



**HAL**  
open science

## Identification of an Endoglin Variant Associated With HCV-Related Liver Fibrosis Progression by Next-Generation Sequencing

Frédégonde About, Stéphanie Bibert, Emmanuelle Jouanguy, Bertrand Nalpas, Lazaro Lorenzo, Vimel Rattina, Mohammed Zarhrate, Sylvain Hanein, Mona Munteanu, Beat Müllhaupt, et al.

### ► To cite this version:

Frédégonde About, Stéphanie Bibert, Emmanuelle Jouanguy, Bertrand Nalpas, Lazaro Lorenzo, et al.. Identification of an Endoglin Variant Associated With HCV-Related Liver Fibrosis Progression by Next-Generation Sequencing. *Frontiers in Genetics*, 2019, 10, 10.3389/fgene.2019.01024 . hal-03974811

**HAL Id: hal-03974811**

**<https://hal.sorbonne-universite.fr/hal-03974811>**

Submitted on 7 Feb 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



# Identification of an Endoglin Variant Associated With HCV-Related Liver Fibrosis Progression by Next-Generation Sequencing

## OPEN ACCESS

### Edited by:

Steven J. Schrodí,  
Marshfield Clinic, United States

### Reviewed by:

Yiming Wu,  
Icahn School of Medicine at Mount  
United States  
Carmelo Bernabeu,  
Spanish National Research Council  
(CSIC), Spain

### \*Correspondence:

Aurélie Cobat  
aurelie.cobat@inserm.fr

<sup>†</sup>These authors share senior  
authorship

### Specialty section:

This article was submitted to  
Statistical Genetics  
and Methodology,  
a section of the journal  
Frontiers in Genetics

**Received:** 13 June 2019

**Accepted:** 24 September 2019

**Published:** 04 November 2019

### Citation:

About F, Bibert S, Jouanguy E,  
Nalpas B, Lorenzo L, Rattina V,  
Zarhrate M, Hanein S, Munteanu M,  
Müllhaupt B, Semela D, Semmo N,  
Casanova J-L, Theodorou I,  
Sultanik P, Poynard T, Pol S,  
Bochud P-Y, Cobat A, and Abel L,  
The Swiss Hepatitis C Cohort Study  
Group and The French ANRS HC EP  
26 Genoscan Study Group (2019)  
Identification of an Endoglin Variant  
Associated  
With HCV-Related  
Liver Fibrosis Progression by  
Next-Generation Sequencing.  
Front. Genet. 10:1024.  
doi: 10.3389/fgene.2019.01024

Frédégonde About<sup>1,2</sup>, Stéphanie Bibert<sup>3</sup>, Emmanuelle Jouanguy<sup>1,2</sup>, Bertrand Nalpas<sup>4</sup>,  
Lazaro Lorenzo<sup>1,2</sup>, Vimel Rattina<sup>1,2</sup>, Mohammed Zarhrate<sup>5</sup>, Sylvain Hanein<sup>2,6</sup>,  
Mona Munteanu<sup>7</sup>, Beat Müllhaupt<sup>8</sup>, David Semela<sup>9</sup>, Nasser Semmo<sup>10</sup>,  
Jean-Laurent Casanova<sup>1,2,11,12,13</sup>, Ioannis Theodorou<sup>14</sup>, Philippe Sultanik<sup>15</sup>, Thierry Poynard<sup>16,17</sup>,  
Stanislas Pol<sup>15</sup>, Pierre-Yves Bochud<sup>3</sup>, Aurélie Cobat<sup>1,2,\*†</sup> and Laurent Abel<sup>1,2,11†</sup>, The Swiss  
Hepatitis C Cohort Study Group and The French ANRS HC EP 26 Genoscan Study Group

<sup>1</sup>Laboratory of Human Genetics of Infectious Diseases, Necker Branch, Inserm U1163, Paris, France, <sup>2</sup>Paris Descartes University, Imagine Institute, Paris, France, <sup>3</sup>Infectious Diseases Service, University Hospital and University of Lausanne, Lausanne, Switzerland, <sup>4</sup>Inserm Scientific Information and Communication Department, Inserm, Paris, France, <sup>5</sup>Genomics Core Facility, Imagine Institute, Research Federative Structure Necker, Inserm U1163 and Inserm US24/CNRS UMS3633, Paris Descartes Sorbonne Paris Cite University, Paris, France, <sup>6</sup>Translational Genetics Platform, Inserm U1163, Imagine Institute, Paris Descartes University, Paris, France, <sup>7</sup>BioPredictive Research, Paris, France, <sup>8</sup>Gastroenterology and Hepatology Service, University Hospital of Zürich, Zürich, Switzerland, <sup>9</sup>Division of Gastroenterology and Hepatology, Kantonsspital Sankt Gallen, Sankt Gallen, Switzerland, <sup>10</sup>Department of Visceral Surgery and Medicine, Department of Hepatology, Inselspital Bern, Bern, Switzerland, <sup>11</sup>St. Giles Laboratory of Human Genetics of Infectious Diseases, Rockefeller Branch, Rockefeller University, New York, NY, United States, <sup>12</sup>Howard Hughes Medical Institute, New York, NY, United States, <sup>13</sup>Pediatric Hematology-Immunology Unit, Necker Hospital for Sick Children, AP-HP, Paris, France, <sup>14</sup>Center for Immunology and Infectious Diseases, Inserm UMR S 1135, Pierre et Marie Curie University, Paris, France, <sup>15</sup>Université Paris Centre; U1223, Institut Pasteur; Liver Department, Hôpital Cochin, APHP; Paris, France, <sup>16</sup>Hepatology Department, Assistance Publique-Hôpitaux de Paris, Pitié-Salpêtrière Hospital, Paris, France, <sup>17</sup>Saint-Antoine Research Center & Institute of Cardiometabolism and Nutrition (ICAN), Inserm, Sorbonne University, Paris, France

Despite the astonishing progress in treating chronic hepatitis C virus (HCV) infection with direct-acting antiviral agents, liver fibrosis remains a major health concern in HCV infected patients, in particular due to the treatment cost and insufficient HCV screening in many countries. Only a fraction of patients with chronic HCV infection develop liver fibrosis. While there is evidence that host genetic factors are involved in the development of liver fibrosis, the common variants identified so far, in particular by genome-wide association studies, were found to have limited effects. Here, we conducted an exome association study in 88 highly selected HCV-infected patients with and without fibrosis. A strategy focusing on TGF- $\beta$  pathway genes revealed an enrichment in rare variants of the endoglin gene (*ENG*) in fibrosis patients. Replication studies in additional cohorts (617 patients) identified one specific *ENG* variant, Thr5Met, with an overall odds ratio for fibrosis development in carriers of 3.04 (1.39–6.69). Our results suggest that endoglin, a key player in TGF- $\beta$  signaling, is involved in HCV-related liver fibrogenesis.

**Keywords:** endoglin, HCV-related liver fibrosis, exome sequencing, TGF-beta, rare-variant association study

## INTRODUCTION

End-stage chronic hepatitis C is the leading cause of liver transplantation in developed countries, despite the advent of potent antiviral drugs (Chhatwal et al., 2016). The pathogenesis of severe complications of chronic HCV infection remains unclear (Chhatwal et al., 2016). Most individuals with chronic HCV infection never develop cirrhosis, whereas others develop severe liver fibrosis in less than 20 years (Patin et al., 2012; Rueger et al., 2015). Known common risk factors for liver fibrosis progression include several host characteristics (e.g., male sex, older age at infection, HIV/HBV co-infection, metabolic syndrome, alcohol consumption) and viral factors (e.g., genotype 3) (Rueger et al., 2015). However, these factors explain only a small proportion of the inter-individual variability in liver fibrosis development (Patin et al., 2012), and there is evidence for the involvement of largely unidentified human genetic factors in fibrosis progression (Matsuura and Tanaka, 2018).

Several association studies have assessed the impact of common variants (polymorphisms) on liver fibrosis progression (Matsuura and Tanaka, 2018). In particular, two large genome-wide association studies (GWAS) on liver fibrosis have been conducted in HCV-infected HIV-negative (i.e., mono-infected) patients. The first, in a Caucasian population, identified four susceptibility loci, three of which are related to genes involved in apoptosis (Patin et al., 2012). The second, conducted in Japan, detected variants within the HLA region (Urabe et al., 2013). Finally, a GWAS in Caucasian patients with HCV/HIV co-infection detected another locus on chromosome 3p25 (Ulveling et al., 2016). These common variants have limited effects and could not explain the considerable variability in liver fibrosis development. Lower-frequency variants with stronger individual effects, particularly in coding regions, may instead be involved, and this hypothesis can now be tested by whole-exome sequencing (WES).

Transforming growth factor- $\beta$  (TGF- $\beta$ ) is a key regulator of fibrosis (Meng et al., 2016) and is strongly involved in chronic liver disease contributing to all stages of disease progression from initial liver injury to fibrosis, cirrhosis, and hepato-cellular carcinoma (Dooley and ten Dijke, 2012; Fabregat et al., 2016). High levels of TGF- $\beta$  result in activation of hepatic stellate cells (HSC) to myofibroblasts and massive hepatocyte cell death, which contributes to the promotion of liver fibrosis and later cirrhosis (Tsuchida and Friedman, 2017; Bansal and Chamroonkul, 2019). TGF- $\beta$ 1 promotes fibrosis through the well described TGF- $\beta$ 1/TGFBR1 canonical pathway (Fabregat et al., 2016; Meng et al., 2016). TGF- $\beta$ 1 binds to TGF- $\beta$  receptor 2 (TGFBR2) which recruits and activates TGFBR1 (also known as ALK5). Active TGFBR1 phosphorylates Smad2 and Smad3 which form a complex with Smad4. These heteromeric complexes translocate to the nucleus to stimulate transcription of profibrotic molecules. Inhibitory Smad7 is a negative regulator of Smad2 and Smad3. In the present work, we focused on the TGF- $\beta$  pathway, and we report here the results of a case-control exome-based association study investigating the role of rare and low frequency variants in 707 genes connected to this pathway in a cohort of highly selected patients with chronic HCV infection.

## MATERIALS AND METHODS

### Study Subjects

Patients of the discovery cohort were selected from the French cohort of a previous GWAS of liver fibrosis progression in HCV mono-infected patients (Patin et al., 2012) [the ANRS Genoscan Study Cohort described elsewhere (Patin et al., 2012; Ulveling et al., 2016)]. We selected patients from Paris Hospitals (Cochin and La Pitié-Salpêtrière) with an extreme fibrosis phenotype determined by the examination of a liver biopsy specimen and the duration of HCV infection until liver biopsy. The severe fibrosis group included 48 patients (cases) who developed severe fibrosis (METAVIR score F3 or F4) in less than 30 years after the presumed date of HCV infection (mean duration = 18.8 y; SD = 6.8 y) without major comorbidities. The control group included 40 patients who did not develop fibrosis (METAVIR score F0 or F1) more than 20 years after the presumed date of HCV infection in the absence of treatment (mean = 30.3 y; SD = 7.8 y; **Table 1**).

### Whole Exome Sequencing (WES) and Variant Detection

WES was performed on an Illumina HiSeq 2500 by Agilent SureSelect All Exons V4 + UTRs (71 Mb) Single-Sample Capture at the Genomics Core Facility of the Imagine Institute (Paris, France). We used the Genome Analysis Software Kit (DePristo et al., 2011) (GATK v3.3) best practice pipeline to analyze our WES data. Reads were aligned with the human reference genome (hg19) with the maximum exact matches algorithm in Burrows-Wheeler Aligner (Li and Durbin, 2010). Local realignment around indels were performed with the GATK (McKenna et al., 2010). PCR duplicates were removed with Picard tools ([broadinstitute.github.io/picard/](http://broadinstitute.github.io/picard/)). The GATK base quality score recalibrator (BQSR) was applied to correct sequencing artifacts. Individual genomic variant call files (gVCF) were generated with GATK HaplotypeCaller, and joint genotyping was performed with GATK GenotypeGVCFs. The calling process targeted regions covered by the WES Agilent 71-Mb Kit, including 200 bps flanking each region.

### Variant Level Quality Control

We filtered out sample genotypes with a coverage < 8X, a genotype quality (GQ) < 20, or a ratio of reads for the less covered allele (reference or variant allele) over the total number of reads covering the position (minor read ratio, MRR) < 20%, using an in-house script (Belkadi et al., 2015). We filtered out variant sites (i) multi-allelic (> 2 alleles), (ii) with more than 5% of missing genotypes, and (iii) flagged as probable false-positive variant sites by GATK Variant Quality Score Recalibrator (VQSR). The VQSR model was trained and applied for single nucleotide variants (SNV) and indels following the GATK best practices. We seek to achieve 99 and 95% sensitivities to the accessible sites in the truth set for SNVs and indels, respectively. A set of 439,790 high quality variants was retained for the analysis.

**TABLE 1** | Clinical and demographic characteristics of chronically HCV infected patients of the discovery cohort (N = 88), the French replication cohort (N = 161), and the Swiss replication cohort (N = 453).

Covariate	Category	Discovery cohort			French replication cohort			Swiss replication cohort		
		No fibrosis N = 40	Severe fibrosis N = 48	Total (%) N = 88	No fibrosis N = 102	Severe fibrosis N = 59	Total (%) N = 161	No fibrosis N = 148	Severe fibrosis N = 308	Total (%) N = 456
Sex	Female	22	25	47 (53.4%)	61	22	83 (51.5%)	74	90	164 (36%)
METAVIR score	F0	13	–	13 (14.8%)	15	–	15 (9.3%)	39	–	39 (9%)
	F1	27	–	27 (30.7%)	87	–	87 (54.0%)	106	–	106 (23%)
	F3	–	32	32 (36.3%)	–	30	30 (18.6%)	–	128	128 (28%)
	F4	–	16	16 (18.2%)	–	29	29 (18.0%)	–	180	180 (40%)
Alcohol consumption (°)	Low	40	42	82 (93.1%)	91	40	131 (81.4%)	65	126	191 (42%)
	High	–	7	7 (7.9%)	11	19	30 (18.6%)	80	181	261 (58%)
HCV mode of acquisition	IDU	9	16	25 (28.4%)	39	21	60 (37.3%)	54	101	155 (34%)
	Blood transfusion	22	23	45 (51.1%)	41	29	70 (43.4%)	35	60	95 (21%)
	Other/NA	9	9	18 (20.5%)	22	9	31 (19.3%)	56	147	203 (45%)
HCV genotype	1	25	34	59 (67.0%)	77	39	116 (72.0%)	90	142	232 (51%)
	2	9	1	10 (11.4%)	6	2	8 (5.0%)	13	38	51 (11%)
	3	1	8	9 (10.2%)	13	12	25 (15.5%)	24	91	115 (25%)
	4	1	2	3 (3.4%)	–	–	–	16	30	46 (10%)
	Other/NA	4	3	7 (7.9%)	6	6	12 (7.5%)	2	7	9 (2%)
	Age at infection (y) (°)	Median (Q1–Q3)	18 (11.5–23)	29 (22–39)		21 (17–27)	36 (21–46)		15 (8–18.5)	23 (17–27.5)
Duration of infection (y) (°)	Median (Q1–Q3)	30 (24–34.5)	18 (14–26)		24 (21–29)	16 (12–19)		32 (27–36)	25 (18–31)	

(°)The threshold to define high alcohol consumption in the French cohorts and in the Swiss replication cohort were  $\geq 30\text{g/d}$  and  $\geq 40\text{g/d}$ , respectively. Alcohol consumption was missing in one patient with severe liver fibrosis in the Swiss replication cohort.

(°)Age at infection was missing in 77 patients with severe liver fibrosis in the Swiss replication cohort.

(°)Duration of infection was missing in 77 patients with severe liver fibrosis in the Swiss replication cohort.

## Sample Level Quality Control

We required that each individual sample had a minimum 8X coverage for at least 80% of the targeted bases. We checked that reported sex was concordant with that inferred from sequence data. We further look at the distribution of the transition/transversion (Ti/Tv) ratio, the het/nonref-hom ratio, the indels/SNPs ratio, and the call rate for high quality variants in consensus coding regions. The median Ti/Tv ratio was 3.14, ranging from 3.07 to 3.24, as expected for coding regions. The median het/nonref-hom ratio was 1.59, ranging from 1.37 to 1.83, as expected for samples mainly from European origin (Wang et al., 2015). The median indel/SNP ratio was 0.017, ranging from 0.016 to 0.019 and consistent with the 1:43 ratio previously observed in coding regions (Ng et al., 2008). Finally, all the samples had a final call rate greater than 90% in coding regions and were kept for further analyses. Principal component analysis of our 88 samples with 1,000 genomes samples using high quality variants with minor allele frequency in the cohort  $\geq 0.02$  (Belkadi et al., 2016) confirmed that most patients were of European origin except five patients. Four patients (two cases and two controls) were of North African origin, and one case was very likely of mixed

European and Asian origin (**Supplementary Figure 1**). We decided to keep those five patients and to adjust all association analyses on the three first principal components.

## Variant Selection

We focus our analyses on the most likely functionally relevant variants based on several variant annotations. Variant effects were predicted using snpEff v4.2 (Cingolani et al., 2012) and the Ensembl GRCh37.75 reference database. For variants that have different annotations due to multiple transcripts of the gene, the highest impact effect for each variant was taken. Additional annotations including variant frequencies in ExAC populations (Lek et al., 2016), the Combined Annotation Dependent Depletion (CADD) score (Kircher et al., 2014), and the gene specific Mutation Significance Cutoff (MSC) (Itan et al., 2016) were performed using variant tools (San Lucas et al., 2012). Candidate coding variants (CCV) were finally defined as missense, splice acceptor, splice donor, stop gained, stop loss, start loss, initiator codon variant, exon loss, frameshift, and in-frame insertions/deletions with a CADD score greater than the MSC, and with an ExAC frequency in the European non-Finnish population below 0.05.

## Gene Selection

In the present work, we focused on the TGF- $\beta$  signaling pathway, the central driver of liver fibrosis (Meng et al., 2016). We selected the seven main core genes of this pathway: *TGFB1*, *TGFB1R1*, *TGFB1R2*, *SMAD2*, *SMAD3*, *SMAD4*, and *SMAD7* and further increased the list of genes of interest to 707 autosomal genes directly connected to those 7 core genes according to the human gene connectome (Itan et al., 2013), i.e., we retained only the genes connected to at least one of the 7 main core genes at a degree of connectivity of 1 (Supplementary Table 1). Within those 707 genes of interest, we identified 1,243 high quality CCV as defined in the variant selection section.

## Statistical Analysis

To test whether rare candidate coding variants contribute to fibrosis, we collapsed candidate coding variants into a single burden variable which takes the value 1 if the patient carries at least one CCV and 0 otherwise and we compared the aggregated carrier frequency of candidate coding variants between cases and controls by logistic regression, as implemented in the R software using the *glm* function. Since principal component analysis identified five patients of non-European origin (Supplementary Figure 1), the analyses were adjusted on the three first principal components. We overcame the problem of the small sample size by restricting our analysis to 116 genes for which candidate coding variants were carried by more than five individuals leading to a Bonferroni corrected threshold of  $p < 0.00043$  to account for multiple testing.

## Replication Studies

To replicate our results, we first performed targeted next-generation sequencing focusing on the coding regions of the endoglin gene on the same Illumina HiSeq2500 using a customized SureSelect gene panel in 161 additional Caucasian HCV mono-infected patients from the ANRS Genoscan Study

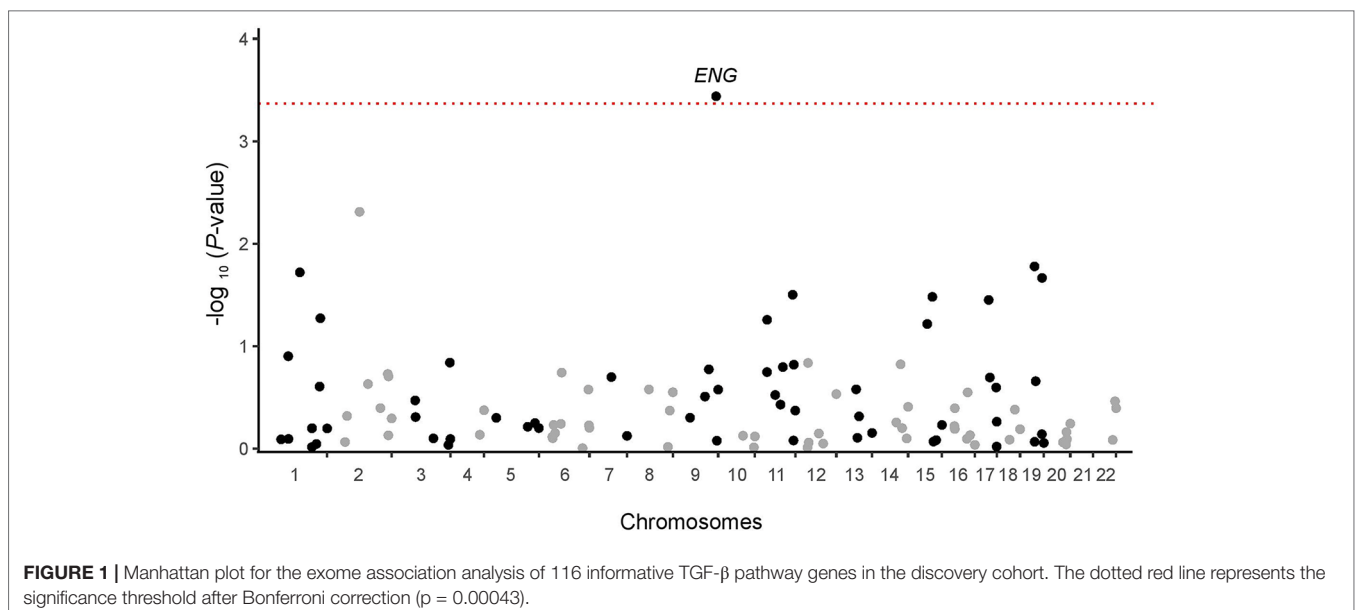
Cohort, including 59 who developed severe liver fibrosis (METAVIR F3–F4) after a mean duration of 16.5 years of infection (SD = 5.5 years), and 102 patients who were fibrosis-free (METAVIR F0–F1) after a mean duration of infection of 25.9 y (SD = 5.8 years; Table 1).

We further investigated the association of rs35400405 with fibrosis in HCV mono-infected patients, in an additional cohort of 456 patients (308 severe fibrosis patients with METAVIR F3–F4 and 148 fibrosis-free patients with METAVIR F0–F1) of Caucasian origin from Switzerland (the Swiss Hepatitis C Cohort Study) with chronic HCV mono-infection (Table 1). In this cohort, rs35400405 was genotyped using a fluorescent based competitive allele specific polymerase chain reaction (KASP, LGC Genomics, Hoddesdon, Herts, United Kingdom). All the methods used were performed in accordance with relevant guidelines and regulations.

## RESULTS

### Discovery Cohort

Results of the exome association study in the discovery cohort of 88 patients for the 116 genes of the TGF- $\beta$  pathway with more than 5 CCV carriers are shown in Figure 1. The main signal observed was that for *ENG*, encoding endoglin, with a  $P$ -value of 0.00037, which remained significant after Bonferroni correction ( $P$ -value<sub>corrected</sub> = 0.043). The  $P$ -value was an order of magnitude larger for the next most closely associated gene (Supplementary Table 2). For *ENG*, there were seven heterozygous carriers of at least one CCV in the severe fibrosis group and none in the control group. There were only three *ENG* candidate coding variants: Thr5Met found in four patients, Gly191Asp in two patients, and Pro131Leu in one patient (Supplementary Table 3). Six of the seven carriers were of European origin, and one, carrying Gly191Asp, was of mixed ancestry, likely Eurasian (Supplementary Figure 1).



**TABLE 2** | Distribution and overall effect of *ENG* Thr5Met on liver fibrosis progression in the discovery cohort and the French and Swiss replication cohorts.

	Discovery		French replication		Swiss replication		All Cohorts		OR [95%CI] <sup>(a)</sup>	P-value <sup>(a)</sup>
	No fibrosis	Severe fibrosis	No fibrosis	Severe fibrosis	No fibrosis	Severe fibrosis	No fibrosis	Severe fibrosis		
N carriers <sup>(b)</sup> (%)	N = 40 0 (0)	N = 48 4 (8.3%)	N = 102 4 (3.9%)	N = 59 6 (10.2%)	N = 148 4 (2.7%)	N = 308 23 (7.5%)	N = 290 8 (2.8%)	N = 415 33 (8.0%)	3.04 (1.39–6.69)	0.0024

<sup>(a)</sup>Computed by means of the likelihood ratio test in the logistic regression framework.

<sup>(b)</sup>All carriers were heterozygous for Thr5Met.

## Replication Studies

For replication of these results, we performed targeted next-generation sequencing focusing on the coding regions of *ENG* in 161 additional Caucasian HCV mono-infected patients. No significant enrichment in *ENG* candidate coding variants was found in this replication cohort, mostly due to the presence of three novel rare candidate coding variants in the control group (**Supplementary Table 3**). However, the frequency of carriers of the most common CCV enriched in our discovery cohort, Thr5Met (c.14C > T), was higher among fibrosis patients (10.2%) than among controls (3.9%) in the replication cohort although not significant ( $P$ -value = 0.12).

We conducted a second replication study by genotyping Thr5Met (rs35400405) in a larger independent Swiss cohort of HCV mono-infected patients including 308 patients with liver fibrosis and 148 without fibrosis (**Table 1**). Thr5Met was found to be significantly more abundant ( $P$ -value = 0.032) in fibrosis patients (7.5%) than in controls (2.7%). Overall, when all cohorts were combined, the frequency of Thr5Met carriers was 8.0% in fibrosis patients and 2.8% in controls, resulting in an odds ratio of 3.04 (1.39–6.69;  $P$ -value = 0.0024; **Table 2**). The frequency of Thr5Met carriers in >50,000 Europeans in the gnomAD database (<http://gnomad.broadinstitute.org/>) was 4.1%, confirming an enrichment of this variant in fibrosis patients, and a depletion in those without fibrosis.

## In Silico Analysis of Thr5Met Variant

The Thr5Met variant maps to exon 1 of *ENG* encoding the signal peptide (**Figure 2A**). Thr5Met was predicted to be benign by PolyPhen and tolerated by SIFT, but it is in very strong linkage disequilibrium with five SNPs (**Figure 2B**) one of which, rs45514697, is predicted to affect transcription factor binding ([www.regulomedb.org](http://www.regulomedb.org)). SNP rs45514697 is predicted to modify a regulatory motif for liver factor (LF)-A1 (also known as hepatocyte nuclear factor 4), which is involved in the expression of several liver-specific genes (Ramji et al., 1991), with a lower affinity for the rare allele (<http://archive.broadinstitute.org/mammals/haploreg/haploreg.php>).

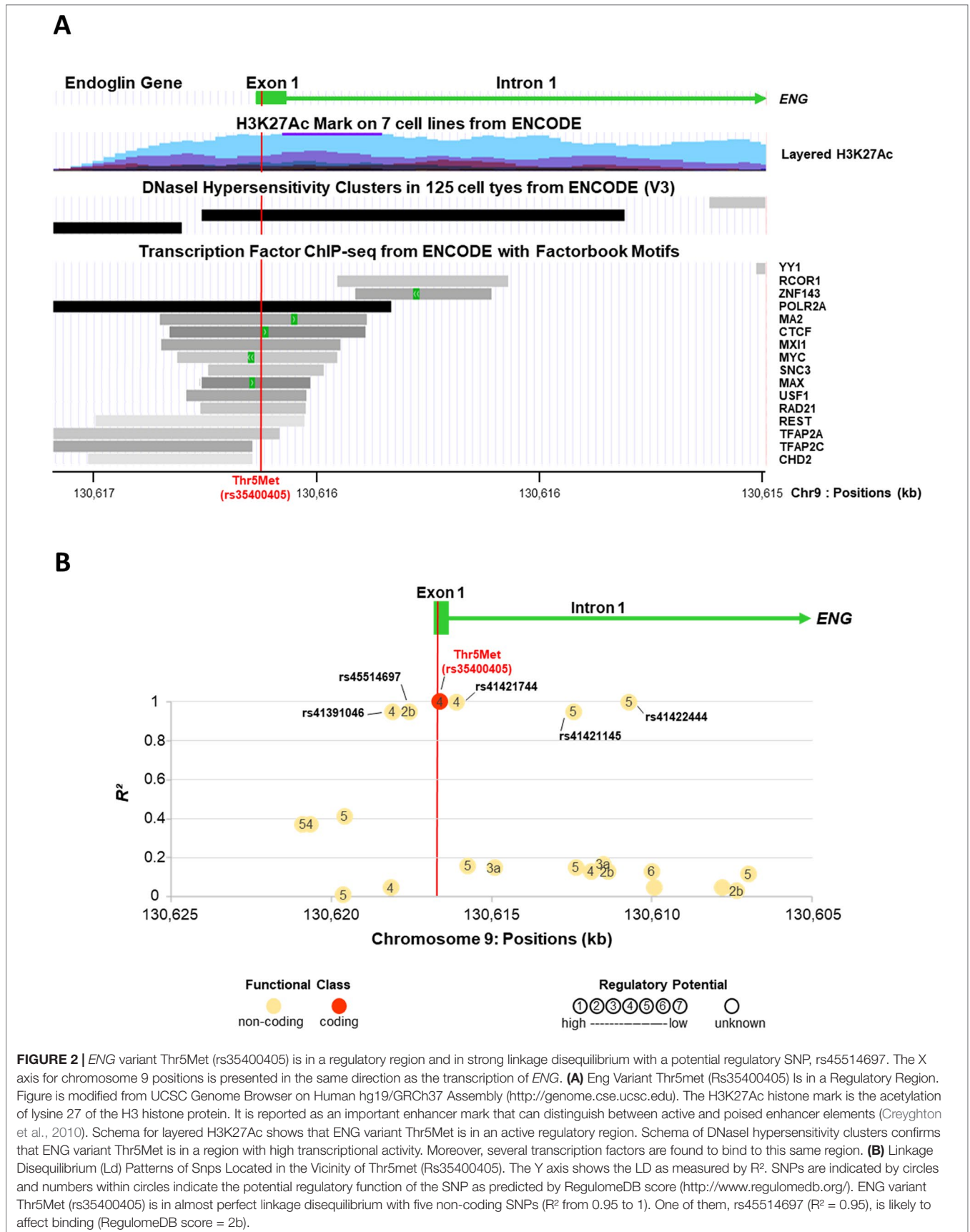
## DISCUSSION

In the present study, we used a three-stage analysis to identify potential protein-coding variants that contribute to liver fibrosis development in HCV mono-infected individuals. In contrast to GWAS investigating common variants, we focused on more

rare potentially damaging coding variants identified by WES in 707 candidate genes linked to the TGF- $\beta$  signaling pathway. We identified one low frequency missense variant (Thr5Met) in *ENG* gene, encoding endoglin, associated with liver fibrosis development. Very interestingly, this variant is also in perfect linkage disequilibrium with rs45514697, an upstream *ENG* variant located in a regulatory region. The rare allele of rs45514697 which is at risk for liver fibrosis is predicted to strongly decrease the DNA-binding affinity for LF-A1, a transcription factor involved in the expression of several liver specific genes (Ramji et al., 1991).

Endoglin, also known as cluster of differentiation CD105, is a TGF- $\beta$  co-receptor. Heterozygous LOF mutations of *ENG* cause hereditary hemorrhagic telangiectasia type 1 (HHT1), also known as Osler–Weber–Rendu disease, a rare autosomal dominant condition characterized by telangiectasias and arteriovenous malformations (Meurer et al., 2014). Interestingly, the three variants identified in the discovery cohort, Thr5Met, Pro131Leu, and Gly191Asp, were previously observed in HHT1 patients (Cymerman et al., 2003; Abdalla et al., 2005; Fernandez et al., 2006; Olivieri et al., 2007; Fontalba et al., 2008). Variants Thr5Met and Gly191Asp are presently considered as non-pathogenic polymorphisms (Abdalla and Letarte, 2006; Olivieri et al., 2007), consistently with their frequency in public databases (MAF > 1% in gnomAD European population). The pathogenic role of Pro131Leu in HHT is more debated as the variant is more rare (MAF = 0.001 in the gnomAD European population), was associated with reduced levels of endoglin (Cymerman et al., 2003; Fernandez et al., 2006), but was also found in healthy individuals (Abdalla and Letarte, 2006; Olivieri et al., 2007). We did not have direct information about HHT symptoms (such as epistaxis or telangiectasia) for the patients carrying *ENG* variants in our cohort. However, they had no contraindication to liver biopsy, and no signs of arterio-venous malformations were reported in the biopsy. Of note, 41–74% of HHT patients present liver involvement with hepatic vascular malformations (Buscarini et al., 2018). Livers with HHT can show nodular hyperplasia and fibrosis around the abnormal vessels which, associated with portal hypertension, may lead to a misdiagnosis of cirrhosis (Buscarini et al., 2018).

In addition to its well-established role in angiogenesis, there is evidence to suggest that endoglin is involved in fibrogenesis (Meurer et al., 2014). Endoglin is strongly expressed in profibrogenic cell types such as mesangial cells, cardiac fibroblasts, scleroderma fibroblasts, and HSCs which are the most fibrogenic cells of the liver. Indeed, endoglin plays a crucial role in TGF- $\beta$  signaling and is thought to modulate the balance between the main profibrotic TGF- $\beta$ 1/TFBR1 signaling pathway and the TGF- $\beta$ 1/ACVRL1



**FIGURE 2 |** *ENG* variant Thr5Met (rs35400405) is in a regulatory region and in strong linkage disequilibrium with a potential regulatory SNP, rs45514697. The X axis for chromosome 9 positions is presented in the same direction as the transcription of *ENG*. **(A)** Eng Variant Thr5met (Rs35400405) Is in a Regulatory Region. Figure is modified from UCSC Genome Browser on Human hg19/GRCh37 Assembly (<http://genome.cse.ucsc.edu>). The H3K27Ac histone mark is the acetylation of lysine 27 of the H3 histone protein. It is reported as an important enhancer mark that can distinguish between active and poised enhancer elements (Creighton et al., 2010). Schema for layered H3K27Ac shows that ENG variant Thr5Met is in an active regulatory region. Schema of DNase-seq hypersensitivity clusters confirms that ENG variant Thr5Met is in a region with high transcriptional activity. Moreover, several transcription factors are found to bind to this same region. **(B)** Linkage Disequilibrium (LD) Patterns of Snps Located in the Vicinity of Thr5met (Rs35400405). The Y axis shows the LD as measured by  $R^2$ . SNPs are indicated by circles and numbers within circles indicate the potential regulatory function of the SNP as predicted by RegulomeDB score (<http://www.regulomedb.org/>). ENG variant Thr5Met (rs35400405) is in almost perfect linkage disequilibrium with five non-coding SNPs ( $R^2$  from 0.95 to 1). One of them, rs45514697 ( $R^2 = 0.95$ ), is likely to affect binding (RegulomeDB score = 2b).

signaling pathway (Velasco et al., 2008), which has a less studied and more controversial role in fibrosis (Munoz-Felix et al., 2013). There are two isoforms of endoglin resulting from alternative splicing, the long (L-Eng) and the short (S-Eng) endoglin (Gougos and Letarte, 1990; Bellon et al., 1993). It was previously found that L-Eng, the predominant form, enhances ACVRL1 (also known as ALK1) (Velasco et al., 2008) while S-Eng promotes TGFBR1 (Blanco et al., 2008; Velasco et al., 2008) signaling pathways. It was also shown that endoglin mediates the crosstalk between the TGF- $\beta$  and the fibronectin/integrin signaling pathways in endothelial cells (Tian et al., 2012). Interestingly, HCV core protein was recently reported to modulate endoglin signaling in liver pathogenesis by upregulating endoglin expression on cell surface and activating its downstream ACVRL1 signaling pathway (Kwon et al., 2017). In addition, patients with chronic HCV infection and liver cirrhosis were found to have high intrahepatic and circulating endoglin levels (Clemente et al., 2006).

Whether endoglin stimulates or inhibits fibrosis is still debated. Several *in vitro* studies have been conducted that support either hypothesis. In numerous cells types, such as mesangial cells, myoblasts, fibroblasts, or chondrocytes, endoglin regulates negatively TGF- $\beta$ -induced extracellular matrix (ECM) protein expression (Diez-Marques et al., 2002; Guo et al., 2004; Obreo et al., 2004; Rodriguez-Barbero et al., 2006; Finnson and Philip, 2012). Other experiments outlined a profibrogenic role of endoglin in HSC cell lines (Meurer et al., 2011; Meurer et al., 2013) or in scleroderma patient fibroblasts (Morris et al., 2011). In an *in vivo* mouse model of kidney fibrosis following ureteral obstruction, L-Eng overexpression was associated with higher amounts of type I collagen and fibronectin (Oujo et al., 2014). More recently, endoglin haploinsufficiency in mice was shown to differentially regulate ECM production in skin and cartilage (Alzahrani et al., 2018). The conflicting results observed on the role of endoglin in fibrosis may lie in the respective contribution of the two splice variants L- and S-endoglin which differ in their cytoplasmic domain and could have antagonizing effects (Velasco et al., 2008; Munoz-Felix et al., 2016), and/or in differential effects according to tissues (Alzahrani et al., 2018) or experimental conditions.

In a more relevant murine model of liver fibrosis, endoglin deficiency in HSC was recently shown to significantly aggravate liver fibrosis in response to injury suggesting a protective role of endoglin against liver fibrosis (Alsamman et al., 2018). This is consistent with a putative hypomorphic role of the rare allele of rs45514697 in liver fibrosis risk, with a predicted decrease in affinity at the corresponding LF-A1 motif. Further functional studies of Thr5Met and rs45514697, in particular, are required in HSCs, to determine their role in endoglin expression and/or function and provide insight into the potential antifibrogenic role of endoglin in the liver. These results may open the way for new treatments aimed at targeting specific signaling pathways involved in liver fibrogenesis (Bansal and Chamroonkul, 2019).

## DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the dbGaP, accession phs001902.v1.p1, [http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study\\_id=phs001902.v1.p1](http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001902.v1.p1)

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the french Committee for the Protection of Persons (CPP) Île de France 3 of the Tarnier-Cochin Hospital (Paris, France) for the Genoscan study. The ethics committees of all participating centers granted ethical approval for the Swiss Hepatitis C Cohort Study: Ethics Committee, Medical Department, Zurich; Ethics Committee, Medical Department, Geneva; Ethics Committee for clinical research, Faculty of Biology and Medicine, Lausanne; Ethics Committee Canton St Gallen; Ethics Committee Canton of Bern; Ethics Committee Canton Ticino; Intercantonal Ethics Committee for Jura, Fribourg and Neuchatel; Ethics Committee of both Basel). The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

AC and LA conceived and designed the study. BN, MM, BM, DS, NS, PS, TP, SP, and P-YB contributed to the recruitment of participants and the clinical measurements. MZ and SH carried out the sequencing of individuals. LL, EJ, VR, and IT contributed reagents/materials/analysis tools. FA and SB performed the data analysis. FA, AC, and LA interpreted the results and wrote the first draft of the manuscript. J-LC contributed to the manuscript in its final form. All authors reviewed the manuscript.

## FUNDING

The French cohort is supported and sponsored by The National Agency for Research on AIDS and Viral Hepatitis (ANRS) (ANRS Study HC EP 26 Genoscan). The SCCS is supported by grants from the Swiss National Science Foundation (3347C0-108782/1), the Swiss Federal Office for Education and Sciences (03.0599), and the European Commission (LSHM-CT-2004-503359; VIRGIL Network of Excellence on Antiviral Drug Resistance). P-YB was supported by the Swiss National Foundation (grant number 324730-144054); the Santos-Suarez Foundation and the Leenaards foundation. FA is the recipient of a fellowship from Fondation pour la Recherche Médicale (FRM, Grant n° FDM20140630671) The laboratory of Human Genetics of infectious Diseases is supported by The National Agency for Research on AIDS and Viral Hepatitis (ANRS), Institut National de la Santé et de la Recherche Médicale (INSERM), Paris Descartes University, the French National Research Agency (ANR) under the "Investments for the future" program (grant no. ANR-10-IAHU-01).

## ACKNOWLEDGMENTS

We thank all members of both branches of the laboratory of Human Genetics of Infectious Diseases for fruitful discussions. We warmly thank all the study subjects.



We thank additional members of the Swiss Hepatitis C Cohort Study Group: Francesco Negro (Geneva, Chairman), Antoine Hadengue (Geneva, Chairman of Scientific Committee), Laurent Kaiser, Laura Rubbia-Brandt (Geneva); Darius Moradpour, Cristina Cellerai (Lausanne); Martin Rickenbach (Lausanne Data Center); Andreas Cerny, Gladys Martinetti (Lugano); Jean-François Dufour, Meri Gorgievski, Virginie Masserey Spicher (Berne); Markus Heim, Hans Hirsch (Basel); Beat Helbling, Stephan Regenass (Zurich); Raffaele Malinverni (Neuchatel); Guenter Dollenmaier (St Gallen); Gieri Cathomas (Liestal).

We thank additional members of the French ANRS HC EP 26 Genoscan Study Group: Laurence Bousquet (Paris), Yen

Ngo (Paris), Pascal Lebray (Paris), Joseph Moussalli (Paris), Yves Benhamou (Paris), Dominique Thabut (Paris), Anaïs Vallet-Pichard (Paris), H el ene Fontaine (Paris), Vincent Mallet (Paris), Philippe Sogni (Paris), Jean-Baptiste Trabut (Paris), Marc Bourli ere (Marseilles), Jean-Fran ois Delfraissy (Paris).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2019.01024/full#supplementary-material>

## REFERENCES

- Abdalla, S. A., Cymerman, U., Rushlow, D., Chen, N., Stoeber, G. P., Lemire, E. G., et al. (2005). Novel mutations and polymorphisms in genes causing hereditary hemorrhagic telangiectasia. *Hum. Mutat.* 25 (3), 320–321. doi: 10.1002/humu.9312
- Abdalla, S. A., and Letarte, M. (2006). Hereditary haemorrhagic telangiectasia: current views on genetics and mechanisms of disease. *J. Med. Genet.* 43 (2), 97–110. doi: 10.1136/jmg.2005.030833
- Alsamman, M., Sterzer, V., Meurer, S. K., Sahin, H., Schaeper, U., Kuscuoglu, D., et al. (2018). Endoglin in human liver disease and murine models of liver fibrosis—A protective factor against liver fibrosis. *Liver. Int.* 38 (5), 858–867. doi: 10.1111/liv.13595
- Alzahrani, A., Chi, Y., Finnson, K. W., Bhati, M., Lussier, B., Kapoor, M., et al. (2018). Endoglin haploinsufficiency is associated with differential regulation of extracellular matrix production during skin fibrosis and cartilage repair in mice. *J. Cell. Commun. Signal.* 12 (1), 379–388. doi: 10.1007/s12079-018-0461-7
- Bansal, M. B., and Chamroonkul, N. (2019). Antifibrotics in liver disease: are we getting closer to clinical use? *Hepatol. Int.* 13 (1), 25–39. doi: 10.1007/s12072-018-9897-3
- Belkadi, A., Bolze, A., Itan, Y., Cobat, A., Vincent, Q. B., Antipenko, A., et al. (2015). Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proc. Natl. Acad. Sci. U. S. A.* 112 (17), 5473–5478. doi: 10.1073/pnas.1418631112
- Belkadi, A., Pedernana, V., Cobat, A., Itan, Y., Vincent, Q. B., Abhyankar, A., et al. (2016). Whole-exome sequencing to analyze population structure, parental inbreeding, and familial linkage. *Proc. Natl. Acad. Sci. U. S. A.* 113 (24), 6713–6718. doi: 10.1073/pnas.1606460113
- Bellon, T., Corbi, A., Lastres, P., Cales, C., Cebrian, M., Vera, S., et al. (1993). Identification and expression of two forms of the human transforming growth factor-beta-binding protein endoglin with distinct cytoplasmic regions. *Eur. J. Immunol.* 23 (9), 2340–2345. doi: 10.1002/eji.1830230943
- Blanco, F. J., Grande, M. T., Langa, C., Oujou, B., Velasco, S., Rodriguez-Barbero, A., et al. (2008). S-endoglin expression is induced in senescent endothelial cells and contributes to vascular pathology. *Circ. Res.* 103 (12), 1383–1392. doi: 10.1161/CIRCRESAHA.108.176552
- Buscarini, E., Gandolfi, S., Alicante, S., Londoni, C., and Manfredi, G. (2018). Liver involvement in hereditary hemorrhagic telangiectasia. *Abdom. Radiol.* 43 (8), 1920–1930. doi: 10.1007/s00261-018-1671-4
- Chhatwal, J., Wang, X., Ayer, T., Kabiri, M., Chung, R. T., Hur, C., et al. (2016). Hepatitis C Disease Burden in the United States in the era of oral direct-acting antivirals. *Hepatol.* 64 (5), 1442–1450. doi: 10.1002/hep.28571
- Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly. (Austin)* 6 (2), 80–92. doi: 10.4161/fly.19695
- Clemente, M., Nunez, O., Lorente, R., Rincon, D., Matilla, A., Salcedo, M., et al. (2006). Increased intrahepatic and circulating levels of endoglin, a TGF-beta1 co-receptor, in patients with chronic hepatitis C virus infection: relationship to histological and serum markers of hepatic fibrosis. *J. Viral. Hepat.* 13 (9), 625–632. doi: 10.1111/j.1365-2893.2006.00733.x
- Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U. S. A.* 107 (50), 21931–21936. doi: 10.1073/pnas.1016071107
- Cymerman, U., Vera, S., Karabegovic, A., Abdalla, S., and Letarte, M. (2003). Characterization of 17 novel endoglin mutations associated with hereditary hemorrhagic telangiectasia. *Hum. Mutat.* 21 (5), 482–492. doi: 10.1002/humu.10203
- DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43 (5), 491–498. doi: 10.1038/ng.806
- Diez-Marques, L., Ortega-Velazquez, R., Langa, C., Rodriguez-Barbero, A., Lopez-Novoa, J. M., Lamas, S., et al. (2002). Expression of endoglin in human mesangial cells: modulation of extracellular matrix synthesis. *Biochim. Biophys. Acta.* 1587 (1), 36–44. doi: 10.1016/S0925-4439(02)00051-0
- Dooley, S., and ten Dijke, P. (2012). TGF-beta in progression of liver disease. *Cell. Tissue. Res.* 347 (1), 245–256. doi: 10.1007/s00441-011-1246-y
- Fabregat, I., Moreno-Caceres, J., Sanchez, A., Dooley, S., Dewidar, B., Giannelli, G., et al. (2016). TGF-beta signalling and liver disease. *FEBS. J.* 283 (12), 2219–2232. doi: 10.1111/febs.13665
- Fernandez, L. A., Sanz-Rodriguez, F., Zarrabeitia, R., Perez-Molino, A., Morales, C., Restrepo, C. M., et al. (2006). Mutation study of Spanish patients with hereditary hemorrhagic telangiectasia and expression analysis of Endoglin and ALK1. *Hum. Mutat.* 27 (3), 295. doi: 10.1002/humu.9413
- Finnson, K. W., and Philip, A. (2012). Endoglin in liver fibrosis. *J. Cell. Commun. Signal.* 6 (1), 1–4. doi: 10.1007/s12079-011-0154-y
- Fontalba, A., Fernandez, L. A., Garcia-Alegria, E., Albinana, V., Garrido-Martin, E. M., Blanco, F. J., et al. (2008). Mutation study of Spanish patients with hereditary hemorrhagic telangiectasia. *BMC. Med. Genet.* 9, 75. doi: 10.1186/1471-2350-9-75
- Gougos, A., and Letarte, M. (1990). Primary structure of endoglin, an RGD-containing glycoprotein of human endothelial cells. *J. Biol. Chem.* 265 (15), 8361–8364.
- Guo, B., Slevin, M., Li, C., Parameshwar, S., Liu, D., Kumar, P., et al. (2004). CD105 inhibits transforming growth factor-beta-Smad3 signalling. *Anticancer. Res.* 24 (3a), 1337–1345.
- Itan, Y., Shang, L., Boisson, B., Ciancanelli, M. J., Markle, J. G., Martinez-Barricarte, R., et al. (2016). The mutation significance cutoff: gene-level thresholds for variant predictions. *Nat. Methods.* 13 (2), 109–110. doi: 10.1038/nmeth.3739
- Itan, Y., Zhang, S. Y., Vogt, G., Abhyankar, A., Herman, M., Nitschke, P., et al. (2013). The human gene connectome as a map of short cuts for morbid allele discovery. *Proc. Natl. Acad. Sci. U. S. A.* 110 (14), 5558–5563. doi: 10.1073/pnas.1218167110
- Kircher, M., Witten, D. M., Jain, P., O’Roak, B. J., Cooper, G. M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* 46 (3), 310–315. doi: 10.1038/ng.2892
- Kwon, Y. C., Sasaki, R., Meyer, K., and Ray, R. (2017). Hepatitis C virus core protein modulates endoglin (CD105) Signaling pathway for liver pathogenesis. *J. Virol.* 91 (21), e01235-17. doi: 10.1128/JVI.01235-17

- Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nat.* 536 (7616), 285–291. doi: 10.1038/nature19057
- Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinf.* 26 (5), 589–595. doi: 10.1093/bioinformatics/btp698
- Matsuura, K., and Tanaka, Y. (2018). Host genetic variations associated with disease progression in chronic hepatitis C virus infection. *Hepatol. Res.* 48 (2), 127–133. doi: 10.1111/hepr.13042
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome. Res.* 20 (9), 1297–1303. doi: 10.1101/gr.107524.110
- Meng, X. M., Nikolic-Paterson, D. J., and Lan, H. Y. (2016). TGF-beta: the master regulator of fibrosis. *Nat. Rev. Nephrol.* 12 (6), 325–338. doi: 10.1038/nrneph.2016.48
- Meurer, S. K., Alsamman, M., Sahin, H., Wasmuth, H. E., Kisseleva, T., Brenner, D. A., et al. (2013). Overexpression of endoglin modulates TGF-beta1-signalling pathways in a novel immortalized mouse hepatic stellate cell line. *PLoS. One.* 8 (2), e56116. doi: 10.1371/journal.pone.0056116
- Meurer, S. K., Alsamman, M., Scholten, D., and Weiskirchen, R. (2014). Endoglin in liver fibrogenesis: Bridging basic science and clinical practice. *World. J. Biol. Chem.* 5 (2), 180–203. doi: 10.4331/wjbc.v5.i2.180
- Meurer, S. K., Tihaa, L., Borkham-Kamphorst, E., and Weiskirchen, R. (2011). Expression and functional analysis of endoglin in isolated liver cells and its involvement in fibrogenic Smad signalling. *Cell. Signal.* 23 (4), 683–699. doi: 10.1016/j.cellsig.2010.12.002
- Morris, E., Chrobak, I., Bujor, A., Hant, F., Mummery, C., Ten Dijke, P., et al. (2011). Endoglin promotes TGF-beta/Smad1 signaling in scleroderma fibroblasts. *J. Cell. Physiol.* 226 (12), 3340–3348. doi: 10.1002/jcp.22690
- Munoz-Felix, J. M., Gonzalez-Nunez, M., and Lopez-Novoa, J. M. (2013). ALK1-Smad1/5 signaling pathway in fibrosis development: friend or foe? *Cytokine. Growth. Factor. Rev.* 24 (6), 523–537. doi: 10.1016/j.cytogfr.2013.08.002
- Munoz-Felix, J. M., Perez-Roque, L., Nunez-Gomez, E., Oujó, B., Arevalo, M., Ruiz-Remolina, L., et al. (2016). Overexpression of the short endoglin isoform reduces renal fibrosis and inflammation after unilateral ureteral obstruction. *Biochim. Biophys. Acta.* 1862 (9), 1801–1814. doi: 10.1016/j.bbadis.2016.06.010
- Ng, P. C., Levy, S., Huang, J., Stockwell, T. B., Walenz, B. P., Li, K., et al. (2008). Genetic variation in an individual human exome. *PLoS. Genet.* 4 (8), e1000160. doi: 10.1371/journal.pgen.1000160
- Obreo, J., Diez-Marques, L., Lamas, S., Duwell, A., Eleno, N., Bernabeu, C., et al. (2004). Endoglin expression regulates basal and TGF-beta1-induced extracellular matrix synthesis in cultured L6E9 myoblasts. *Cell. Physiol. Biochem.* 14 (4-6), 301–310. doi: 10.1159/000080340
- Olivieri, C., Pagella, F., Semino, L., Lanzarini, L., Valacca, C., Pilotto, A., et al. (2007). Analysis of ENG and ACVRL1 genes in 137 HHT Italian families identifies 76 different mutations (24 novel). Comparison with other European studies. *J. Hum. Genet.* 52 (10), 820–829. doi: 10.1007/s10038-007-0187-5
- Oujó, B., Munoz-Felix, J. M., Arevalo, M., Nunez-Gomez, E., Perez-Roque, L., Pericacho, M., et al. (2014). L-Endoglin overexpression increases renal fibrosis after unilateral ureteral obstruction. *PLoS. One.* 9 (10), e110365. doi: 10.1371/journal.pone.0110365
- Patin, E., Kutalik, Z., Guergnon, J., Bibert, S., Nalpas, B., Jouanguy, E., et al. (2012). Genome-wide association study identifies variants associated with progression of liver fibrosis from HCV infection. *Gastroenterol.* 1431244-1252 (5), e1241–e1212. doi: 10.1053/j.gastro.2012.07.097
- Ramji, D. P., Tadros, M. H., Hardon, E. M., and Cortese, R. (1991). The transcription factor LF-A1 interacts with a bipartite recognition sequence in the promoter regions of several liver-specific genes. *Nucleic. Acids. Res.* 19 (5), 1139–1146. doi: 10.1093/nar/19.5.1139
- Rodriguez-Barbero, A., Obreo, J., Alvarez-Munoz, P., Pandiella, A., Bernabeu, C., and Lopez-Novoa, J. M. (2006). Endoglin modulation of TGF-beta1-induced collagen synthesis is dependent on ERK1/2 MAPK activation. *Cell. Physiol. Biochem.* 18 (1-3), 135–142. doi: 10.1159/000095181
- Rueger, S., Bochud, P. Y., Dufour, J. F., Müllhaupt, B., Semela, D., Heim, M. H., et al. (2015). Impact of common risk factors of fibrosis progression in chronic hepatitis C. *Gut.* 64 (10), 1605–1615. doi: 10.1136/gutjnl-2014-306997
- San Lucas, F. A., Wang, G., Scheet, P., and Peng, B. (2012). Integrated annotation and analysis of genetic variants from next-generation sequencing studies with variant tools. *Bioinf.* 28 (3), 421–422. doi: 10.1093/bioinformatics/btr667
- Tian, H., Myhre, K., Golzio, C., Katsanis, N., and Blobel, G. C. (2012). Endoglin mediates fibronectin/alpha5beta1 integrin and TGF-beta pathway crosstalk in endothelial cells. *EMBO. J.* 31 (19), 3885–3900. doi: 10.1038/emboj.2012.246
- Tsuchida, T., and Friedman, S. L. (2017). Mechanisms of hepatic stellate cell activation. *Nat. Rev. Gastroenterol. Hepatol.* 14 (7), 397–411. doi: 10.1038/nrgastro.2017.38
- Ulveling, D., Le Clerc, S., Cobat, A., Labib, T., Noirel, J., Laville, V., et al. (2016). A new 3p25 locus is associated with liver fibrosis progression in human immunodeficiency virus/hepatitis C virus-coinfected patients. *Hepatol.* 64 (5), 1462–1472. doi: 10.1002/hep.28695
- Urabe, Y., Ochi, H., Kato, N., Kumar, V., Takahashi, A., Muroyama, R., et al. (2013). A genome-wide association study of HCV-induced liver cirrhosis in the Japanese population identifies novel susceptibility loci at the MHC region. *J. Hepatol.* 58 (5), 875–882. doi: 10.1016/j.jhep.2012.12.024
- Velasco, S., Alvarez-Munoz, P., Pericacho, M., Dijke, P. T., Bernabeu, C., Lopez-Novoa, J. M., et al. (2008). L- and S-endoglin differentially modulate TGFbeta1 signaling mediated by ALK1 and ALK5 in L6E9 myoblasts. *J. Cell. Sci.* 121 (Pt 6), 913–919. doi: 10.1242/jcs.023283
- Wang, J., Raskin, L., Samuels, D. C., Shyr, Y., and Guo, Y. (2015). Genome measures used for quality control are dependent on gene function and ancestry. *Bioinf.* 31 (3), 318–323. doi: 10.1093/bioinformatics/btu668

**Conflict of Interest:** Author MM was employed by company BioPredictive Research, France.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 About, Bibert, Jouanguy, Nalpas, Lorenzo, Rattina, Zarhrate, Hanein, Munteanu, Müllhaupt, Semela, Semmo, Casanova, Theodorou, Sultanik, Poynard, Pol, Bochud, Cobat and Abel, The Swiss Hepatitis C Cohort Study Group and The French ANRS HC EP 26 Genoscan Study Group. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.