Genetics of PlGF plasma levels highlights a role of its receptors and supports the link between angiogenesis and immunity

Daniela Ruggiero1,2*, Teresa Nutile1, Stefania Nappo3, Alfonsina Tirozzi2, Celine Bellenguez4, Anne-Louise Leutenegger5,6 & Marina Ciullo1,2*

Placental growth factor (PlGF) is a member of the vascular endothelial growth factor family and is involved in bone marrow-derived cell activation, endothelial stimulation and pathological angiogenesis. High levels of PlGF have been observed in several pathological conditions especially in cancer, cardiovascular, autoimmune and inflammatory diseases. Little is known about the genetics of circulating PlGF levels. Indeed, although the heritability of circulating PlGF levels is around 40%, no studies have assessed the relation between PlGF plasma levels and genetic variants at a genome-wide level. In the current study, PlGF plasma levels were measured in a population-based sample of 2085 adult individuals from three isolated populations of South Italy. A GWAS was performed in a discovery cohort (N = 1600), followed by a de novo replication (N = 468) from the same populations. The meta-analysis of the discovery and replication samples revealed one signal significantly associated with PlGF circulating levels. This signal was mapped to the PlGF co-receptor coding gene NRP1, indicating its important role in modulating the PlGF plasma levels. Two additional signals, at the PlGF receptor coding gene FLT1 and RAPGEF5 gene, were identified at a suggestive level. Pathway and TWAS analyses highlighted genes known to be involved in angiogenesis and immune response, supporting the link between these processes and PlGF regulation. Overall, these data improve our understanding of the genetic variation underlying circulating PlGF levels. This in turn could lead to new preventive and therapeutic strategies for a wide variety of PlGF-related pathologies.

Placental growth factor (PlGF) is a member of the vascular endothelial growth factor (VEGF) family. It is a cytokine with a non essential role in healthy conditions, but with a specific involvement in several malignant, inflammatory and ischemic disorders1-3. High levels of circulating PlGF have been observed in individuals with various diseases such as cancer (breast4, melanoma5, leukemia6), immune diseases (rheumatoid arthritis7, Systemic Lupus Erythematous8), metabolic syndrome9, coronary artery disease10 and neovascular age-related macular degeneration (nAMD)11. PlGF has recently emerged as a predictor of survival and cardiovascular risk and for cardiovascular risk stratification in patients with chronic kidney disease (CKD)12. Women affected by preeclampsia, a pregnancy-specific disorder characterized by the development of hypertension and proteinuria in the later stages of gestation13, show lower PlGF circulating levels compared to healthy pregnant women14-16. All these data show that the amount of PlGF has a relevant impact on the determination of pathological conditions. Moreover, since PlGF levels can be pharmacologically modifiable, understanding the determinants of circulating PlGF could support efforts directed at risk prediction, prevention and therapy.
Although the heritability of circulating PlGF levels is around 40%\textsuperscript{17}, the knowledge about the genetic factors modulating circulating PlGF levels is limited. Indeed, no studies have assessed the relation between circulating PlGF levels and genetic variants at the genome-wide level. In the present work, we have conducted the first genome-wide association study (GWAS) for PlGF in a general population sample with a deep genomic coverage based on imputation to the 1000 Genomes panel, to identify genetic variants that explain variation in circulating PlGF concentrations.

Results
GWAS of PlGF plasma levels: discovery, replication and meta-analysis. A GWAS meta-analysis of PlGF circulating levels was performed in 2068 individuals from three isolated populations of Cilento, South Italy\textsuperscript{18,19}. Characteristics of the study participants are presented in Table 1. A sample of 1600 individuals was used as discovery cohort. An additional sample of 468 individuals from the same populations was used in the de novo replication stage. The replication sample was younger than the discovery sample, no other differences between the two cohorts were observed.

In the discovery GWAS 8,281,256 autosomal SNPs were investigated for association with PlGF plasma levels. A Manhattan plot and a Q-Q plot of the association results are reported in Fig. 1 and S1 Figure, respectively. Although none of the signals in the discovery cohort reached genome-wide significance (\(p\)-value \(< 5 \times 10^{-8}\)), 88 variants were associated at \(p\)-value \(< 1 \times 10^{-5}\) and, among those, 7 variants were associated at \(p\)-value \(< 1 \times 10^{-6}\). A linkage disequilibrium (LD)-based clumping procedure revealed 5 independent signals, suggestively associated with PlGF plasma concentration at a \(p\)-value \(< 1 \times 10^{-6}\). The regional association plots reported in the S2 Figure provide a detailed overview of those loci. Those variants were carried forward to a de novo replication. Summary results for main associations for PlGF plasma levels are shown in Table 2. Overall, in the meta-analysis of discovery and replication, a variant on chromosome 10p11.22 reached a genome-wide significant \(p\)-value; two additional signals, on chromosomes 13q12.3 and 7p15.3, although didn’t reach the genome-wide significance, showed an effect in the same direction between discovery and replication and a lower \(p\)-value in the meta-analysis compared to that of the discovery stage. Therefore, based on these criteria, we considered the 13q12.3 and 7p15.3 loci as replicated signals. In the locus on chromosome 10p11.22, the most significantly associated variant was rs17296631 (\(p\)-value = 8.36 \(\times 10^{-9}\)) located upstream Neuropilin 1 (NRP1) gene, which encodes for a co-receptor of PlGF protein. The rs9551465 variant (\(p\)-value = 7.84 \(\times 10^{-8}\)), identified on chromosome 13q12.3, was located in the 3’UTR of the Fms Related Tyrosine Kinase 1 (FLT1) gene, coding for the PlGF receptor Flt-1/VEGFR-1. The variant rs77619310 (\(p\)-value = 5.27 \(\times 10^{-7}\)) on chromosome 7p15.3 was located in an intron of the RAPGEF5 gene.

The three most associated loci collectively explained 3.68% of the PlGF phenotypic variance.

Colocalization analysis. To verify if the variants associated with PlGF circulating levels directly act through gene expression regulation in a particular tissue, and considering that, given the involvement of PlGF in different pathological conditions, it might exert a role in several tissues, we tested the three signals on chromosomes 7, 10 and 13 for colocalization with expression quantitative trait loci (eQTL) in all the 48 tissues from Genotype-Tissue Expression (GTEx)\textsuperscript{20} using the coloc program\textsuperscript{21} (https://rdrr.io/cran/coloc/man/coloc.abf.html).

We highlighted a very high posterior probability of colocalization (H4 = 0.96) in the Artery Tibial tissue for the rs9551465 variant, which is the best GWAS hit on chromosome 13q12.3, located upstream Neuropilin 1 (NRP1) gene, which encodes for a co-receptor of PlGF protein. The rs9551465 variant (\(p\)-value = 7.84 \(\times 10^{-8}\)), identified on chromosome 13q12.3, was located in the 3’UTR of the Fms Related Tyrosine Kinase 1 (FLT1) gene, coding for the PlGF receptor Flt-1/VEGFR-1. The variant rs77619310 (\(p\)-value = 5.27 \(\times 10^{-7}\)) on chromosome 7p15.3 was located in an intron of the RAPGEF5 gene.

The A allele, reported to determine an increase of RAPGEF5 mRNA levels, was also associated with higher levels of PlGF. This analysis also revealed that the rs9551465, the most associated variant on chromosome 13, showed the best colocalization signal (H4 = 0.79) in the Thyroid tissue, where it is also reported as an eQTL for the FLT1 gene. In this case the allele associated with an increase of PlGF levels was correlated with a lower expression of the FLT1 gene.

| Cohort       | Discovery | Replication |
|--------------|-----------|-------------|
| No of Individuals | 1600 | 468         |
| Women (%)    | 55.5     | 54.1        |
| Age (mean ± SD) | 52.5 ± 19.7 | 35.2 ± 25.2 |
| PlGF (pg/ml) |           |             |
| All          | 12.5     | 12.2        |
| Median       | 4.8—51.2 | 5.6—69.3    |
| Men          | 13.2     | 13.0        |
| Median       | 5.6—34.2 | 5.6—57.7    |
| Women        | 12.0     | 11.7        |
| Median       | 4.8—51.2 | 5.9—69.3    |

Table 1. Characteristics of study participants from the discovery and replication cohorts.
The best signal on chromosome 10, the rs17296631, is in complete LD with rs145141871, the second most associated variant in this locus, and both are reported as eQTLs for the NRP1 gene in the Brain Anterior cingulate cortex tissue. For both variants, the alleles that lead to higher expression levels of the NRP1 gene are also associated with an increase of PlGF levels. Despite this evidence, coloc did not reveal significant colocalization evidence, confirming that in such a situation of complete LD between two variants, the program shows uncertainty in identifying the causal one21.

Transcription-wide and gene-based association analyses. To further identify loci associated with PlGF levels, we performed a transcription-wide association analysis (TWAS) and a gene-based analysis.

In our study, TWAS analysis identified several genes in different genomic regions whose expression in particular tissues, is associated with circulating levels of PlGF protein. The list of the 52 gene/tissue pairs, showing a significant association at FDR (< 0.05), is reported in the S1 Table. Interestingly, the majority of them are located on chromosome 6p22.1-p21.33, in the HLA region. These results suggest that multiple variants, each likely with
a marginal level of significance and located in the HLA genomic region, might contribute, in aggregate, to the regulation of the PlGF circulating levels.

Pathway analysis. To discover biological pathways involved in the modulation of the circulating levels of PlGF, we performed an enrichment analysis using the GSA-SNP2 program\(^\text{23}\). The analysis highlighted 27 Gene Ontology (GO) significantly enriched pathways (q-value < 0.05) (16 Biological Processes, 5 Cellular Components and 6 Molecular Functions), reported in Table 3. In line with the results obtained in the TWAS and the gene-based association analysis, the majority of the significantly enriched pathways were related to immune response. In fact, the GO terms summarization analysis performed by REVIGO program\(^\text{24}\) evidenced that 15 out of the 16 Biological Processes were linked to the immunoglobulin mediated immune response, and that 2 Cellular Components and one Molecular Function were related to the MHC class II complex. Also, 3 GO terms were linked to the lumenal side of the endoplasmic reticulum membrane. Other significantly over-represented pathways were the N-glycan processing, the oxidoreductase activity, the mismatched DNA binding, the folic acid binding and the AU-rich element binding.

**Table 3.** Significantly enriched pathways. Gene Ontologies: Molecular Function (MF), Biological Process (BP), Cellular Component (CC). Dispensability represents a measure of the non-redundancy of a GO term with respect to other semantically close terms. Representative, non-redundant GO terms are given in black, the other cluster members are in italics.

| Gene ontologies | GO term                                                                 | q-value      | Dispensability |
|-----------------|------------------------------------------------------------------------|--------------|----------------|
| BP              | GO:0006491N-glycan processing                                          | 0.042        | 0              |
| BP              | GO:0016064 immunoglobulin mediated immune response                      | 7.23 × 10^{-4} | 0              |
| BP              | GO:0006959 humoral immune response                                      | 0.036        | 0.692          |
| BP              | GO:0042129 regulation of T cell proliferation                           | 0.035        | 0.637          |
| BP              | GO:0002377 immunoglobulin production                                    | 3.69 × 10^{-3} | 0.632         |
| BP              | GO:0002889 regulation of immunoglobulin mediated immune response        | 0.040        | 0.947          |
| BP              | GO:0002381 immunoglobulin production involved in immunoglobulin mediated immune response | 0.035        | 0.952          |
| BP              | GO:0041112 positive regulation of homotypic cell–cell adhesion         | 0.036        | 0.722          |
| BP              | GO:00019724 B cell mediated immunity                                    | 1.42 × 10^{-3} | 0.924         |
| BP              | GO:0002712 regulation of B cell mediated immunity                       | 0.040        | 0.947          |
| BP              | GO:0002440 production of molecular mediator of immune response         | 5.25 × 10^{-3} | 0.627         |
| BP              | GO:0002640 adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin super-family domains | 0.043        | 0.933          |
| BP              | GO:0002637 regulation of immunoglobulin production                     | 0.049        | 0.888          |
| BP              | GO:1903039 positive regulation of leukocyte cell–cell adhesion         | 0.040        | 0.979          |
| BP              | GO:0002455 humoral immune response mediated by circulating immunoglobulin | 0.035        | 0.952          |
| BP              | GO:050870 positive regulation of T cell activation                      | 0.039        | 0.933          |
| CC              | GO:0042613 MHC class II protein complex                                 | 1.83 × 10^{-3} | 0              |
| CC              | GO:0042611 MHC protein complex                                          | 3.21 × 10^{-3} | 0.501         |
| CC              | GO:0071556 integral component of lumenal side of endoplasmic reticulum membrane | 4.37 × 10^{-3} | 0.032         |
| CC              | GO:0098576 lumenal side of membrane                                     | 0.035        | 0.513          |
| CC              | GO:0098553 lumenal side of endoplasmic reticulum membrane               | 4.37 × 10^{3}  | 0.606         |
| MF              | GO:0016641 oxidoreductase activity, acting on the CH-NH2 group of donors, oxygen as acceptor | 6.87 × 10^{-3} | 0              |
| MF              | GO:0030983 mismatched DNA binding                                       | 1.42 × 10^{-3} | 0              |
| MF              | GO:0032395 MHC class II receptor activity                               | 8.59 × 10^{-3} | 0              |
| MF              | GO:000542 folic acid binding                                            | 0.036        | 0.066          |
| MF              | GO:0017091 AU-rich element binding                                      | 0.040        | 0.149          |
| MF              | GO:0016638 oxidoreductase activity, acting on the CH-NH2 group of donors | 0.035        | 0.299          |

Discussion

Our study is the first GWAS of circulating PlGF levels. It was undertaken in 2068 individuals from three population isolates of the Cilento area, South Italy, and represents the largest survey of PlGF measurement in a population-based sample. In this study, we have identified 3 chromosomal regions (7p15.3, 10p11.22 and 13q12.3) harboring variants associated with PlGF plasma levels. The three top variants in these loci explain about 4% of PlGF phenotypic variance.
The newly identified regions include many interesting and plausible candidate genes. The lead variant on chromosome 10, rs17296631, is located in an intergenic region upstream the NRP1 gene. This gene encodes for the Neuropilin-1, a membrane protein devoid of tyrosine kinase activity which acts as a co-receptor for PlGF. Recent studies suggest that NRP-1, through the binding of PlGF and other growth factors, can play an important role in the activation of specific signal transduction pathways also independently of other membrane receptors. A positive correlation of expression of PlGF and NRP-1 has been observed in breast cancer, and PlGF activation of NRP-1 could also promote tumor cell survival in a paracrine manner in a mouse model of medulloblastoma. These observations are in line with the results of the eQTL analysis, in which the T allele of the rs17296631 variant is associated with an increase of both PlGF levels and NRP1 gene expression.

On chromosome 13, the rs9551465 is located in an intron of the FLT1 gene. This gene encodes for a tyrosine-protein kinase, Flt-1, belonging to the vascular endothelial growth factor receptor (VEGFR) family, which functions as a receptor for VEGFA, VEGFB and PlGF. The protein is also present in a soluble form, sFlt-1, lacking of the transmembrane and intracellular domains. The PI GF binding to Flt-1 stimulates angiogenesis via both direct and indirect mechanisms: the activation of Flt-1 by PlGF results in phosphorylation of specific tyrosine residues in Flt-1 and downstream signaling different from those activated by VEGFA binding; also, it has been proposed, based on in vitro data and overexpression studies, that the binding of PlGF to Flt-1 induces pro-angiogenic effects as PlGF shifts VEGFA from Flt-1 to VEGFR2, enhancing the effects of VEGFA. In the present study, we have found that the PlGF associated rs9551465 variant in the FLT1 gene is also an eQTL for FLT1: in particular, the variant allele (T) is associated with higher PI GF protein levels and with a lower FLT1 gene expression. This is consistent with the observation that a higher expression of Flt-1, acting as a “decoy” for PlGF, can determine lower detectable circulating levels of the ligand protein.

The third associated variant, rs77619310, is located on chromosome 7p15.3 in an intron of the RAPGEF5 gene. This gene, encoding for a guanine nucleotide exchange factor (GEF) for the GTPases Rap1, Rap2 and M-RAS, serves as RAS activator by promoting the acquisition of GTP to maintain the active GTP-bound state and is a key link between cell surface receptors and RAS activation. Interestingly, Rap1, activated by different GEFs, including RAPGEF5, acts as a regulator of several basic cellular functions such as adhesion, polarity, differentiation and growth. In the endothelium, Rap1 is a key positive regulator of angiogenic process. It functions downstream of the main angiogenic growth-controlling receptors in endothelial cells, including VEGFA. Also, Rap1 promotes VEGF-mediated angiogenesis through the activation of VEGFR2 and is capable of regulating endothelial barrier permeability by VEGF stimulation.

In a previous candidate gene study, we reported the association of the PI GF gene locus with the levels of the protein in plasma. More recently, another study has found an association between an additional polymorphism in the PI GF gene and the protein plasma levels. In the current study, no variants in the PI GF gene region reached the statistical significance; however, some SNPs located between 5 and 20 kb upstream of the gene were nominally associated with PI GF levels (rs4903273, p-value = 2.04 × 10⁻³) and showed a moderate LD (r² = 0.56) with the rs2268614 SNP described in our previous work.

An implication of HLA genes in the regulation of PlGF levels was detected at statistical significance in the TWAS and was also suggested by a gene-based association analysis, confirming an increased power of TWAS analysis which takes advantage from gene expression level data. In particular, such genes belong to the HLA class III and are involved in both the immune system function and the angiogenic process. In detail, gene expression levels of LYZGB5 and LYGSC, belonging to the cluster of leukocyte antigen-6 (LY6) genes, showed an association with PlGF circulating levels in multiple tissues. Little is known about these two genes, however, other members of the LY6 protein family, expressed in various types of tissues and at different stages of cell differentiation, are involved in cell proliferation, cell migration, cell–cell interactions, immune cell maturation, macrophage activation, and cytokine production, and their overexpression or dysregulation is associated with tumorigenesis and autoimmune diseases. Also, the C4A gene, encoding for the acidic form of the complement factor 4, plays a pivotal role in the activation of immune defenses and the clearance of immune complexes or apoptotic debris in vitro and in vivo studies. A lower number of gene copies of C4A has been linked to an increased risk for different autoimmune diseases, such as Systemic Lupus Erythematosus, Type 1 Diabetes and Juvenile Dermatomyositis. In addition, the deficiency of the C4A gene has also been associated with pre eclampsia, a well-established PlGF related disease, with a lower gene copy number associated with an increase of disease severity, supporting the importance of the classical pathway of the complement system in this pathology. Also, higher levels of C4A have been observed in patients with neovascular age-related macular degeneration, an ocular pathology also characterized by an increase of PI GF levels. Finally, some evidences link the pseudogene MCD to the immune system: in fact, SNPs in this gene region have been associated with eosinophil, basophil and granulocyte count and with different autoimmune disease, such as vitiligo, Graves’ disease and psoriatic arthritis.

In the same region, the Dimethylarginine dimethylaminohydrolase-2 (DDAH2) and the TCF19 genes are implicated in the angiogenic process. The DDAH2 gene encodes for an enzyme that positively regulates the nitric oxide (NO) generation by metabolizing the asymmetric dimethylarginines (ADMA), which are inhibitors of the nitric oxide synthase (NOS) activity. DDAH2 acts as a key regulator of the angiogenesis: in fact, its overexpression enhances the proliferation and migration of endothelial cells through the induction of expression and secretion of VEGF, both regulating the production of endothelial NO through the stimulation of endothelial NOS (eNOS) activity, and in an eNOS-independent manner, via the activation of Sp1 binding site of the VEGF promoter. DDAH2 is considered an antiatherosclerotic molecule: in fact, hypermethylation of DDAH2 promoter, accompanied by its reduced expression, correlates with endothelial dysfunction in patients affected by Coronary Artery Disease. The inhibition of DDAH2 has also been correlated with an attenuation of aberrant angiogenesis and with the improvement of vascular regeneration in mice models of oxygen-induced retinopathy.
plasma levels.17. Also, decreased expression levels of PlGF secretion was upregulated in isolated T-cells, suggesting PlGF as a regulator of T-cell differentiation. In greater inflammatory TH1 and TH17 helper T-cell differentiation, thereby emphasizing the role of PlGF in T-cell

of plasma were immediately prepared and stored at −80°C and were subsequently used for the assessment of

participating to the study.

According to the criteria set by the declaration of Helsinki and each subject signed an informed consent before

University of Naples “Federico II” (Research committee, in 2013 protocol #231/13). The study was conducted

mice with collagen-induced arthritis and that PlGF concentrations also correlated with IL-17 concentrations in

the Flt1 and NRP1 receptors. Also, they show that the overexpression of PlGF in T cells exacerbates disease in

glioma cells to PlGF, upregulated the expression of the TGF-β.69. (TB-403), a (PlGF)-specific inhibitor, completely suppressed the expression of the TGF-β, while the exposure of

naïve B cells captured the PlGF-containing exosomes from glioma cells, they differentiated into TGF-β+Bregs

removed glioma tissues, they demonstrated that glioma cells release exosomes carrying PlGF. When purified

overall role in regulating immune responses involved in inflammation, autoimmunity and cancer. Using surgically

removed glioma tissues, they demonstrated that glioma cells release exosomes carrying PI GF. When purified naïve B cells captured the PI GF-containing exosomes from glioma cells, they differentiated into TGF-β+Bregs able to suppress the CD8+ T cell activities. Further, the treatment of glioma cells with an anti-PI GF antibody (TB-403), a (PI GF)-specific inhibitor, completely suppressed the expression of the TGF-β, while the exposure of glioma cells to PI GF, upregulated the expression of the TGF-β.69.

A recent work has demonstrated that PI GF is selectively secreted by the helper T cells (Th17), a subset of inflammatory T cells that, producing IL-17, contribute to autoimmunity and tissue damage22 and which dysregulation is associated with various autoimmune diseases, including multiple sclerosis and rheumatoid arthritis11,72. In the same article, the authors demonstrate that T cell-produced PI GF is functionally active in promoting angiogenesis and that PI GF stimulates Th17 cell differentiation by activating STAT3 via binding to the Flt1 and NRP1 receptors. Also, they show that the overexpression of PI GF in T cells exacerbates disease in mice with collagen-induced arthritis and that PI GF concentrations also correlated with IL-17 concentrations in synovial fluid from patients with rheumatoid arthritis. Overall, these findings provide an insight into the links between angiogenesis, Th17 cell development, inflammation and autoimmunity, emphasizing the importance of PI GF in these processes59. Previously, Kang et al., using a PI GF-overexpressing transgenic mouse model, showed that PI GF secretion was upregulated in isolated T-cells, suggesting PI GF as a regulator of T-cell differentiation. In addition, the authors evidenced that the CD4+ T-cells isolated from the spleen of transgenic mice indicated greater inflammatory Th1 and Th17 helper T-cell differentiation, thereby emphasizing the role of PI GF in T-cell differentiation and development59.

In conclusion, in this study we have identified some genes and pathways, known to be implicated in the angiogenesis process and the immune response, that are associated with the variation of circulating PI GF. The identification of those genes corroborates the link between PI GF protein levels and immune system function and could lead to new preventive and therapeutic strategies in immune and/or angiogenesis-related diseases in which PI GF has been implicated.

Methods
Population samples and PI GF measurement. The discovery sample includes 1600 volunteer individu-

als recruited through a population-based sampling strategy in three small isolated villages of the Cilento region, South Italy (Campora, Gioi and Cardile). A de novo replication was performed in additional 468 subjects from the same villages18,19. A subset of this sample (N = 871) was used in our previous study on the genetics of PI GF plasma levels12.

The Cilento study was approved by the ethics committee of "Azienda Sanitaria Locale Napoli 1" (Medical committee, in 2003 protocol #291 and #113, and in 2007 protocol #556) and the ethics committee "Carlo Romano" University of Naples “Federico II” (Research committee, in 2013 protocol #231/13). The study was conducted according to the declaration of Helsinki and each subject signed an informed consent before participating to the study.

Blood samples were collected in the morning after the participants had been fasting for at least 12 h. Aliquots of plasma were immediately prepared and stored at –80°C and were subsequently used for the assessment of
PIGF levels. PIGF levels (pg/ml) were measured using an electrochemiluminescence immunoassay on the Elecsys 2010 analyzer (Roche Diagnostics, Mannheim, Germany).

Pregnant women were excluded from the study because of their high level of PIGF in the plasma.

A logarithmic transformation was applied to the PIGF levels to normalize the trait distribution and the transformed trait was used in all subsequent statistical analyses.

The Mann–Whitney U test was used to compare median PIGF plasma levels among the samples.

GWAS and replication study. Genotyping in the discovery sample was performed with 370 K and Omni-Express Illumina chips, phasing and imputation were conducted separately by platform with the MaCH software (https://csg.sph.umich.edu/abecasis/mach/index.html) and minimac (https://genome.sph.umich.edu/wiki/Minimac) software respectively, using 1000 Genomes Phase 1 v3 data as reference. Quality control filters applied before imputation were call rate > 95% for SNPs and samples and minor allele frequency (MAF) > 0.01. GWAS was carried out through a mixed model linear regression where the variance/covariance matrix is the genomic kinship to account for relatedness between individuals. Age and gender were used as covariates and an additive genetic model was considered. The analysis was performed with GenABEL R package (https://cran.r-project.org/src/contrib/Archive/GenABEL/) for genotyped SNPs and ProbABEL (https://github.com/GenABEL-Project/ProbABEL) for imputed data. SNPs with imputation quality (Rsq in MaCH) < 0.8 or MAF < 0.01 were excluded.

To select linkage disequilibrium (LD)-based independent association signals among the PIGF-associated SNPs from the discovery phase, we conducted the clumping procedure implemented in PLINK (https://www.coggen.com/PLINK/) and picked the index SNPs with the most significant association p-value from each clumped association region based on the GWAS. The 1000 Genomes Phase 1 v3 genotypes were used as reference panel; the following thresholds for clumping were applied: association p-value < 1 x 10^-8, physical distance > 1 Mb, and r^2 < 0.01.

Independently associated SNPs in the discovery were de-novo genotyped in the 468 individuals of the replication sample using TaqMan SNP genotyping assays, following the manufacturer’s instructions (Bio-Rad, USA).

To assess evidence for replication, test-statistics of discovery and in silico replication samples were meta-analysed using a fixed effect model weighted by inverse variance, using Metal (http://csg.sph.umich.edu/abecasis/metal/index.html). SNPs were considered replicated if the effect was in the same direction between discovery and replication, and the p-value in the meta-analysis was lower than in the discovery sample.

The proportion of phenotypic variance explained by the PIGF-associated variants was estimated fitting 3 linear mixed effect models, in which PIGF levels were regressed, respectively, on: (1) no covariate; (2) gender and age; (3) gender, age and additive effect of each of the three SNPs. The variance explained by the associated variants was estimated using the gaston R package (https://CRAN.R-project.org/package=gaston) lmm.aireml function (https://search.r-project.org/CRAN/refmans/gaston/html/lmm.aireml.html), which uses the genomic kinship matrix to correct for relatedness between individuals.

Colocalization analysis. Colocalization analysis is a method used to identify shared regulatory variants between a GWAS and an eQTL analysis: if a GWAS trait and a gene expression analysis share the same associated SNP, it may suggest a regulatory role of the SNP mediated through the gene on the GWAS trait.

Colocalization analysis of the PIGF-associated loci with gene expression was conducted using the discovery GWAS results and the cis-eQTL results from 48 tissues in the GTEx Project (Version 7). We considered the three PIGF-associated loci on chromosomes 7, 10 and 13, and, for each of them, we identified all transcripts and all tissue transcript pairs with reported eQTLs within ±500 kb of each GWAS index SNP. We used the colocalization method outlined by Giambartolomei et al., and implemented in the coloc.abf function from the R package coloc (https://rdrr.io/cran/coloc/man/coloc.abf.html), applying the default parameters. Evidence for colocalization was defined as an H4 ≥ 0.8, which represents the posterior probability that the association with PIGF and gene expression is due to the same underlying variant.

Analysis of the imputed genetically regulated gene expression (TWAS). TWAS is a way of integrating expression data and genome-wide association studies, allowing the discovery of genes associated with traits of interest. TWAS analysis typically consists of two steps: first, a model is trained to predict gene expressions from local genetic variants near the focal genes, using a reference dataset containing both genotype and expression data; then, the pre-trained model is used to predict expressions from genotypes in the dataset under study, which contains genotypes and phenotypes. The predicted expressions are then associated with the phenotype of interest.

The genetically regulated gene expression was imputed using PrediXcan v.7 data (https://github.com/hakyimlab/PrediXcan). PrediXcan was used to impute the transcriptome of the 1600 individuals who were included in the discovery GWAS using the SNP weights derived from models trained on reference transcriptome datasets of the 48 tissues in the GTEx Project (Version 7), downloaded from PredicDB, and the genome-wide imputed variants showing MAF > 0.01 and R2 > 0.8. The residuals of the regression of PIGF levels on sex, age and kinship matrix were used as phenotype. We tested for association 1866–8753 protein-coding genes (depending on the tissue) using a linear regression. An FDR < 0.05 was considered as threshold for statistical significance of associations.

Gene-based analysis. A gene-based association analysis was performed using the VEGAS2 software (https://vegas2.qimrberghofer.edu.au/). VEGAS2 is an extension of the VErsatile Gene-based Association Study (VEGAS) approach that uses 1000 Genomes reference data to estimate LD between variants. All the variants from the discovery GWAS were used for the analysis. A ‘−20kbloc’ parameter, which assigns all variants in the
Received: 27 January 2021; Accepted: 28 July 2021
Published online: 19 August 2021

References

1. Carmeliet, P. et al. Synergism between vascular endothelial growth factor and placental growth factor contributes to angiogenesis and plasma extravasation in pathological conditions. Nat. Med. https://doi.org/10.1038/87904 (2001).

2. Fischer, C., Mazzone, M., Jonckx, B. & Carmeliet, P. FLT1 and its ligands VEGF and PlGF: Drug targets for anti-angiogenic therapy?. Nat. Rev. Cancer https://doi.org/10.1038/nrc23524 (2008).

3. Van de Veire, S. et al. Further pharmacological and genetic evidence for the efficacy of PlGF inhibition in cancer and eye disease. Cell https://doi.org/10.1016/j.cell.2010.02.039 (2010).

4. Naik, A. et al. Neuruplin-1 associated molecules in the blood distinguish poor prognosis breast cancer: A cross-sectional study. Sci. Rep. https://doi.org/10.1038/s41598-017-03280-0 (2017).

5. Pagani, E. et al. Placenta growth factor and neuruplin-1 collaborator in promoting melanoma aggressiveness. Int. J. Oncol. https://doi.org/10.3892/ijo.2016.3362 (2016).

6. Schmidt, T. et al. Loss or inhibition of stromal-derived PlGF prolongs survival of mice with imatinib-resistant Bcr-Abl1(+) leukemia. Cancer Cell 19, 740–753 (2011).

7. Yoo, S. A. et al. Role of placenta growth factor and its receptor flt-1 in rheumatoid inflammation: A link between angiogenesis and inflammation. Arthritis Rheum. https://doi.org/10.1002/art.24289 (2009).

8. Zhou, L., Lu, G., Shen, L., Wang, L. & Wang, M. Serum levels of three angiogenic factors in systemic lupus erythematosus and their clinical significance. Biomed Res. Int. https://doi.org/10.1155/2014/627126 (2014).

9. Siervo, M. et al. Mechanism of regulation of the Epac family of cAMP-dependent RapGEFs. J. Cell Sci. https://doi.org/10.1242/jcs.15238 (2003).

10. Fong, S. W. et al. Plasma levels of hypoxia-regulated factors in patients with age-related macular degeneration. Graefes Arch. Clin. Exp. Ophthalmol. https://doi.org/10.1007/s00417-017-3846-z (2018).

11. Levine, R. J. et al. Soluble endoglin and other circulating antiangiogenic factors in preeclampsia. N. Engl. J. Med. https://doi.org/10.1056/NEJMoa053252 (2006).

12. Matsui, M. et al. Synergism between vascular endothelial growth factor and placental growth factor contributes to angiogenesis and plasma extravasation in pathological conditions. Nat. Med. https://doi.org/10.1038/87904 (2001).

13. Wagner, L. K. Diagnosis and management of preeclampsia. Am. J. Obstet. Gynecol. 213, 565–570 (2015).

14. Maynard, S. E. et al. Excess placental soluble fms-like tyrosine kinase 1 (sFlt1) may contribute to endothelial dysfunction hypertension, and proteinuria in preeclampsia. J. Clin. Invest. https://doi.org/10.1172/JCI17189 (2003).

15. Colonna, V. et al. Comparing population structure as inferred from genealogical versus genetic information. Eur. J. Hum. Genet. https://doi.org/10.1038/ejhg.2009.97 (2009).

16. Lonsdale, J. et al. The genotype–tissue expression (GTEx) project. Nat. Genet. https://doi.org/10.1038/ng.2653 (2013).

17. Giambartolomei, C. et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS ONE https://doi.org/10.1371/journal.pone.0042537 (2012).

18. Chau, K., Hennessy, A. & Makris, A. Placental growth factor and pre-eclampsia. J. Hum. Hypertens. https://doi.org/10.1038/jhh.2017.61 (2017).

19. Sorice, R. et al. Genetic and environmental factors influencing the placental growth factor (PlGF) variation in two populations. PLoS ONE https://doi.org/10.1371/journal.pone.0042537 (2012).

20. Colonna, V. et al. Comparing population structure as inferred from genealogical versus genetic information. Eur. J. Hum. Genet. https://doi.org/10.1038/ejhg.2009.97 (2009).

21. Boettner, B. & Van Aelst, L. Control of cell adhesion dynamics by Rap1 signaling. Curr. Opin. Cell Biol. https://doi.org/10.1016/j.ceb.2009.06.004 (2009).

22. Tjwa, M., Luttun, A., Autiero, M. & Carmeliet, P. VEGF and PIGF: Two pleiotropic growth factors with distinct roles in development and homeostasis. Cell Tissue Res. https://doi.org/10.1007/s00441-003-0776-3 (2003).

23. Boet, J. et al. Mechanism of regulation of the Epac family of cAMP-dependent RapGEFs. J. Biol. Chem. https://doi.org/10.1074/jbc.M001113200 (2000).

24. Rebhun, J. F., Castro, A. E. & Quilliam, L. A. Identification of guanine nucleotide exchange factors (GEFs) for the Rap1 GTPase. J. Biol. Chem. https://doi.org/10.1074/jbc.M005327200 (2000).

25. Supek, F., Böšnjak, M., Škunca, N. & Šmuc, T. Revigo summarizes and visualizes long lists of gene ontology terms. BMC Bioinformatics https://doi.org/10.1186/1471-2105-8-276 (2007).

26. Campora: A young genetic isolate in South Italy. Hum. Hered. https://doi.org/10.1159/000101964 (2007).

27. De Rooij, J. et al. Comparing population structure as inferred from genealogical versus genetic information. Eur. J. Hum. Genet. https://doi.org/10.1038/ejhg.2009.97 (2009).

28. Caron, E. Cellular functions of the Rap1 GTP-binding protein: A pattern emerges. J. Cell Sci. https://doi.org/10.1242/jcs.00238 (2003).
31. Chrzanowska-Wodnicka, M. Regulation of angiogenesis by a small GTPase Rap1. Vascul. Pharmacol. https://doi.org/10.1016/j.vph.2010.03.003 (2010).
32. Tawa, H. et al. Role of afadin in vascular endothelial growth factor- and sphenogin 1-phosphate-induced angiogenesis. Circ. Res. https://doi.org/10.1161/CIRCRESAHA.110.216747 (2010).
33. Lakshmikanthan, S. et al. Rap1 promotes VEGFR2 activation and angiogenesis by a mechanism involving integrin αvβ3. Blood https://doi.org/10.1182/blood-2011-04-349328 (2011).
34. Lakshmikanthan, S. et al. Rap1B promotes VEGF-induced endothelial permeability and is required for dynamic regulation of the endothelial barrier. J. Cell Sci. https://doi.org/10.1242/jcs.207605 (2018).
35. Vodolazkaia, A. et al. Vascular endothelial growth factor pathway in endometriosis: Genetic variants and plasma biomarkers. Fertil. Steril. https://doi.org/10.1016/j.fertnstert.2015.12.016 (2016).
36. Gamazon, E. R. et al. A gene-based association method for mapping traits using reference transcriptome data. Nat. Genet. https://doi.org/10.1038/ng.3367 (2015).
37. Kong, H. K. & Park, J. H. Characterization and function of human Ly-6-uPAI molecules. RMB Rep. 45, 595–603 (2012).
38. Loughner, C. L. et al. Organization, evolution and functions of the human and mouse Ly6/uPAI family genes. Hum. Genomics 10, 10 (2016).
39. Upadhyay, G. Emerging role of lymphocyte antigen-6 family of genes in cancer and immune cells. Front. Immunol. https://doi.org/10.3389/fimmu.2019.00819 (2019).
40. Carroll, M. C. The role of complement and complement receptors in induction and regulation of immunity. Annu. Rev. Immunol. https://doi.org/10.1146/annurev.immunol.16.1.1345 (1998).
41. Ross, S. C. & Densen, P. Complement deficiency states and infection: Epidemiology, pathogenesis and consequences of neisserial and other infections in an immune deficiency. Med. (United States) https://doi.org/10.1097/00005109-19840900-00001 (1984).
42. Lintner, K. E. et al. Gene copy-number variations (CNVs) of complement C4 and C4A deficiency in genetic risk and pathogenesis of juvenile dermatomyositis. Ann. Rheum. Dis. https://doi.org/10.1136/annrheumdis-2015-207762 (2016).
43. Li, N. et al. Association between C4, C4A, and C4B copy number variations and susceptibility to autoimmune diseases: A meta-analysis. Sci. Rep. https://doi.org/10.1038/srep42628 (2017).
44. Mason, M. J. et al. Low HERV-K(C4) copy number is associated with type 1 diabetes. Diabetes https://doi.org/10.2337/db13-1382 (2014).
45. Zeidler, H. et al. Predictive value of the sFlt-1:PIGF ratio in women with suspected preeclampsia. N. Engl. J. Med. https://doi.org/10.1056/NEJMoa1414838 (2016).
46. Verloren, S. et al. New gestational phase-specific cutoff values for the use of the soluble fms-like tyrosine kinase-1/placental growth factor ratio as a diagnostic test for preeclampsia. Hypertension 63, 346–352 (2014).
47. Inkeri Lokki, A. et al. Complement activation and regulation in preclamptic placenta. Front. Immunol. 5, 312 (2014).
48. Lechner, J. et al. Higher plasma levels of complement C3a, C4a and C5a increase the risk of subretinal fibrosis in neovascular age-related macular degeneration. Immun. Ageing https://doi.org/10.1186/s12979-016-0060-5 (2016).
49. Omori, T. et al. Evidence for activation of lectin and classical pathway complement components in aqueous humor of neovascular age-related macular degeneration. Ophthalmic Res. https://doi.org/10.1159/000503328 (2019).
50. Astle, W. J. et al. The allelic landscape of human blood cell trait variation and links to common complex disease. Cell https://doi.org/10.1016/j.cell.2016.10.042 (2016).
51. Jin, Y. et al. Genome-wide association studies of autoimmune vitiligo identify 23 new risk loci and highlight key pathways and regulatory variants. Nat. Genet. https://doi.org/10.1038/ng.3680 (2016).
52. Jin, Y. et al. Variant of TYR and autoimmunity susceptibility loci in generalized vitiligo. N. Engl. J. Med. https://doi.org/10.1056/NEJMoa0908547 (2010).
53. Nakabayashi, K. et al. Identification of independent risk loci for Graves disease within the MHC in the Japanese population. J. Hum. Genet. https://doi.org/10.1038/jhg.2011.99 (2011).
54. Aterido, A. et al. Genetic variation at the glycosaminoglycan metabolism pathway contributes to the risk of psoriatic arthritis but not psoriasis. Ann. Rheum. Dis. https://doi.org/10.1136/annrheumdis-2018-214158 (2019).
55. Lu, C. W., Xiong, Y. & He, P. Dimethylarginine dimethylaminohydrolase-2 overexpression improves impaired nitric oxide synthesis of endothelial cells induced by glycated protein. Nitric Oxide - Biol. Chem. https://doi.org/10.1016/j.niox.2006.07.006 (2007).
56. Diederich, L. R. & Wojciak-Stothard, B. The DDAH/ADMA pathway in the control of endothelial cell migration and angiogenesis. Biochem. Soc. Trans. 37, 1243–1247 (2009).
57. Hasegawa, K. et al. Dimethylarginine dimethylaminohydrolase 2 increases vascular endothelial growth factor expression through Sirt1 transcription factor in endothelial cells. Arterioscler. Thromb. Vasc. Biol. https://doi.org/10.1161/01.ATV.0000219615.88323.b4 (2006).
58. Niu, P. P. et al. Hypermethylation of DDAH2 promoter contributes to the dysfunction of endothelial progenitor cells in coronary artery disease patients. J. Transl. Med. https://doi.org/10.1186/1475-2872-12-170 (2014).
59. Lange, C. et al. Dimethylarginine dimethylaminohydrolase-2 deficiency promotes vascular regeneration and attenuates pathological angiogenesis. Exp. Eye Res. https://doi.org/10.1016/j.exer.2016.05.007 (2016).
60. Anderssohn, M. et al. Severely decreased activity of placental dimethylarginine dimethylaminohydrolase in pre-eclampsia. Eur. J. Obstet. Gynecol. Reprod. Biol. https://doi.org/10.1016/j.ejogr.2011.12.032 (2012).
61. Aziizi, F. et al. Modified methylation and expression patterns of genes regulating placental nitric oxide pathway in patients with severe preeclampsia. Hum. Antibodies https://doi.org/10.3233/HAB-180356 (2019).
62. Krautkramer, K. A. et al. Tcf19 is a novel islet factor necessary for proliferation and survival in the INS-1 β-cell line. Am. J. Physiol. Endocrinol. Metab. https://doi.org/10.1152/ajpendo.00147.2013 (2013).
63. Kim, Y. J. et al. A genome-wide association study identified new variants associated with the risk of chronic hepatitis B. Hum. Mol. Genet. https://doi.org/10.1093/hmg/ddl166 (2013).
64. Cheung, Y. H., Watkinson, J. & Anastassiou, D. Conditional meta-analysis stratifying on detailed HLA genotypes identifies a novel type 1 diabetes locus around TCF19 in the MHC. Hum. Genet. https://doi.org/10.1007/s00439-010-0908-2 (2011).
65. Mahajan, A. et al. Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes article. Nat. Genet. https://doi.org/10.1038/ng.1588-018-0084-1 (2018).
66. Xian Zeng, C. et al. TCF19 enhances cell proliferation in hepatocellular carcinoma by activating the ATK/FOXO1 signaling pathway. Neoplasma 66, 46–53 (2019).
67. Zhou, Z. H. et al. TCF19 contributes to cell proliferation of non-small cell lung cancer by inhibiting FOXO1. Cell Biol. Int. https://doi.org/10.1002/cbin.11189 (2019).
68. Lin, Y.-L., Liang, Y.-C. & Chiang, B.-L. Placental growth factor down-regulates type 1 T helper immune response by modulating the function of dendritic cells. J. Leukoc. Biol. https://doi.org/10.1189/jlb.0307164 (2007).
69. Han, S. et al. Glioma cell-derived placental growth factor induces regulatory B cells. Int. J. Biochem. Cell Biol. https://doi.org/10.1016/j.biocel.2014.10.005 (2014).
70. Yoo, S. A. et al. Placental growth factor regulates the generation of TH17 cells to link angiogenesis with autoimmunity. Nat. Immunol. https://doi.org/10.1038/nn.4199-019-0456-4 (2018).
71. Dong, C. TH17 cells in development: An updated view of their molecular identity and genetic programming. Nat. Rev. Immunol. https://doi.org/10.1038/nri2295 (2008).
72. Volin, V. & Shahzada, S. Roles of TH-17 cells in rheumatic and other autoimmune diseases. Rheumatol. Curr. Res. https://doi.org/10.4172/2161-1149.1000104 (2011).

73. Kang, M. et al. Placental growth factor (PIGF) is linked to inflammation and metabolic disorders in mice with diet-induced obesity. Endocrin. J. https://doi.org/10.1507/endocrj.EIJ7-0363 (2018).

74. Li, Y., Willer, C. J., Ding, J., Scheet, P. & Abecasis, G. R. MaCH: Using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genet. Epidemiol. https://doi.org/10.1002/gepi.20533 (2010).

75. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G. R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nat. Genet. https://doi.org/10.1038/ng.2354 (2012).

76. Aulchenko, Y. S., Ripke, S., Isaacs, A. & van Duijn, C. M. GenABEL: An R library for genome-wide association analysis. Bioinformatics https://doi.org/10.1093/bioinformatics/btm108 (2007).

77. R Development Core Team. R: A language and environment for statistical computing. R Found. Stat. Comput. Vienna, Austria. ISBN 3-900, (2017).

78. Aulchenko, Y. S., Struchalin, M. V. & van Duijn, C. M. ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinform. https://doi.org/10.1186/1471-2105-11-134 (2010).

79. Purcell, S. et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. https://doi.org/10.1086/519795 (2007).

80. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: Fast and efficient meta-analysis of genomewide association scans. Bioinformatics https://doi.org/10.1093/bioinformatics/btq340 (2010).

81. Gilmour, A. R., Thompson, R. & Cullis, B. R. Average information REML: An efficient algorithm for variance parameter estimation in linear mixed models. Biometrics 51, (1995).

82. Liu, J. Z. et al. A versatile gene-based test for genome-wide association studies. Am. J. Hum. Genet. https://doi.org/10.1016/j.ajhg.2010.06.009 (2010).

Acknowledgements

We address special thanks to the populations of Campora, Gioi, and Cardile for their participation in the study. We thank dr. Debora Chirico, dr. Michelina De Cristofaro, dr. Raffaele D’Urso and dr. Giovanni D’Arena for Cilento data collection. We thank Mariarosaria Aletta for bibliographic support.

Author contributions

T.N. and D.R. carried out all the analyses of Cilento data. A.T. and S.N. performed the direct genotyping. D.R wrote the initial manuscript. D.R., T.N., and M.C. formulated the study design and main analyses strategies. C.B. and A.L.L. also advised on additional analysis approaches. All authors contributed to the final redaction of the paper.

Funding

This work was supported by grants from the Italian Ministry of Universities (PON03PE_00060_7, IDF SHARID ARS01_01270), FP6 (Vasoplus037254), the Assessorato Ricerca Regione Campania (POR CAMPANIA 2000/2006 MISURA 3.16), Alzheimer’s Association (AARG-NTF-20-683992).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-96256-0.

Correspondence and requests for materials should be addressed to D.R. or M.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021