a caveolae-associated integral membrane protein, and its expression has been implicated in asthma risk via Transcriptome-Wide Association Study (TWAS)[47].

We also found interactions between genotype and puberty on expression of B cell (*IGKV1-27* in males, Fig. 5B and F) and T cell (*TRBV30* in females, Fig. 5D and H) antigen-recognition proteins, with genetic effects on gene expression disappearing by late puberty. We had previously identified genotype-specific effects of lymphocyte and neutrophil fractions as well as nightly asthma symptoms on expression of *TRBV30* in the same cohort[48], which suggests that this association may be mediated by changes in blood cell composition that accompany puberty[49]. Interestingly, a sex-biased genetic effect on expression of *TRBV30* had been previously reported[50] with a significant effect in adult male but not female spleen. Our findings suggest that the genetic effect on *TRBV30* expression may be active in pre-pubertal females but disappear due to different mechanisms of gene expression regulation in post-pubertal females.

In boys, we found an interaction between pubertal development stage and the genotype at rs116200096 on the expression of the *PGAP1* (Fig. 5C and G) gene. This gene encodes an enzyme catalyzing the inositol deacylation of glycosylphosphatidylinositol, which ensures proper folding of proteins involved in embryogenesis, neurogenesis, immunity, and fertilization[51]. Previously, we had found that expression of *PGAP1* was modulated by the interaction with genotype across many environments including cell composition and cell composition and psychosocial factors, such as the extent of self-disclosure of thoughts and feelings, and frequency of verbal arguments within the family[48]. Lastly, we observed a strong effect of the T allele at rs28630791 on the expression of the nearby *DSC1* gene only in mid-pubertal girls (Fig. 5E and I). *DSC1* encodes desmocollin 1 - a cell surface protein forming part of the desmosome that ensures tight junctions between epithelial cells[52] but is also highly expressed by macrophages where it may impair HDL biogenesis[53].

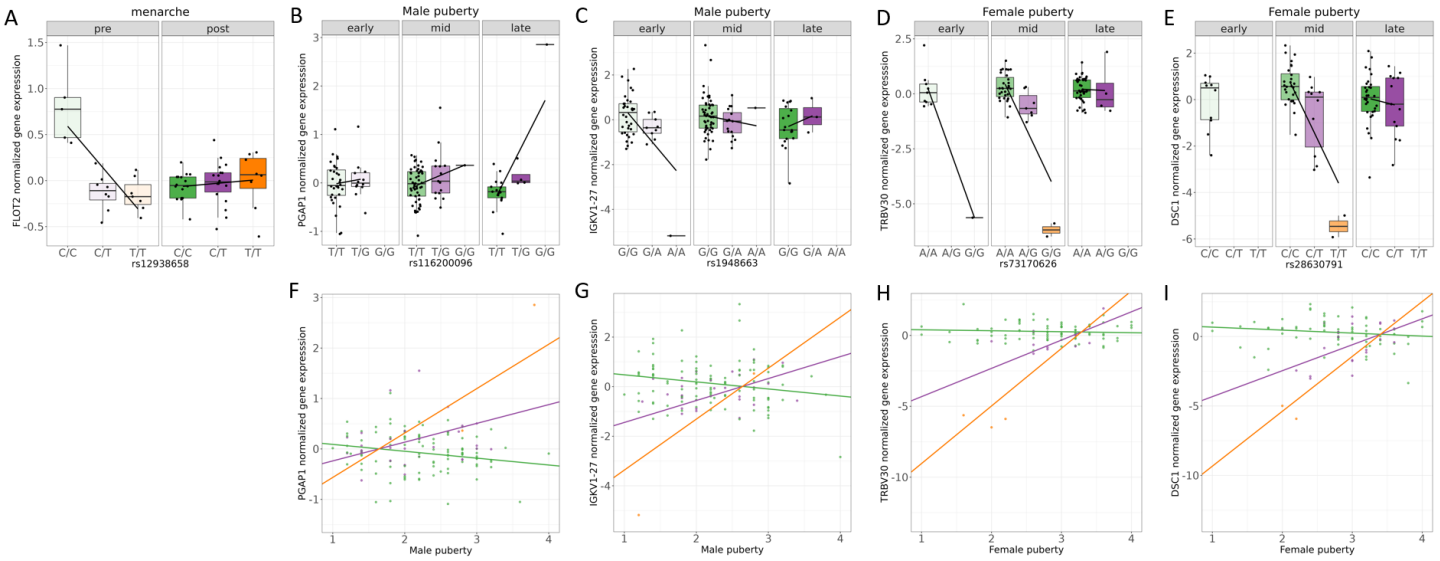

Fig. 5. Genetic variation interacts with pubertal status to influence gene expression. Boxplots (top row) and scatterplots (bottom row) depict gene expression across the three genotype classes. A – Menarche interacts with rs12938658 to affect expression of the *FLOT2* gene, B and F – Pubertal development stages interact with rs116200096 to affect expression of the *PGAP1* gene in boys, C and G – Pubertal development stages interact with rs1948663 affect expression of the *IGKV1-27* gene in boys, D and H – Pubertal development stages interact with rs73170626 to affect expression of the *TRBV30* gene in girls, E and I – Pubertal development stages interact with rs28630791 to affect expression of the *DSC1* gene in girls.

Changes in gene expression during puberty are associated with age at menarche and asthma risk

We have limited information on the genetic underpinnings of pubertal development. As a consequence, our knowledge of the molecular mechanisms accompanying developmental changes in the different body sites during puberty is limited. Our study identified changes in the expression of several genes in immune cells. To confirm that these gene expression changes are markers of pubertal development in independent cohorts, we considered GWAS data for age at menarche from the UK Biobank[46] jointly with gene expression data in whole blood from the GTEx study[54]. To investigate expression of which genes is associated with age at menarche, we performed a Transcriptome-wide association study (TWAS). TWAS allows the integration of gene expression data with GWAS data to identify plausible candidate genes driving trait variability in the population. We performed TWAS for age at menarche by incorporating whole blood eQTL data from GTEx(v8) with a GWAS of age at menarche from 370,000 individuals[46] using SPrediXcan. Among 7,174 tested genes, we detected a total of 108 genes associated with age at menarche at a genome-wide significant level (Bonferroni corrected p-value<0.05, Fig. 6A, File S13). These genes associated with age at menarche were mostly enriched in immune-related pathways or diseases: MHC protein complex and autoimmune thyroid disease or antigen processing and presentation, (10% FDR). Notably, we found that 27 of the genes associated with age at menarche in the TWAS were undergoing significant expression changes in peri-pubertal girls: 17 genes with expression changes associated with age, 10 genes whose expression differed between pre- and post-menarcheal girls, and one gene whose expression was higher when girls were at later pubertal stages (Fig. 6B, File S14). For example, the gene *EEFSEC* (Selenocysteine-Specific Elongation Factor) is upregulated as girls become older. Previous studies implicated this gene in endometrial cancer[55] and preterm birth[56], and a meta-analysis found a protective effect of late menarche on endometrial cancer risk[57].

Pubertal transition is associated with a shift in the sex distribution of asthma prevalence - in childhood boys have higher asthma prevalence, while in adulthood women have higher asthma prevalence and morbidity/severity[13,58]. A previous TWAS[47] associated expression of 1621 genes with asthma risk. We investigated whether pubertal development affected expression of these genes. In both sexes, we observed

gene expression changes over time during the peri-pubertal period for 124 genes associated with asthma risk. Additionally, expression of 56 genes associated with asthma risk differed between pre- and post-menarcheal girls, and four genes associated with asthma risk changed expression as girls progressed through puberty (Fig. 6C, File S15).

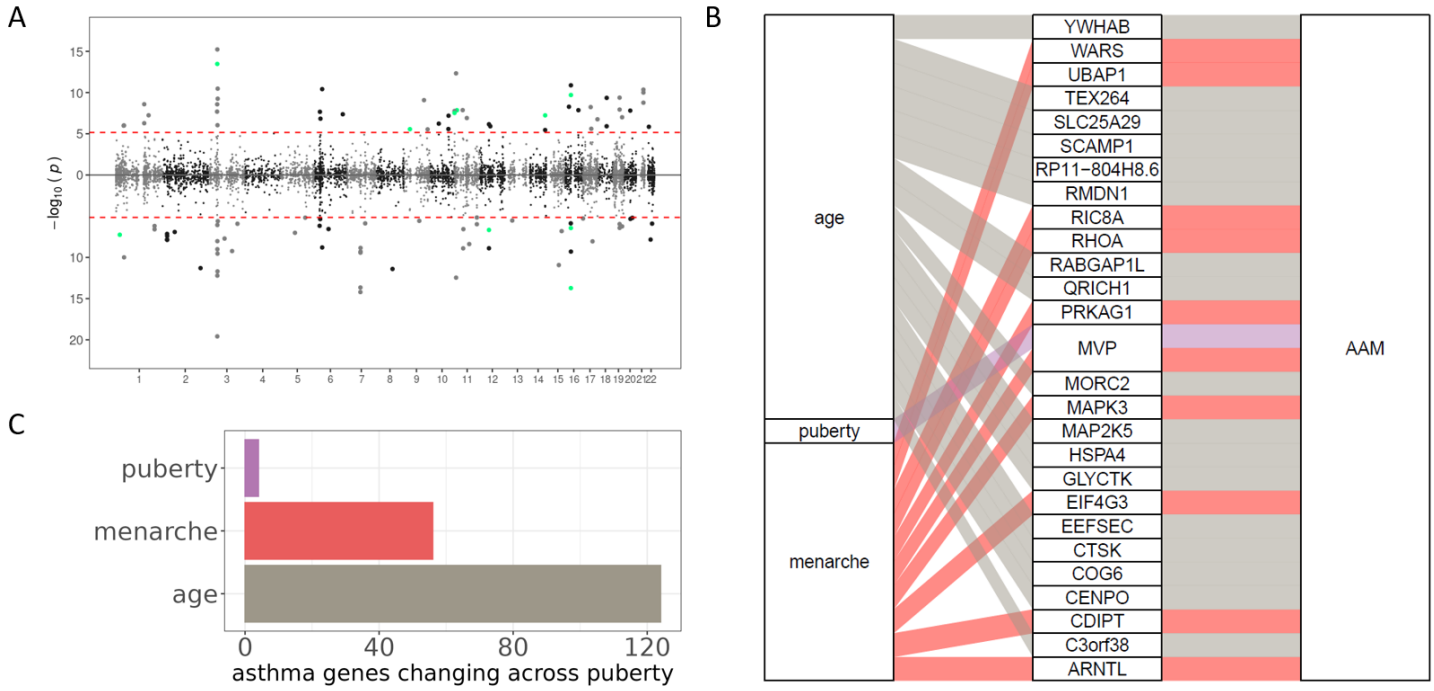

Fig. 6. Changes in gene expression during puberty are associated with age at menarche and asthma risk. A - Miami plot representing TWAS p-values for age at menarche across the genome. Each dot is a gene; red line represents Bonferroni significance threshold. Genes with higher expression in girls undergoing menarche later are plotted above the x axis, genes with higher expression in girls undergoing menarche earlier are plotted below the x axis. Color represents TWAS genes which were also significantly downregulated in post-menarcheal girls. B - Expression of genes associated with age at menarche differs across puberty in girls. Traits (left column) are connected to their respective differentially expressed genes (middle column) which are connected to the TWAS trait of age at menarche, C - Barplot of the number of genes previously implicated in asthma whose expression levels differed across pubertal traits (y axis).

Discussion

Previous genomic studies have focused on exploring DNA methylation changes between pre- and post-puberty[22–26] in peripheral blood immune cells. Yet, despite the clear evidence of epigenetic reprogramming during pubertal development, gene expression changes throughout puberty remain understudied. Here, we presented a comprehensive analysis of immune cell gene expression during the peri-pubertal period by analyzing the association of age, pubertal development stage and menarche status with gene expression in a sample of 251 children. Unlike previous approaches, we sampled the entire span of pubertal development thus capturing both early and late changes associated with puberty. Furthermore, we augmented our longitudinal study design with cross-sectional analyses to corroborate findings on gene expression changes accompanying the peri-pubertal period.

Changes to human physiology across the lifespan have been extensively studied for decades[59–64]. Yet, the effect of age on gene expression in humans has not been explored thoroughly. Particularly, we lack insight into the transcriptional changes across the peri-pubertal period. Here, we discovered widespread effects of age on immune cell gene expression in both males and females. To directly compare gene expression changes associated with age between the two sexes, we applied multivariate adaptive shrinkage, which revealed a remarkable consistency of age effect on gene expression between peri-pubertal boys and girls. Genes overexpressed when children were older were enriched for genes involved in biological processes related to viral gene expression and translation initiation, which may reflect the higher demand for protein anabolism when

children are growing most rapidly[37]. We observed a small but significant enrichment of genes differentially expressed with age in our peri-pubertal sample within genes differentially expressed with age in the elderly[36], which implies that some age effects on gene expression may be universal across the lifespan, or at least shared between puberty and senescence. However, age had the opposite effect on gene expression in youth and the elderly for 35% of those genes, suggesting that senescence may be associated with reversal of some of the same physiological processes leading to sexual maturity.

While the epigenetic changes accompanying pubertal development have been explored previously in several studies[22–26], the likely subsequent transcriptional changes are largely unknown. We explored the gene expression changes in boys and girls tracked longitudinally through puberty. However, we only detected significant longitudinal gene expression changes in girls but not boys, despite the larger sample size of boys in our study. In our sample girls and boys were of similar age. However, because females experience puberty earlier than males, girls in our sample were overrepresented at later pubertal stages while boys were overrepresented at earlier pubertal stages (Fig. 1C, t-test p-value=$3.54*10^{-6}$). Our results thus suggest that in immune cells, gene expression changes during later pubertal stages are more pronounced than those occurring during earlier pubertal development. Our sample size should be adequate to detect large changes at any stage if those were to occur. Thus an alternative explanation would be that the gene expression changes associated with pubertal development in boys were very subtle compared to those of girls during the time frame analyzed. Furthermore, our results are in line with previous findings of more genes differentially methylated between pre- and post-pubertal girls than boys[22,24,28]. Accordingly, our study adds to the existing evidence of a stronger effect of pubertal development on the functioning of the immune system in females compared to males. This reprogramming of the immune system by female sex hormones[65] may contribute to higher incidence of autoimmune diseases post-puberty in women than in men. However, we found only a small overlap between genes changing expression with pubertal development in girls and those previously reported to change methylation status between pre- and post-puberty in girls[22]. One biological explanation for this lack of enrichment may be that some genes are epigenetically poised for transcription even before puberty but are not expressed until activated by sex hormones. Alternatively, the small overlap between genes differentially expressed as girls advance through puberty and genes associated with regions differentially methylated in pre- and post-pubertal girls may suggest that some of the differentially expressed genes we discovered are specific to the pubertal transition period.

Expectedly, gene expression changes associated with pubertal development in girls were well correlated with age effects. This finding helps further validate the use of self-reported pubertal development scores in research settings. However, we discovered more genes differentially expressed with age than pubertal development. This could be because some of the girls in our sample had completed pubertal development and only differed by age, thus allowing us to discover additional genes for which age effects may be subtler but continue past puberty.

A recent study of DNA methylation changes associated with the five markers of pubertal development used in our study found that menarche status was the puberty marker associated with the highest number of regions differentially methylated[24]. In agreement with this finding, we found more genes whose expression differed between pre- and post-menarcheal girls compared to genes discovered to change expression as girls progress through pubertal development. This is likely because the onset of first menstruation is a less ambiguous marker of the shift into sexual maturity.

Previous studies of DNA methylation changes between pre- and post-puberty in girls found overrepresentation of genes involved in immune and inflammatory responses[22,23,25], reproductive hormone signaling[22], and other physiological changes related to growth and development[24]. Our results allowed a more fine-grained insight into the changes in the immune system that accompany sexual maturation of girls. We discovered an enrichment of processes related to innate immunity within genes more highly expressed in pre-menarcheal girls and enrichment for T cell receptor signaling in genes more highly expressed in post-menarcheal

females. The latter result corroborates a previous finding of T cell receptor signaling being overrepresented within genes near differentially methylated regions which were co-expressed in post-menarcheal girls[22]. Our results suggest a shift from innate to adaptive immunity past menarche. This is supported by the known effects of sex hormones on the immune system. For example, estrogen has been shown to promote adaptive immunity[66] and may promote or restrict innate immunity[67].

While sex-biased genetic effect on gene expression had been reported before[50], our study is the first to demonstrate that genetic regulation of gene expression varies during pubertal development. For example, the genetic effect on the expression of the B cell antigen recognizing protein *IGKV1-27* gene disappears by late puberty in males. We found hormone response elements for androgen and estrogen receptors near *IGKV1-27*, suggesting that the different hormonal context between early and late puberty may explain the observed interactions between genetic effect and puberty at this locus. In girls, we discovered that the genotype at rs73170626 is associated with the expression of *TRBV30* gene in early puberty, but this genetic effect disappears in late puberty. This gene encodes a T cell receptor. Interestingly, a sex-biased eQTL was previously reported for this gene, with genetic effects on gene expression present only in adult males but not females[50]. This implicates puberty in origins of sex-biased genetic effects on gene expression, which may translate into disease risk in women later in life.

A previous TWAS identified 205 genes whose expression in neural tissues is associated with age at menarche[46]. Using a similar TWAS approach here we demonstrated an association between blood expression levels of 108 genes and age at menarche. For 31 of these genes, our results confirmed previous observations for causal or pleiotropic effects between blood gene expression and age at menarche that were based on Mendelian randomization. This overlap represents a significant enrichment over chance (OR=195, p-value<2.2e-16)[46]. We demonstrated that expression of 27 genes associated with age at menarche changed across puberty, including a gene implicated in endometrial cancer risk[55], which may help explain the link between earlier age at menarche and increased endometrial cancer risk[57]. Additionally, we demonstrated that pubertal transition was associated with changes in expression of genes associated with asthma risk. Further investigations of genes that are both associated with asthma risk and differentially expressed between pre- and post-menarche females may provide clues to understanding increased asthma risk in females post puberty.

Our study is the first genome-wide characterization of patterns of gene expression changes in peri-pubertal boys and girls, and complements the existing body of literature on the epigenetic reprogramming of immune cells during puberty. Our results also add to the literature on lifespan changes in gene expression by focusing on early adolescence. Overall, our findings suggest a shift from a more innate to a more adaptive immune response past puberty in girls, which might be a mechanism for the higher incidence of autoimmune disease in adult women compared to men. Additionally, our study is the first to demonstrate that the associations between gene expression and pubertal development can be modified by genetic variation, which may contribute to inter-individual variability in incidence and severity of diseases with peri-pubertal onset.

Methods

Study participants

Participants were drawn from the Asthma in the Lives of Families Today study (ALOFT; samples collected November 2010-July 2018). This multi-ethnic cohort was established to investigate the effects of psychosocial experiences on asthma symptoms in youth living in the Detroit Metropolitan Area. The total sample included in the current study consisted of 251 participants (103 girls and 148 boys) aged 10-17 at recruitment for whom we successfully collected RNA-seq and genome-wide genotype data (cohort described previously[48]). The ALOFT study included three measurements at annual intervals, but only the first two sampled time points were included

for individuals with more than two longitudinal data points in the ALOFT study. This strategy allowed us to maximize the sample size for the longitudinal analysis. For 65% of the participants, we were able to collect longitudinal gene expression data from at least two timepoints. Average time interval between the two time points was 1.27 years (SD=0.46 years).

Gene expression data

Leukocyte gene expression data for all timepoints and genotypes imputed to the whole genome were obtained as described previously[48]. In brief, RNA was extracted using LeukoLOCK (Thermo Fisher) and only samples with RNA Integrity Number (RIN) of at least 6 measured on Agilent Bioanalyzer were included. Library preparation was done in batches of up to 96 samples following standard Illumina TruSeq Stranded mRNA protocol, with longitudinal samples from the same individual processed together in the same batch. Samples were sequenced on the Illumina NextSeq500 Desktop Sequencer to obtain 75bp paired-end reads. A total of 21 million (M) to 76M reads were collected per sample, with a mean of 41M reads/sample. The RNA-seq data were aligned to the human genome version GRCh37 using HISAT2[68] and counted across all genes using HTSeq. Samples with excess PCR duplicate rate (>60%) were excluded and sample identity was confirmed by comparing the genotypes obtained from the RNA-seq data with those obtained from the DNA samples. For all gene expression analyses, genes on sex chromosomes and genes with expression below 6 reads or 0.1 counts per million in at least 20% of samples were dropped. Data for the first time point for each individual and longitudinal data for 120 of the participants have been previously published[48] and are available through dbGaP accession #phs002182.v1.p1. The additional new data included in this study will be deposited in a new version of the same repository.

Genotype data

VCF files were generated from low-coverage (~0.4×) whole-genome sequencing of one blood-derived DNA sample per individual and imputed to 37.5M variants using the 1000 Genomes database by Gencove (New York, NY). These data are accessible at dbGaP accession #phs002182.v1.p1.

Puberty measurements

Pubertal development was measured at each timepoint prior to the blood draw for gene expression measurement via a sex-specific self-report questionnaire[29]. The questions and possible responses can be found in Table S2. The overall pubertal development score is a mean of the responses to specific questions 1-5, given that at least four valid responses are provided. Pubertal development stages are analogous to the Tanner stages[39,40] and were treated as a continuous variable in all analyses. For female participants we also considered self-reported menarche status (pre-menarche/post-menarche).

Differential Gene Expression Analysis

Longitudinal analysis: We used DESeq2 to test for gene expression changes over time and across puberty stages in the longitudinal sample. Specifically, we used the following models:

gene expression ~ fraction of reads mapping to exons + fraction of high-quality reads + RIN + participant ID + age

gene expression ~ fraction of reads mapping to exons + fraction of high-quality reads + RIN + participant ID + pubertal stage

In the DESeq model, we corrected for fraction of reads mapping to exons, fraction of high-quality reads (after removing PCR duplicates), RNA Integrity Number (RIN), and individual effect.

These analyses were run in each sex separately. A subset of 24 participants reported a lower pubertal development score in the second timepoint compared to baseline. To account for these mis-reported data, we corrected the second timepoint to match the pubertal development score measured for the same individual at baseline.
Significance threshold for differentially expressed genes was set at 10% FDR using Benjamini-Hochberg correction method.

Cross-sectional analysis: We used DESeq2 to test for gene expression associated with sex, age, puberty, and menarche status in each sex separately. To account for possible confounding factors, we corrected for library preparation batch, fraction of reads mapping to exons, fraction of high-quality reads, RNA Integrity Number (RIN), and the top three genotype PCs.

gene expression ~ batch + fraction of reads mapping to exons + fraction of high-quality reads + RIN + genotype PCs + age + **sex**

gene expression ~ batch + fraction of reads mapping to exons + fraction of high-quality reads + RIN + genotype PCs + **age**

gene expression ~ batch + fraction of reads mapping to exons + fraction of high-quality reads + RIN + genotype PCs + **pubertal stage**

gene expression ~ batch + fraction of reads mapping to exons + fraction of high-quality reads + RIN + genotype PCs + **menarche**

Multivariate adaptive shrinkage (mash): To investigate gene expression changes associated with age in both sexes, we used the multivariate adaptive shrinkage (mash) method[31] using the mashr package v.0.2.40 in R v4.0. As input, we provided the log-fold change and corresponding standard error from the DESeq analysis in each sex, subset for genes tested in both sexes (18341 genes). We considered genes with local false sign rate (LFSR)<0.1 to be differentially expressed.

Hormone receptor binding site analysis

To test for enrichment of estrogen receptor binding sites near differentially expressed genes, we downloaded position weight matrices (PWM) for ESR1 from the JASPAR Vertebrate Core Motifs database version 2020[69]. We counted the number of genes with at least one strongly-predicted (log(2) odds over genomic background>10) binding motif within 10kb of the transcription start site within the set of differentially expressed genes and the set of all tested genes for each binding motif separately. We tested for enrichment using Fisher's exact test and used P-value<0.05 as threshold for significance.

Gene Ontology Enrichment Analysis

We used the clusterProfiler[70] package for R to run Gene Ontology enrichment analysis for the differentially expressed genes for each variable, correcting for the background of all expressed genes in each sex. The significance threshold was set at 10% FDR, and we only considered categories that included at least three genes from the test set.

To investigate pathway-level changes between pre- and post-menarcheal females, we calculated pathway scores of interest as follows: (1) calculated residual expression values by regressing the effects of confounders used in differential gene expression analysis; (2) extracted the subset of differentially expressed genes between pre- and post-menarcheal girls and belonging to a specific GO term or KEGG pathway; (3) scaled mean-centered gene expression for each gene across individuals; (4) calculated the average values as a specific pathway score across genes for each individual.

Interaction eQTL mapping

To prepare the data for eQTL mapping, we quantile-normalized the gene expression data for the cross-sectional sample of 251 participants using the voom(method="quantile") function in *limma* (Law et al. 2014). To remove the effect of technical confounders, we regressed the effects of: RIN, fraction of reads mapping to exons, fraction of high-quality reads, data collection batch, library preparation batch, genotype PC1, genotype PC2, genotype PC3, age, weight, and height. We first performed eQTL mapping in both sexes combined to identify genetic variants associated with gene expression levels. To do this, we used FastQTL[71] with adaptive permutations (1000-10,000). For each gene, we tested all genetic variants within 1 Mb of the transcription start site (TSS) and with cohort MAF > 0.1. We defined significant eGenes atFDR<10%. For interaction eQTL mapping, we tested the SNP with the lowest p-value for each eGene. We quantile-normalized the gene expression data in each sex separately using the voom(method="quantile") function in *limma* and regressed the effects of: RIN, fraction of reads mapping to exons, fraction of high quality reads, library preparation batch, genotype PC1, genotype PC2, genotype PC3. To identify genetic variants that interacted with puberty stage, age or menarche status to influence expression of the nearby gene, we tested the interaction between genotype dosage and each of these variables while controlling for the marginal effects using the following models in each sex separately:

Normalized gene expression ~ genotype dosage + age + **genotype dosage*age**

Normalized gene expression ~ genotype dosage + pubertal stage + **genotype dosage*pubertal stage**

Normalized gene expression ~ genotype dosage + menarche + **genotype dosage*menarche**

We used 1000 permutations to correct the p-values as described previously[48] before applying Storey's q-value method to control for FDR.

Transcriptome-wide association study of age at menarche

We performed TWAS for age at menarche (AAM) by running the module SPrediXcan in the MetaXcan software[72]. To perform TWAS, we first downloaded GWAS summary statistics of AAM from the ReproGen consortium (https://www.reprogen.org/)[46], which were generated based on data from 329,345 women from 40 studies. This data set contains over 10 million genetic markers. After transferring marker name from "chr:position" to "rs id" using annovar pipeline[73], over 4 million markers remained in the following TWAS analysis. As input for SPrediXcan, we also provided covariances file (single-tissue LD reference files) and gene expression model in the blood tissue from the GTEx (V8) project[54], available from (https://predictdb.org/). We used gene expression models trained in the blood tissue based on the Elastic Net model from the GTEx (V8) project. To validate our results, we used Fisher's exact test to consider the enrichment of genes associated with age at menarche through TWAS within genes previously implicated in age at menarche through Mendelian randomization[46].

Acknowledgments:

We thank Carole Ober for helpful comments on an earlier version of this manuscript and members of the Luca, Pique-Regi and Zilioli groups for helpful discussions.

Bibliography:

1. Rey, R. A., Campo, S. M., Ropelato, M. G. & Bergadá, I. Hormonal Changes in Childhood and Puberty. in *Puberty: Physiology and Abnormalities* (eds. Kumanov, P. & Agarwal, A.) 23–37 (Springer International Publishing, 2016).
2. Growth Charts. https://www.cdc.gov/growthcharts (2019).
3. Mendle, J., Beltz, A. M., Carter, R. & Dorn, L. D. Understanding Puberty and Its Measurement: Ideas for Research in a New Generation. *J. Res. Adolesc.* **29**, 82–95 (2019).
4. Rogol, A. D., Roemmich, J. N. & Clark, P. A. Growth at puberty. *J. Adolesc. Health* **31**, 192–200 (2002).
5. Halmi, K. A., Casper, R. C., Eckert, E. D., Goldberg, S. C. & Davis, J. M. Unique features associated with age of onset of anorexia nervosa. *Psychiatry Res.* **1**, 209–215 (1979).
6. Raab, J. *et al.* Continuous rise of insulin resistance before and after the onset of puberty in children at increased risk for type 1 diabetes - a cross-sectional analysis. *Diabetes. Metab. Res. Rev.* **29**, 631–635 (2013).
7. Keselman, A. & Heller, N. Estrogen Signaling Modulates Allergic Inflammation and Contributes to Sex Differences in Asthma. *Front. Immunol.* **6**, 568 (2015).
8. Charania, S. N. *et al.* Bullying Victimization and Perpetration Among US Children with and Without Tourette Syndrome. *Journal of Developmental & Behavioral Pediatrics* vol. Publish Ahead of Print (2021).
9. Huang, J. L., Yao, T. C. & See, L. C. Prevalence of pediatric systemic lupus erythematosus and juvenile chronic arthritis in a Chinese population: a nation-wide prospective population-based study in Taiwan. *Clin. Exp. Rheumatol.* **22**, 776–780 (2004).
10. Ceovic, R. *et al.* Psoriasis: female skin changes in various hormonal stages throughout life--puberty, pregnancy, and menopause. *Biomed Res. Int.* **2013**, 571912 (2013).
11. Waubant, E. Effect of puberty on multiple sclerosis risk and course. *Mult. Scler.* **24**, 32–35 (2018).
12. Ingram, J. R. The epidemiology of hidradenitis suppurativa. *Br. J. Dermatol.* **183**, 990–998 (2020).

13. Vink, N. M., Postma, D. S., Schouten, J. P., Rosmalen, J. G. M. & Boezen, H. M. Gender differences in asthma development and remission during transition through puberty: the TRacking Adolescents' Individual Lives Survey (TRAILS) study. *J. Allergy Clin. Immunol.* **126**, 498–504.e1–6 (2010).
14. Day, F. R., Perry, J. R. B. & Ong, K. K. Genetic Regulation of Puberty Timing in Humans. *Neuroendocrinology* **102**, 247–255 (2015).
15. Minelli, C. *et al.* Age at puberty and risk of asthma: A Mendelian randomisation study. *PLoS Med.* **15**, e1002634 (2018).
16. Lakshman, R. *et al.* Early age at menarche associated with cardiovascular disease and mortality. *J. Clin. Endocrinol. Metab.* **94**, 4953–4960 (2009).
17. Elks, C. E. *et al.* Age at menarche and type 2 diabetes risk: the EPIC-InterAct study. *Diabetes Care* **36**, 3526–3534 (2013).
18. Eliasson, O., Scherzer, H. H. & DeGraff, A. C., Jr. Morbidity in asthma in relation to the menstrual cycle. *J. Allergy Clin. Immunol.* **77**, 87–94 (1986).
19. Vrieze, A., Postma, D. S. & Kerstjens, H. A. m. Perimenstrual asthma: a syndrome without known cause or cure. *J. Allergy Clin. Immunol.* **112**, 271–282 (2003).
20. Arathimos, R. *et al.* Genetic and observational evidence supports a causal role of sex hormones on the development of asthma. *Thorax* **74**, 633–642 (2019).
21. Almstrup, K. *et al.* Pubertal development in healthy children is mirrored by DNA methylation patterns in peripheral blood. *Sci. Rep.* **6**, 28657 (2016).
22. Thompson, E. E. *et al.* Global DNA methylation changes spanning puberty are near predicted estrogen-responsive genes and enriched for genes involved in endocrine and immune processes. *Clin. Epigenetics* **10**, 62 (2018).
23. Han, L. *et al.* Changes in DNA methylation from pre- to post-adolescence are associated with pubertal exposures. *Clin. Epigenetics* **11**, 176 (2019).
24. Chen, S. *et al.* Age at onset of different pubertal signs in boys and girls and differential DNA methylation at age 10 and 18 years: an epigenome-wide follow-up study. *Hum Reprod Open* **2020**, hoaa006 (2020).
25. Moore, S. R. *et al.* Distinctions between sex and time in patterns of DNA methylation across puberty. *BMC*

Genomics **21**, 389 (2020).

26. De Vito, R. *et al.* Differentially methylated regions and methylation QTLs for teen depression and early puberty in the Fragile Families Child Wellbeing Study. *bioRxiv* 2021.05.20.444959 (2021) doi:10.1101/2021.05.20.444959.
27. Horvath, S. & Raj, K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat. Rev. Genet.* **19**, 371–384 (2018).
28. Almstrup, K. *et al.* Erratum: Pubertal development in healthy children is mirrored by DNA methylation patterns in peripheral blood. *Sci. Rep.* **6**, 30664 (2016).
29. Carskadon, M. A. & Acebo, C. A self-administered rating scale for pubertal development. *J. Adolesc. Health* **14**, 190–195 (1993).
30. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
31. Urbut, S. M., Wang, G., Carbonetto, P. & Stephens, M. Flexible statistical methods for estimating and testing effects in genomic studies with multiple conditions. *Nat. Genet.* **51**, 187–195 (2019).
32. Gay, C. M., Zygmunt, T. & Torres-Vázquez, J. Diverse functions for the semaphorin receptor PlexinD1 in development and disease. *Dev. Biol.* **349**, 1–19 (2011).
33. White, B. J., Rogol, A. D., Brown, K. S., Lieblich, J. M. & Rosen, S. W. The syndrome of anosmia with hypogonadotropic hypogonadism: a genetic study of 18 new families and a review. *Am. J. Med. Genet.* **15**, 417–435 (1983).
34. Tornberg, J. *et al.* Heparan sulfate 6-O-sulfotransferase 1, a gene involved in extracellular sugar modifications, is mutated in patients with idiopathic hypogonadotrophic hypogonadism. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 11524–11529 (2011).
35. Miraoui, H. *et al.* Mutations in FGF17, IL17RD, DUSP6, SPRY4, and FLRT3 are identified in individuals with congenital hypogonadotropic hypogonadism. *Am. J. Hum. Genet.* **92**, 725–743 (2013).
36. Balliu, B. *et al.* Genetic regulation of gene expression and splicing during a 10-year period of human aging. *Genome Biol.* **20**, 230 (2019).
37. Mauras, N., Rogol, A. D., Haymond, M. W. & Veldhuis, J. D. Sex steroids, growth hormone, insulin-like

growth factor-1: neuroendocrine and metabolic regulation in puberty. *Horm. Res.* **45**, 74–80 (1996).

38. Abreu, A. P. & Kaiser, U. B. Pubertal development and regulation. *Lancet Diabetes Endocrinol* **4**, 254–264 (2016).
39. Marshall, W. A. & Tanner, J. M. Variations in pattern of pubertal changes in girls. *Archives of Disease in Childhood* vol. 44 291–303 (1969).
40. Marshall, W. A. & Tanner, J. M. Variations in the pattern of pubertal changes in boys. *Arch. Dis. Child.* **45**, 13–23 (1970).
41. Ding, N. *et al.* Human serum N-glycan profiles are age and sex dependent. *Age Ageing* **40**, 568–575 (2011).
42. Pucic, M. *et al.* Changes in plasma and IgG N-glycome during childhood and adolescence. *Glycobiology* **22**, 975–982 (2012).
43. Loomba-Albrecht, L. A. & Styne, D. M. Effect of puberty on body composition. *Curr. Opin. Endocrinol. Diabetes Obes.* **16**, 10–15 (2009).
44. Euling, S. Y., Selevan, S. G., Pescovitz, O. H. & Skakkebaek, N. E. Role of environmental factors in the timing of puberty. *Pediatrics* **121 Suppl 3**, S167–71 (2008).
45. Day, F. R. *et al.* Shared genetic aetiology of puberty timing between sexes and with health-related outcomes. *Nat. Commun.* **6**, 8842 (2015).
46. Day, F. R. *et al.* Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. *Nat. Genet.* **49**, 834–841 (2017).
47. Zhang, Y. *et al.* PTWAS: investigating tissue-relevant causal molecular mechanisms of complex traits using probabilistic TWAS analysis. *Genome Biol.* **21**, 232 (2020).
48. Resztak, J. A. *et al.* Psychosocial experiences modulate asthma-associated genes through gene-environment interactions. *Elife* **10**, (2021).
49. Pérez-de-Heredia, F. *et al.* Influence of sex, age, pubertal maturation and body mass index on circulating white blood cell counts in healthy European adolescents—the HELENA study. *Eur. J. Pediatr.* **174**, 999–1014 (2015).
50. Oliva, M. *et al.* The impact of sex on gene expression across human tissues. *Science* **369**, (2020).

51. Liu, Y.-S. & Fujita, M. Mammalian GPI-anchor modifications and the enzymes involved. *Biochem. Soc. Trans.* **48**, 1129–1138 (2020).

52. Donetti, E. *et al.* Desmocollin 1 and desmoglein 1 expression in human epidermis and keratinizing oral mucosa: a comparative immunohistochemical and molecular study. *Arch. Dermatol. Res.* **297**, 31–38 (2005).

53. Choi, H. Y. *et al.* Desmocollin 1 is abundantly expressed in atherosclerosis and impairs high-density lipoprotein biogenesis. *Eur. Heart J.* **39**, 1194–1202 (2018).

54. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).

55. Kho, P. F. *et al.* Multi-tissue transcriptome-wide association study identifies eight candidate genes and tissue-specific gene expression underlying endometrial cancer susceptibility. *Commun Biol* **4**, 1211 (2021).

56. Zhang, G. *et al.* Genetic Associations with Gestational Duration and Spontaneous Preterm Birth. *N. Engl. J. Med.* **377**, 1156–1167 (2017).

57. Gong, T.-T., Wang, Y.-L. & Ma, X.-X. Age at menarche and endometrial cancer risk: a dose-response meta-analysis of prospective studies. *Sci. Rep.* **5**, 14051 (2015).

58. Shah, R. & Newcomb, D. C. Sex Bias in Asthma Prevalence and Pathogenesis. *Front. Immunol.* **9**, 2997 (2018).

59. Spencer, R. P. VARIATION OF INTESTINAL ACTIVITY WITH AGE: A REVIEW. *Yale J. Biol. Med.* **37**, 105–129 (1964).

60. Somerville, D. A. The effect of age on the normal bacterial flora of the skin. *Br. J. Dermatol.* **81**, Suppl 1:14+ (1969).

61. Kay, M. M. An overview of immune aging. *Mech. Ageing Dev.* **9**, 39–59 (1979).

62. Ludwig, F. C. & Smoke, M. E. The measurement of biological age. *Exp. Aging Res.* **6**, 497–522 (1980).

63. Reaven, G. M. & Reaven, E. P. Effects of age on various aspects of glucose and insulin metabolism. *Mol. Cell. Biochem.* **31**, 37–47 (1980).

64. Dutta, S. & Sengupta, P. Men and mice: Relating their ages. *Life Sci.* **152**, 244–248 (2016).

65. Shepherd, R., Cheung, A. S., Pang, K., Saffery, R. & Novakovic, B. Sexual Dimorphism in Innate Immunity: The Role of Sex Hormones and Epigenetics. *Front. Immunol.* **11**, 604000 (2020).

66. Gubbels Bupp, M. R. Sex, the aging immune system, and chronic disease. *Cell. Immunol.* **294**, 102–110 (2015).

67. Kovats, S. Estrogen receptors regulate innate immune cells and signaling pathways. *Cell. Immunol.* **294**, 63–69 (2015).

68. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).

69. Fornes, O. *et al.* JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **48**, D87–D92 (2020).

70. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).

71. Ongen, H., Buil, A., Brown, A. A., Dermitzakis, E. T. & Delaneau, O. Fast and efficient QTL mapper for thousands of molecular phenotypes. *Bioinformatics* **32**, 1479–1485 (2016).

72. Barbeira, A. N. *et al.* Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat. Commun.* **9**, 1825 (2018).

73. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).

Supplementary Figures

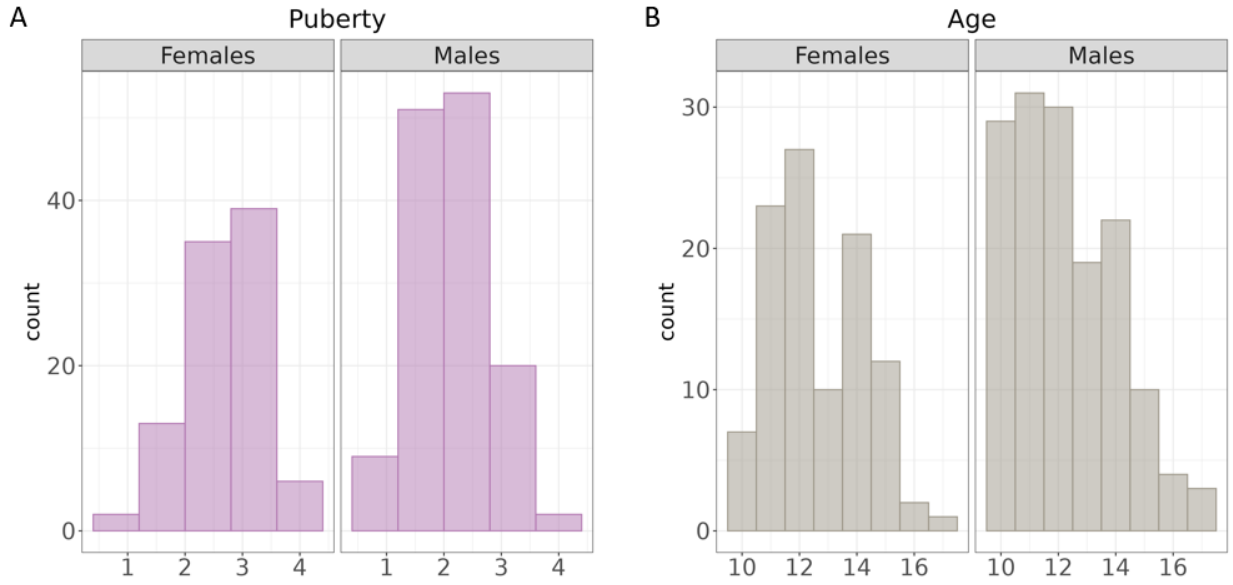

Fig. S1. Distribution of puberty and age in the cross-sectional sample. A – Histogram representing the distribution of pubertal development stages in girls (left panel) and boys (right panel), B - Histogram representing the distribution of age in years in girls (left panel) and boys (right panel).

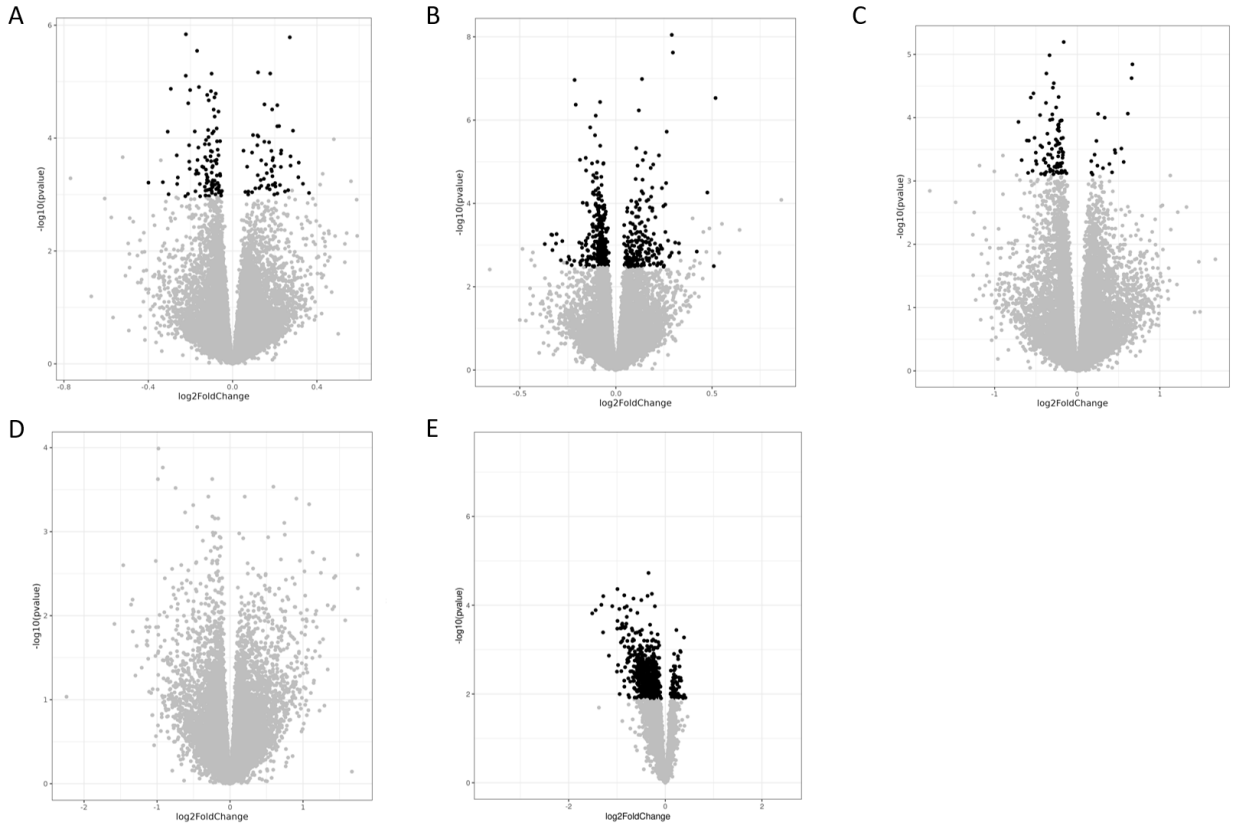

Fig. S2. Volcano plots of differential gene expression (DGE) analysis effect size (log2 fold-change, x axis) and p-values (y axis). A – longitudinal DGE analysis across time in girls, B – longitudinal DGE analysis across time in boys, C – longitudinal DGE analysis across puberty stages in girls, D – longitudinal DGE analysis across puberty stages in boys, E – cross-sectional DGE analysis of pre- and post-menarche in girls. Black denotes significant values (10% FDR), grey denotes not significant.

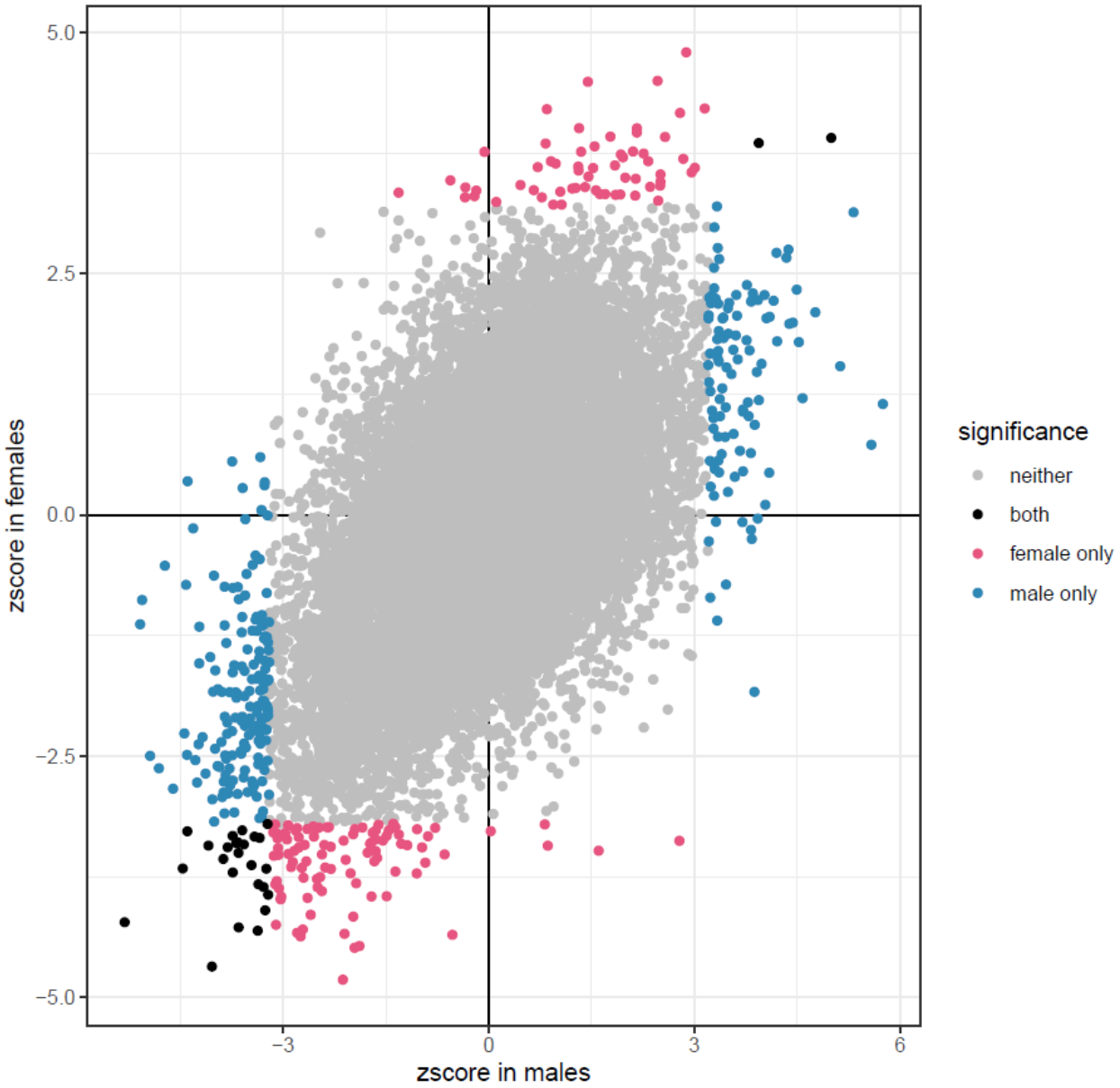

Fig. S3. Age effects on gene expression are consistent between peri-pubertal girls and boys. Scatterplot of normalized effect sizes (z-scores) from longitudinal differential gene expression analysis across pubertal stages in boys (x axis) and girls (y axis). Color denotes significance in each analysis as indicated by the legend.

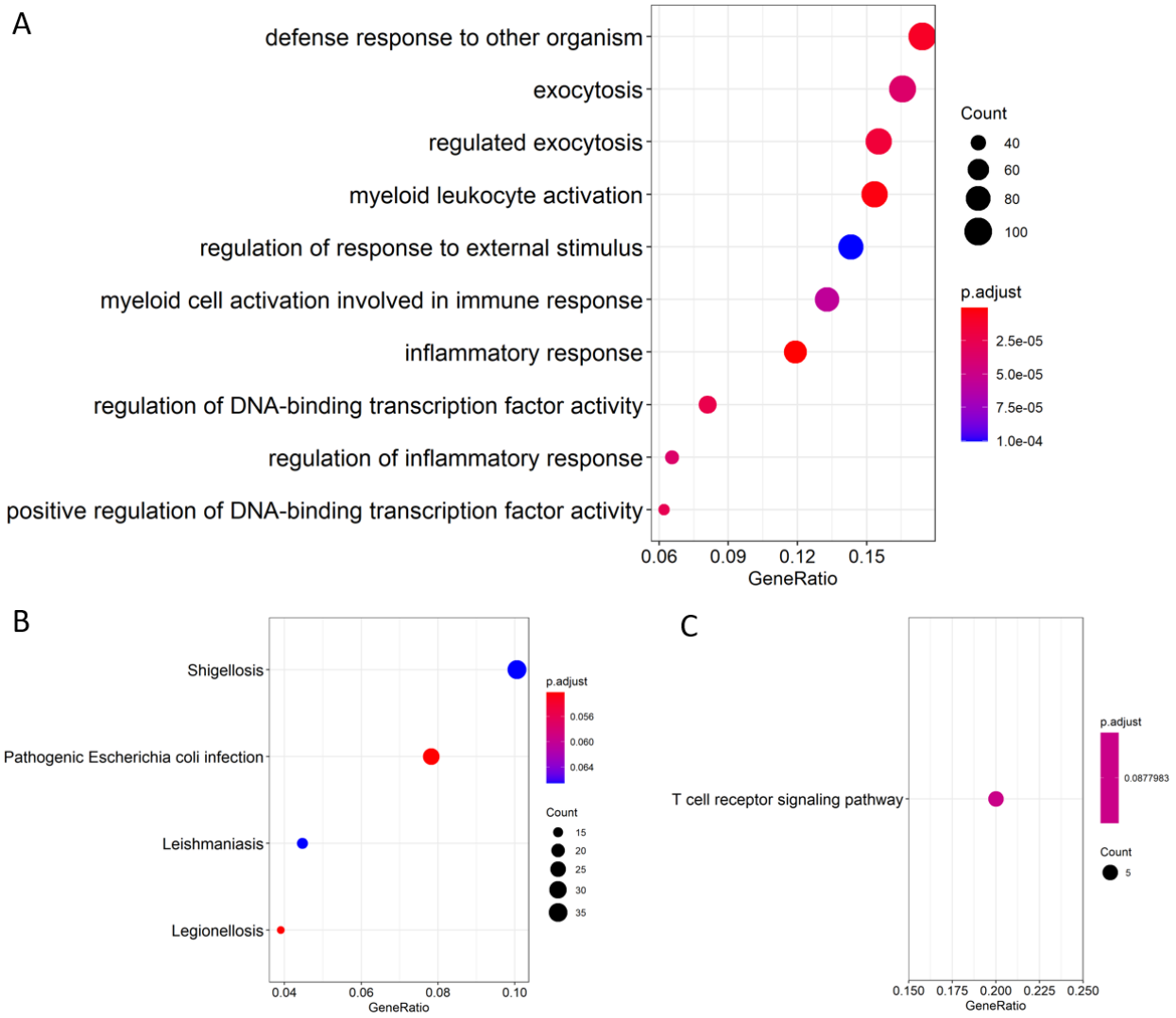

Fig. S4. Biological processes and pathway enriched within genes whose expression differs between pre- and post-pubertal girls. A - dotplot represents enrichment of Gene Ontology biological processes within genes whose expression is lower in girls post menarche, B – dotplot represents enrichment of KEGG pathways within genes whose expression is lower in girls post menarche, C – dotplot represents enrichment of KEGG pathways within genes whose expression is higher in girls post menarche.

Supplementary Tables

Table S1. Sample size. Reported are the sample size (N) and age range for each analysis, and age range for first timepoint (T0) and second timepoint (T1) of the longitudinal analysis).

Age					
Subset	**N (Cross-Sectional)**	**Age Range**	**N (Longitudinal)**	**Age Range T0**	**Age Range T1**
Total	251	10-17 yrs	163	10-15 yrs	11-16 yrs
Males	148	10-17 yrs	97	10-15 yrs	11-16 yrs
Females	103	10-17 yrs	66	10-15 yrs	11-16 yrs
Puberty					
Subset	**N (Cross-Sectional)**	**Age Range**	**N (Longitudinal)**	**Age Range T0**	**Age Range T1**
Total	240	10-17 yrs	142	10-15 yrs	11-16 yrs
Males	141	10-17 yrs	85	10-15 yrs	11-16 yrs
Females	99	10-17 yrs	57	10-15 yrs	11-16 yrs
Menarche					
Subset	**N (Cross-Sectional)**	**Age Range**			
Females	66	10-16 yrs			

Table S2. Pubertal Development Questionnaire.

Question	Answers	Numerical value
Girls' questionnaire		
Would you say your growth in height (getting taller)…	Has not yet began to spurt ('spurt' means more growth than usual)	1
	Has barely started to spurt	2
	Has definitely started to spurt, but has not finished	3
	Seems complete (you're about as tall as you're going to get)	4
How about the growth of your body hair? ("Body hair" means hair any place other than your head, such as under your arms).	Has not started growing	1
	Has barely started growing	2
	Has definitely started growing, but has not finished	3
	Seems complete (you have as much body hair as you're going to get)	4
Have you noticed any skin changes, especially pimples?	Skin has not yet started showing changes	1
	Skin has barely started showing changes	2
	Skin changes have definitely started but are not finished	3
	Skin changes seem complete	4
Have you noticed that your breasts have begun to grow?	Have not yet started growing	1
	Have barely started growing	2
	Breast growth has definitely started, but is not finished	3
	Breast growth seems completed	4
Have you begun to menstruate? ("menstruate" means to get your period)	Yes	4
	No	1
Boys' questionnaire		
Would you say your growth in height (getting taller)…	Has not yet began to spurt ('spurt' means more growth than usual)	1
	Has barely started to spurt	2
	Has definitely started to spurt, but has not finished	3
	Seems complete (you're about as tall as you're going to get)	4
How about the growth of your body hair? ("Body hair" means hair any place other than your head, such as under your arms).	Has not started growing	1
	Has barely started growing	2
	Has definitely started growing, but has not finished	3
	Seems complete (you have as much body hair as you're going to get)	4
Have you noticed any skin changes, especially pimples?	Skin has not yet started showing changes	1
	Skin has barely started showing changes	2
	Skin changes have definitely started but are not finished	3
	Skin changes seem complete	4
Have you noticed a deepening of your voice?	Voice has not yet started changing	1
	Voice has barely started changing	2

	Voice has definitely started changing, but is not finished	3
	Voice change seems complete	4
Have you started to grow hair on your face?	Facial hair has not started growing	1
	Facial hair has barely started growing	2
	Facial hair growth has definitely started but is not finished	3
	Facial hair growth seems complete	4

Supplementary Files

File S1. Results of longitudinal differential gene expression analysis across time in girls.
http://genome.grid.wayne.edu/puberty/cage1_Female_stats_longit.txt.gz

File S2. Results of longitudinal differential gene expression analysis across time in boys.
http://genome.grid.wayne.edu/puberty/cage1_Male_stats_longit.txt.gz

File S3. Results of cross-sectional differential gene expression analysis across age in girls.
http://genome.grid.wayne.edu/puberty/cage1_Female_stats_cs.txt.gz

File S4. Results of cross-sectional differential gene expression analysis across age in boys.
http://genome.grid.wayne.edu/puberty/cage1_Male_stats_cs.txt.gz

File S5. Significance of longitudinal multivariate adaptive shrinkage analysis across time in both sexes (LFSR).
http://genome.grid.wayne.edu/puberty/cage1_LFSR.txt

File S6. Effect size estimates from longitudinal multivariate adaptive shrinkage analysis across time in both sexes.
http://genome.grid.wayne.edu/puberty/cage1_beta.txt

File S7. Results of longitudinal differential gene expression analysis across puberty stages in girls.
http://genome.grid.wayne.edu/puberty/cgpd_stats_longit.txt.gz

File S8. Results of longitudinal differential gene expression analysis across puberty stages in boys.
http://genome.grid.wayne.edu/puberty/cbpd_stats_longit.txt.gz

File S9. Results of cross-sectional differential gene expression analysis across puberty stages in girls.
http://genome.grid.wayne.edu/puberty/cgpd_stats_cs.txt.gz

File S10. Results of cross-sectional differential gene expression analysis across puberty stages in boys.
http://genome.grid.wayne.edu/puberty/cbpd_stats_cs.txt.gz

File S11. Results of cross-sectional differential gene expression of pre and post-menarche in girls.
http://genome.grid.wayne.edu/puberty/cgpd5_stats_cs.txt.gz

File S12. Results of cis interaction eQTL mapping.
http://genome.grid.wayne.edu/puberty/GxPuberty_all.txt.gz

File S13. Results of Transcriptome-Wide Association Study of age at menarche.
http://genome.grid.wayne.edu/puberty/Asthma-TWAS_Blood.csv

File S14. Overlap of genes associated with age at menarche via TWAS and differentially expressed genes.
http://genome.grid.wayne.edu/puberty/AAM-TWAS_DEG_overlap.txt

File S15. Overlap of genes associated with asthma via TWAS (Zhang et al, 2019) and differentially expressed genes.
http://genome.grid.wayne.edu/puberty/Astma-TWAS_DEG_overlap.txt