Hereditary mixed polyposis syndrome is caused by a 40kb upstream duplication that leads to increased and ectopic expression of the BMP antagonist *GREM1*

Emma Jaeger[#1], **Simon Leedham**[#1,2], **Annabelle Lewis**[1], **Stefania Segditsas**[1], **Martin Becker**[1], **Pedro Rodenas Cuadrado**[1], **Hayley Davis**[1], **Kulvinder Kaur**[3], **Karl Heinimann**[1,4], **Kimberley Howarth**[1], **The HMPS Collaboration**[6], **James East**[2], **Jenny Taylor**[3], **Huw Thomas**[5], and **Ian Tomlinson**[1,3]

[1]Molecular and Population Genetics Laboratory, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, United Kingdom [2]Translational Gastroenterology Unit, John Radcliffe Hospital, Oxford OX3 9DU, United Kingdom [3]NIHR Comprehensive Biomedical Research Centre, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, United Kingdom [4]Abt. Medizinische Genetik, Universitaetskinderspital beider Basel, Burgfelderstrasse 101, Haus J CH-4055 Basel, Switzerland [5]Family Cancer Clinic, Imperial College School of Medicine, St Mark's Hospital, Watford Road, Harrow A1 3UJ, United Kingdom

[#] These authors contributed equally to this work.

Abstract

The hereditary mixed polyposis syndrome (HMPS) was first described about 50 years ago in a large Ashkenazi Jewish family from St Mark's Hospital, London. The family showed apparent autosomal dominant inheritance of multiple types of colorectal polyp, with colorectal carcinoma in a high proportion of individuals. In the last 15 years, we have mapped the HMPS gene to chromosome 15q13.3 and identified an ancestral haplotype common to all the known HMPS families. Here, we have used genetic mapping, copy number analysis, exclusion of mutations by high-throughput sequencing, gene expression analysis and functional assays to show that HMPS is caused by a large duplication spanning the 3′ end of the *SCG5* gene and a region upstream of the *GREM1* locus. This mutation has no effect on *SCG5* expression, but is associated with greatly increased, allele-specific *GREM1* expression. Whilst *GREM1* is expressed in intestinal sub-epithelial myofibroblasts in controls, HMPS patients predominantly express *GREM1* in the

epithelium of the large bowel. The HMPS duplication contains predicted transcriptional enhancer elements; we have shown that some of these interact with the *GREM1* promoter and are capable of driving gene expression *in vitro*. Increased *GREM1* expression is predicted to lead to reduced bone morphogenetic protein pathway activity, a mechanism that also underlies tumorigenesis in juvenile polyposis of the large bowel. The pathogenic mechanism in HMPS is extremely unusual in Mendelian cancer syndromes and highlights ectopic gene expression as a mechanism of tumorigenesis.

Colorectal cancers that arise as a result of high-penetrance germline mutations in genes such as *APC, MSH2* and *MLH1* are associated with tumours of particular morphology and with variable, but specific extra-colonic features. HMPS is an unusual disease in that patients can develop polyps of multiple and mixed morphologies – including serrated lesions, Peutz-Jeghers polyps, juvenile polyps, conventional adenomas and colorectal cancer (CRC) [1] – in the absence of any identifiable extra-colonic features (http://omim.org/entry/601228). A single, large Jewish HMPS family (SM96) with dominantly-inherited polyposis was initially identified at St Mark's Hospital, London, and germline mutations in the known CRC predisposition genes were excluded over time [2]. Using linkage analysis in SM96 and independently in another Jewish polyposis kindred, SM1311, we mapped the HMPS gene (also known as *CRAC1*) to chromosome 15q13.3 and showed that the two families shared a haplotype within this region [2,3]. Subsequently, we identified additional HMPS kindreds, based on their clinical features and Ashkenazi descent, and demonstrated that all families shared a disease haplotype between 30.735Mb and 31.370Mb on chromosome 15q13.3 [4]. To date, no non-Ashkenazi HMPS families have convincingly been identified, although there is overlap between some of the clinical features of HMPS and those of juvenile polyposis syndrome (JPS) [5] and hyperplastic polyposis syndrome (HPPS) [6].

The possibility that a large proportion of HMPS cases was caused by a single founder mutation led us to hypothesize that the underlying genetic change might be somewhat unusual: we reasoned that disease-causing mutations might have to be of a very specific type that occurred infrequently as spontaneous events. Consistent with this notion, we found no convincing evidence of loss-of-heterozygosity at 15q13.3 in 12 HMPS tumours (details not shown), suggesting that the gene was not a classical tumour suppressor. We also sequenced three HMPS patients and two unaffected relatives who did not share the disease haplotype for the coding regions of the 3 genes (*GREM1*, *SCG5* and *FMN1*) that lay within the region of the shared haplotype. We found known polymorphisms, but no novel, potentially pathogenic changes that were consistently present in patients and absent in their unaffected relatives.

We therefore wondered whether the underlying HMPS mutation might take the form of a copy number change. We designed a custom oligonucleotide array to search for copy number variation in the region. This analysis showed the presence of a heterozygous, single-copy duplication of about 40kb centred on chr15:30.77Mb in two HMPS patients and in none of three unaffected relatives (Figure 1, Supplementary Figure 1). PCR amplification across the duplication breakpoints subsequently mapped it to chr15:30,752,231-30,792,051. The change was found to be a simple, tandem tail-head duplication with the insertion of a

30bp sequence of unknown origin and no homology to known sequences between the duplicons. The duplication extended from intron 2 of *SCG5* to a site just upstream of the *GREM1* CpG island (Figure 1). In order to investigate the duplication further, we designed PCR primers that spanned the duplication boundary and produced a unique, specific amplification product of 190bp. As a control, we chose a region upstream of *GREM1* present in all individuals that generated a product of 435bp. We tested 40 affected individuals (Supplementary Figure 2) and 50 unaffected individuals (either polyp-free, but at-risk and aged >40 years old, or spouses) from 6 putative HMPS families. There was perfect concordance between presence of the 190bp product and affected status (Supplementary Figure 1, Supplementary Table 1). We then tested 188 unselected Ashkenazi controls and none showed amplification of the 190bp product; this group included one duplication-negative individual who shared the HMPS haplotype, suggesting that the duplication had arisen on that haplotype background.

We then wondered whether duplication of chromosome 15q13.3 – either identical to that in HMPS or a similar change – might be present in a set of familial colorectal cancer cases from the general population of the United Kingdom. We therefore used the QuantiSNP program [7] to analyse our existing Illumina Hap550 SNP array data from 718 CRC cases and 935 controls enrolled in the CORGI study of CRC genetics [8]. One case, but no controls, was found to have a duplication close to *GREM1* and *SCG5* and we showed that this change was identical to that found in the HMPS cases. The patient concerned had been excluded from the CORGI genome-wide association study because ancestry was uncertain on principal component analysis. On further investigation, we found that the family was of Ashkenazi descent and had a phenotype of multiple polyps and CRC, entirely compatible with HMPS, but not previously diagnosed as such (Supplementary Figure 2). We found no non-Ashkenazi patients with duplications in the HMPS region.

Although the *SCG5-GREM1* duplication was highly likely to be pathogenic and was an unusual type of mutation consistent with the apparent monophyletic origin of HMPS, we wished to exclude the presence of other potentially pathogenic variants on the HMPS haplotype. We therefore screened the entire ancestral haplotype region (chr15:30,735,098-31,369,755) in 2 cases and 3 unaffected individuals from 3 different HMPS families using 454 amplicon sequencing. In addition, we sequenced the region in each of two somatic cell hybrids that had been made, respectively carried one copy of the mutant and one copy of the "wildtype" chromosome 15 from an HMPS patient. After filtering out variants in dbSNP and a small number of artefacts, only 3 variants were specifically present in all cases and in no controls. All these variants were single nucleotide changes deep within introns of *FMN1* Supplementary Table 2). None of the changes was predicted to have profound functional effects, although one change at a conserved location early in *FMN1* (base 31,038,855) lay within a DNAse I hypersensitive site and thus had potential effects on gene transcription. Differences in *FMN1* transcription were subsequently excluded (see below).

On the basis that the *SCG5-GREM1* duplication was by far the most likely pathogenic mutation in the HMPS region, we sought to determine which gene(s) was affected by that change. Since the duplication involved the latter half of the coding region of *SCG5,* in

addition to a large part of the region between *SCG5* and *GREM1,* we initially attempted to detect aberrant *SCG5* transcripts in our patients based on predictions from the duplicated exons. However, only the normal *SCG5* mRNA species was found (details not shown). We then wondered whether the duplication could act as an enhancer of gene transcription. mRNA levels for *SCG5, GREM1* and *FMN1* were assayed in 87 individual normal colonic crypts from 8 HMPS patients, and 55 individual crypts from 8 United Kingdom controls as well as 8 crypts from 1 HMPS family member without the duplication. These control patients had clear colonoscopies and no history of colorectal tumours. We found massively increased *GREM1* transcript levels in the normal epithelium of HMPS patients compared with the controls, but no significant differences in the expression of *SCG5* or *FMN1* (Figure 2). Allele-specific expression (ASE) analysis showed significantly increased expression of the duplicated allele in HMPS crypts (Mann Whitney p=0.0009). Attempts to confirm that the duplicated allele was expressed more strongly in lymphoblastoid cells from our HMPS patients were unsuccessful because we found *GREM1* levels to be very low in these cells. We did, however, examine our somatic cell hybrids carrying the mutant and "wildtype" copies of chromosome 15, since these cells were based on mouse fibroblast lines that did express *GREM1*. We detected 4.6-fold higher *GREM1* mRNA expression in the cells with the mutant human chromosome 15 compared with those carrying the "wildtype" copy.

We then examined the localisation of *GREM1* mRNA in the normal colorectal crypts of HMPS patients and controls. In controls and HMPS family members without the duplication, *GREM1* mRNA was restricted to the sub-epithelial myofibroblasts (ISEMFs) at the crypt bases, as has previously been reported [9,10]. In HMPS patients, by contrast, *GREM1* was expressed not only in basal ISEMFs, but also at very high levels in epithelial cells (predominantly colonocytes), extending most of the way up the sides of the crypt (Figure 2). Increased *GREM1* expression was also seen in HMPS polyps, albeit less dramatically than in normal epithelium.

Given the increased, ectopic expression of *GREM1* in HMPS patients, we tested whether elements within the HMPS duplication could enhance gene transcription. Although ENCODE data (http://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=219213609&c=chr15&g=wgEncodeReg) suggested that at least two regions within the duplication might have enhancer activity (Supplementary Figure 3), we focussed on a 3kb region (chr15:30,779,000-30,782,000), since we had previously demonstrated this to show histone H4K4 acetylation and H3K4 methylation marks indicative of active chromatin [11] and subsequently shown with chromatin immunoprecipitation to contain transcription factor binding sites (M. Becker, A. Lewis, I. Tomlinson*,* unpubl.). Using a luciferase reporter assay, we found that a fragment encompassing this 3kb region was capable of 4-fold enhancement of gene expression in the colorectal cancer cell line SW948 that expresses *GREM1* (P=0.00002, t test; Supplementary Table 3). We then used chromatin conformation capture (3C) to show that the 3kb region interacts directly with the *GREM1* promoter in *GREM1*-expressing colorectal cancer cell lines (details not shown). 3C analysis in the chromosome 15 somatic cell hybrids found a much stronger promoter interaction in the hybrid with the duplication than in the control (Figure 3).

The BMP signalling pathway is not well characterised compared with other pathways. The canonical pathway signals through BMP receptors and drives expression of the ID transcriptional repressor genes through SMAD1/5/8. A non-canonical pathway acts through decreased PTEN expression and increased levels of active AKT [12]. BMP2 and BMP4, the main pathway ligands in the coloretum, have been postulated to act against the stem cell phenotype and hence increased GREM1 would be expected to reduce BMP ligand levels and promote the stem cell phenotype. In keeping with that hypothesis, we found a small, but significant increase in expression of the stem/progenitor cell marker *LGR5* in whole HMPS colonic crypts compared with crypts from controls (P=0.03, Mann-Whitney test). There was no detectable difference in phospho-SMAD1,5,8 staining in HMPS lesions and we were not able to show an increase in ID1/2/3 expression in crypts from HMPS patients. We did however, detect a decrease in PTEN expression in some HMPS polyps (but not tissue of normal appearance), together with patches of complete PTEN loss (Supplementary Figure 4), a finding of note given the presence of hamartomatous polyps in the intestines of some Cowden syndrome patients, who carry germline *PTEN* mutations [13]. There was a small increase in cytoplasmic phospho-AKT expression in HMPS polyps (Supplementary Figure 4), although this was not spatially identical to the changes in PTEN expression.

In summary, we have shown that the Mendelian colorectal polyposis syndrome, HMPS, results from a duplication of about 40kb upstream of the gene that encodes the secreted BMP antagonist GREM1. Although the duplication includes some sequence within the neighbouring *SCG5* gene, no differences in SCG5 expression were found. Whilst the duplication predisposes to benign lesions which are hypothesised to progress to CRC in some cases, it is of note that some sporadic adenomas and CRCs acquire low-level, ectopic epithelial expression of *GREM1* (Leedham *et al,* unpubl.)*,* suggesting that the increased levels in HMPS may act to promote multiple stages of colorectal tumorigenesis. To date, we have only found the HMPS duplication in the Ashkenazi Jewish descendants of a single founder, but we do not exclude the possibility that rare patients of other ethnic origins carry pathogenic copy number changes or other mutations in *GREM1* that considerably or modestly increase their risk of CRC. In this regard, it is interesting that a single CRC case with whole-gene duplication of *GREM1* has been reported [14], although, if pathogenic rather than incidental, that mutation produces features of hereditary non-syndromic CRC [15] rather than HMPS. It is also notable that our genome-wide association studies have shown common single nucleotide polymorphisms (rs16969681 and rs11632715) within the HMPS duplication to be associated with relatively modest differences in CRC risk in the general population [11]. The HMPS duplication is associated, we contend causally, with increased, ectopic *GREM1* transcription in colorectal epithelium. It is of note in this regard that the limb deformity mouse phenotype is known to result from deletion of a limb bud-specific *Grem1* control region within *Fmn1* [16]; we wonder, therefore whether the HMPS duplication involves an intestine-specific *GREM1* control region partly within *SCG5*. Although no reliable GREM1 antibody exists, it is expected that the increased mRNA causes increased epithelial GREM1 protein secretion and hence a reduction in BMP signalling to the crypt cells. A reduction in BMP signalling is also the likely cause of juvenile polyposis syndrome, in which either the type 1A BMP receptor or the downstream BMP effector *SMAD4* is mutated and functionally defective [17,18]. The detailed functional consequences of the HMPS

duplication, including full mapping of regulatory sites and the downstream effects of inactivating BMP signalling, await further investigation in the context of HMPS, but plausible modes of action include aberrant stem cell numbers or function and effects on the non-canonical BMP pathway *via* PTEN. It seems highly unlikely that all the cells over-expressing *GREM1* in the normal bowel of HMPS patients represent early tumour clones. One potential underlying pathogenic mechanism in HMPS is an increase in the number of cells that are susceptible to tumour-causing mutations, probably through microenvironmental effects of increased *GREM1* secretion.

Methods

Patients and samples

Individuals and families were originally classified as possibly having a diagnosis of HMPS on the basis of early-onset, multiple colorectal polyps (including some of serrated or mixed morphology), coupled with no obvious excess of the extra-colonic tumours typical of other Mendelian CRC syndromes. Germline mutations in *APC, MUYTH, MSH2, MLH1, MSH6, PMS2, LKB1/STK11, SMAD4* and *BMPR1A* were excluded. Following identification of the Ashkenazi HMPS haplotype, this was used to confirm the diagnosis and this testing indicated that all Ashkenazi HMPS families showed evidence of dominant inheritance of the disease. Any non-Jewish patients (or Jewish patients without the HMPS haplotype) with features suggestive of HMPS were classified as "Possible HMPS", although very few of these had evidence of dominantly-inherited disease and many were given a presumptive primary diagnosis of hyperplastic polyposis syndrome [6].

Ethical approval for the study was provided by Southampton and South-West Hampshire Research Ethics Committee A. After obtaining informed consent, peripheral blood samples were taken and DNA extracted from these. Tumour samples, including colorectal carcinomas, were obtained from Histopathology archives. Fresh biopsies of normal bowel were additionally taken from a number of HMPS patients at colonoscopy, and from tumour-free individuals who were having investigative colonoscopy for symptoms subsequently shown to have innocent origins.

Lymphoblastoid cell lines were made from several HMPS patients and unaffected family members, and from one affected individual, two somatic cell hybrids were made in a murine NIH-3T3 background, each hybrid respectively containing a single copy of chromosome 15 carrying the mutant and wildtype haplotypes at the HMPS locus.

Genotyping and variant screening

Microsatellite genotyping employed fluorescence-based fragment analysis on ABI3730 sequencers.

Copy number analysis was undertaken using an Oxford Gene Technology (http://www.ogt.co.uk/gsarrayCGH.asp) custom-design oligonucleotide microarray comprising 6,761 probes spanning a 2.2Mb region that was delimited based on critical flanking recombinants from linkage analysis data available at the time. In brief, test (HMPS patient or unaffected family member) DNA was labelled with Cy5 and control (standard reference)

DNA was labelled with Cy3. DNAs were simultaneously hybridised to the arrays and relative Cy3/Cy5 intensities measured. Copy number gain was assessed as a consistent relative increase in Cy5 and loss as an increase in Cy3 using proprietary OGT software.

Ilummina Infinium Hap550 arrays were used for SNP genotyping and subsequently for copy number analysis. HMPS samples were analysed alongside a larger sample set from the CORGI genome-wide association study of CRC susceptibility in the general UK population. Quality control was performed as described for the CORGI study [11]. Copy number analysis on the Hap550 data was performed using the QuantiSNP program [7].

DNA sequencing was originally performed using conventional Sanger sequencing on the ABI3730 and analysed by visual inspection of sequence electropherograms from the mutant and "wildtype" somatic cell hybrids, based on human-specific amplicons. More recently, sequencing was performed using the Roche454 GSFlex system and software. The region of interest was tiled with long range PCR amplimers and the amplification products pooled at equal stoichiometric concentration. Sequencing libraries were prepared using the Roche 454 Rapid Library preparation method and each sample tagged with a unique molecular identifier before pooling and emulsion PCR amplification. 200Mb of sequencing data was generated in total, which was demultiplexed and analysed using the Roche gsMapper software.

All PCR primer sequences and reaction conditions are available on request. Human Genome Build36 was used as a reference throughout this manuscript.

Real-time quantitative PCR (qRT-PCR)

The Taqman system was used to analyse the expression of *GREM1*, *SCG5*, *FMN1* and *LGR5*. Small biopsies of normal bowel were taken from the ileum and throughout the colorectum of HMPS cases and controls. Tissue was incubated in 5mls of Dulbecco's modified Eagle medium (Invitrogen) with added 30mmol/L ethylenediaminetetraacetic acid (EDTA; Sigma, UK), 0.1 mol/L dithiothreitol, and 100ul RNA later solution (Ambion, Austin, TX, USA) for 15 minutes. Tissue was transferred to phosphate buffered saline (PBS) and then vigorously shaken for 20 seconds. Individual intestinal crypts were then drawn up using glass pipettes and transferred to RNA lysis buffer (Qiagen, Crawley, UK). After several cycles of shaking with fresh PBS washes, complete epithelial denudement occurred and the residual mesenchymal tissue was collected and processed. Extracted mRNA was analysed by qRT-PCR on the ABI7700 system. Primers and probes for test mRNAs (details on request) were obtained from the ABI catalogue. A *GAPDH* control probe was used. Expression levels were determined using the standard DDCt method.

Allele specific-expression (ASE)

ASE analysis was done following the protocol of Lo *et al* [19]. A Taqman SNP genotyping assay for rs12915554, within the 3′UTR of the *GREM1* gene was used to identify two HMPS and three control patient heterozygotes. Phase was assigned using existing genotyping data, including those from the somatic cell hybrids. Control homozygote DNA was mixed in different ratios and the fluorescent intensity data was used to generate a linear

regression standard curve line. ASE for the heterozygote HMPS and control cDNA from individual crypts was then measured using qRT-PCR, and the gene expression allelic ratio measured by intercepting the log of the fluorescent intensity ratio on the standard curve.

In situ tissue analyses

4 micron sections were de-waxed and rehydrated by standard methods. Endogenous peroxidase was blocked with 3% H2O2 in methanol for 10 minutes. Antigen retrieval was achieved by 10 minutes pressure-cooking in sodium citrate buffer at pH 6. Slides were blocked in goat serum for 30 minutes. The following primary antibodies were used for immunohistochemistry: rabbit antibody to phospho-SMAD1/5/8 which recognises the phosphorylated forms of SMAD1 (Ser463/465), SMAD 5 (Ser463/465) and SMAD8 (Ser426/428) (1:50 Cell Signaling Technology, Beverly, MA) rabbit antibody to PTEN (1:100 Cell Signaling Technology), rabbit antibody to Ser473-phosphorylated AKT (1:50 Cell Signaling Technology). This was followed by 1:200 dilution in biotinylated goat anti-rabbit secondary antibodies (DAKO, Glostrup, Denmark) before incubation in ABC (contains avidin and biotinylated horseradish peroxidise) (Vector Labs, Peterborough, UK). Each layer was applied for 30 minutes and three 5-minute PBS washes were performed between layers. Sections were then developed with 3,3-diaminobenzidine-tetrahydrochloride solution (DAB; Sigma, Gillingham, UK) for 2 minutes, followed by rinsing in tap water and light haematoxylin counterstaining. Negative controls underwent all steps but were incubated with PBS instead of the primary antibody solution.

Since suitable antibodies continue to be unavailable, expression levels of *GREM1, FMN1* and *LGR5* were analysed by mRNA *in situ* hybridisation using the method described by Poulsom *et al* [20] with ^{35}S-labelled riboprobes. *SCG5* was primarily analysed in this way, although immunohistochemistry was also used to provide supporting evidence. Details of the riboprobes used are available on request. Expression was scored in a qualitative fashion (by SL and SS) as present or absent, or as stronger or weaker than in control samples analysed alongside.

Assessment of control of GREM1 expression

For luciferase reporter assays, the region chr15:30,778,456-30,782,505 was cloned into the pGL3-Promoter vector (Promega) to test enhancer function in forward and reverse orientations. SW948 cells were transiently co-transfected (in triplicate) with the appropriate pGL3-promoter constructs and the Renilla luciferase pGL4.75 vector (Promega), as a control for transfection efficiency. Forty-eight hours after transfection, luciferase activity was measured (Dual-Glo™ Luciferase Assay System, Promega) and firefly luciferase intensities for enhancer regions were normalized to Renilla values for each sample. Background levels of firefly luciferase activity were assayed by transfecting cells with a control plasmid, where a non-coding 2.2kb stretch of plasmid sequence (taken from pENTR1A plasmid, Invitrogen) was cloned into the pGL3-promoter vector. These values were further used to normalise test luciferase intensities.

3C was done following the protocol of Miele *et al* [21]. About 10^7 cells (90-100% confluence) were cross-linked with 1% formaldehyde for 20 min and quenched with 125 mM Glycine.

After shaking the cells for more than 30 mins in lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.5 % NP-40, 1 % Triton X-100 and protease inhibitors) nuclei were pelleted (1800 rpm, 5 min) and resuspended in 360 μl H2O and 60 μl of the 10x restriction buffer for BglII. After 1h incubation in 3.75 % SDS and subsequent quenching with 37.5 % triton X-100 400-800U of BglII was added and incubated overnight at 37°C. The ligation reaction used a 10X higher volume to favour ligation events between cross-linked DNA strands using 50 U T4 Ligase (Roche, 5U/μl) and incubation at 16°C over 1-2 nights. To de-crosslink the samples, 30μl Proteinase K (10 mg/ml) was added and incubated over night at 65°C. After phenol extraction and ethanol precipitation the samples were recovered in 500 μl TE. 4μl of elute was used per qRT-PCR, using TaqMan Fast Universal PCR System (Applied Biosystems). Data were normalized to crosslinking frequencies in BAC RP11-420B6. The ratio of the test data points to the adjacent upstream primer data was used to compare results between cell lines.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. Whitelaw SC, et al. Clinical and molecular features of the hereditary mixed polyposis syndrome. Gastroenterology. 1997; 112:327–34. [PubMed: 9024286]

2. Jaeger EE, et al. An ancestral Ashkenazi haplotype at the HMPS/CRAC1 locus on 15q13-q14 is associated with hereditary mixed polyposis syndrome. Am J Hum Genet. 2003; 72:1261–7. [PubMed: 12696020]

3. Tomlinson I, et al. Inherited susceptibility to colorectal adenomas and carcinomas: evidence for a new predisposition gene on 15q14-q22. Gastroenterology. 1999; 116:789–95. [PubMed: 10092300]

4. Jaeger E, et al. Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk. Nat Genet. 2008; 40:26–8. [PubMed: 18084292]

5. Cheah PY, et al. Germline bone morphogenesis protein receptor 1A mutation causes colorectal tumorigenesis in hereditary mixed polyposis syndrome. Am J Gastroenterol. 2009; 104:3027–33. [PubMed: 19773747]

6. East JE, Saunders BP, Jass JR. Sporadic and syndromic hyperplastic polyps and serrated adenomas of the colon: classification, molecular genetics, natural history, and clinical management. Gastroenterol Clin North Am. 2008; 37:25–46. v. [PubMed: 18313538]

7. Colella S, et al. QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. Nucleic Acids Res. 2007; 35:2013–25. [PubMed: 17341461]

8. Tomlinson IP, et al. COGENT (COlorectal cancer GENeTics): an international consortium to study the role of polymorphic variation on the risk of colorectal cancer. Br J Cancer. 2010; 102:447–54. [PubMed: 19920828]

9. Kosinski C, et al. Gene expression patterns of human colon tops and basal crypts and BMP antagonists as intestinal stem cell niche factors. Proc Natl Acad Sci U S A. 2007; 104:15418–23. [PubMed: 17881565]

10. Segditsas S, et al. Putative direct and indirect Wnt targets identified through consistent gene expression changes in APC-mutant intestinal adenomas from humans and mice. Hum Mol Genet. 2008; 17:3864–75. [PubMed: 18782851]

11. Tomlinson IP, et al. Multiple common susceptibility variants near BMP pathway loci GREM1, BMP4, and BMP2 explain part of the missing heritability of colorectal cancer. PLoS Genet. 2011; 7:e1002105. [PubMed: 21655089]

12. Tian Q, He XC, Hood L, Li L. Bridging the BMP and Wnt pathways by PI3 kinase/Akt and 14-3-3zeta. Cell Cycle. 2005; 4:215–6. [PubMed: 15655376]

13. Hobert JA, Eng C. PTEN hamartoma tumor syndrome: an overview. Genet Med. 2009; 11:687–94. [PubMed: 19668082]

14. Venkatachalam R, et al. Identification of candidate predisposing copy number variants in familial and early-onset colorectal cancer patients. Int J Cancer. 2010; 129:1635–42. [PubMed: 21128281]

15. Papaemmanuil E, et al. Deciphering the genetics of hereditary non-syndromic colorectal cancer. Eur J Hum Genet. 2008; 16:1477–86. [PubMed: 18628789]

16. Zuniga A, et al. Mouse limb deformity mutations disrupt a global control region within the large regulatory landscape required for Gremlin expression. Genes Dev. 2004; 18:1553–64. [PubMed: 15198975]

17. Howe JR, et al. Germline mutations of the gene encoding bone morphogenetic protein receptor 1A in juvenile polyposis. Nat Genet. 2001; 28:184–7. [PubMed: 11381269]

18. Howe JR, et al. Mutations in the SMAD4/DPC4 gene in juvenile polyposis. Science. 1998; 280:1086–8. [PubMed: 9582123]

19. Lo HS, et al. Allelic variation in gene expression is common in the human genome. Genome Res. 2003; 13:1855–62. [PubMed: 12902379]

20. Poulsom R, Longcroft JM, Jeffery RE, Rogers LA, Steel JH. A robust method for isotopic riboprobe in situ hybridisation to localise mRNAs in routine pathology specimens. Eur J Histochem. 1998; 42:121–32. [PubMed: 9728289]

21. Miele A, Gheldof N, Tabuchi TM, Dostie J, Dekker J. Mapping chromatin interactions by chromosome conformation capture. Curr Protoc Mol Biol. 2006:11. **Chapter 21**, Unit 21. [PubMed: 18265379]
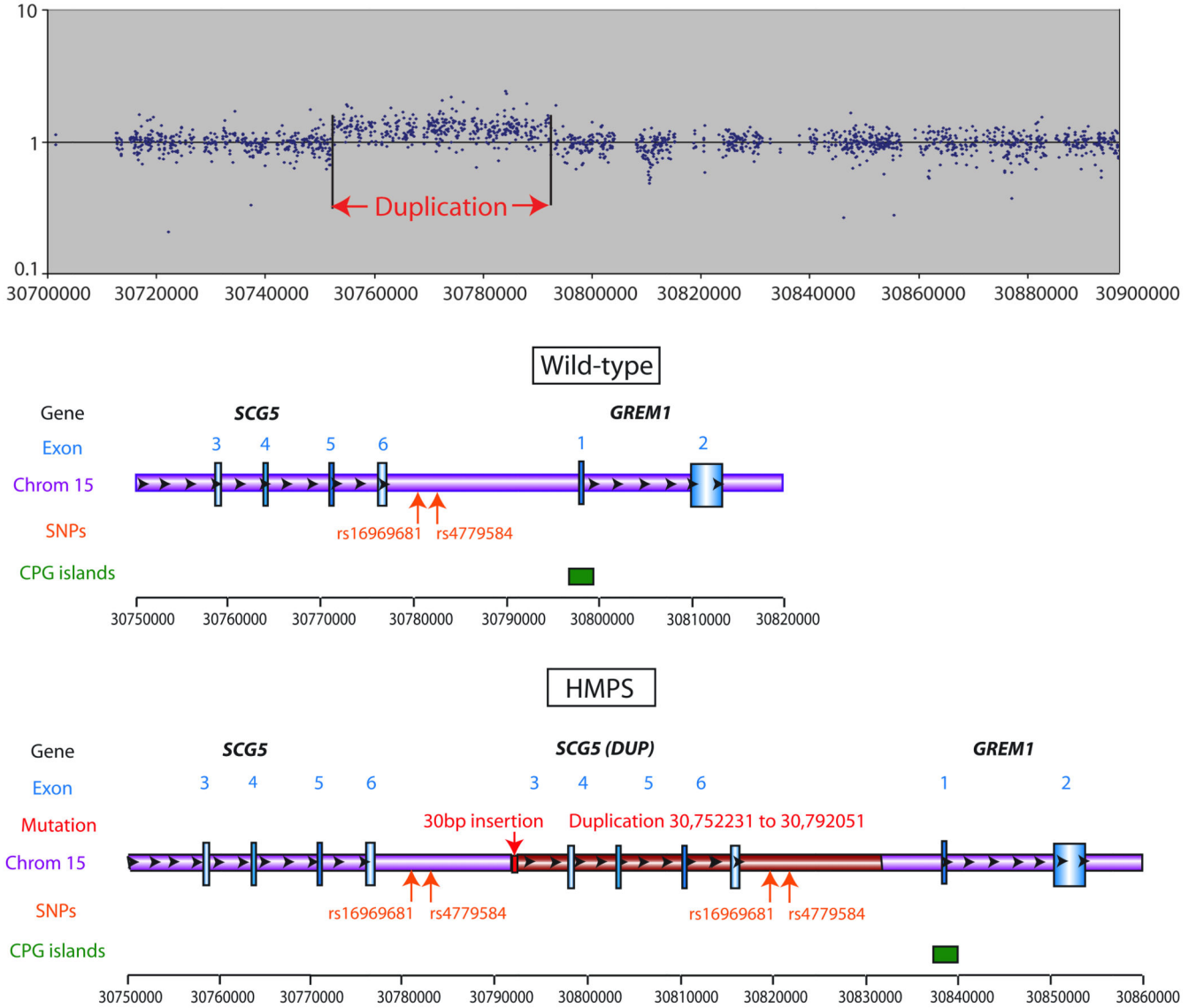
Figure 1. The HMPS duplication

A, array-CGH analysis of the region around the HMPS duplication. The region of copy number gain, estimated to be 3:2 ratio relative to controls and hence resulting from a single-copy gain, is indicated. Results from the full region analysed are shown in Supplementary Figure 1.

B, schematic of the 40kb duplication showing the involvement of the 3′ half of SCG5 extending to a region upstream of GREM1.The locations on chromosome 15 of coding sequences, introns, selected SNPs and CpG islands are shown. Arrows show direction of transcription.
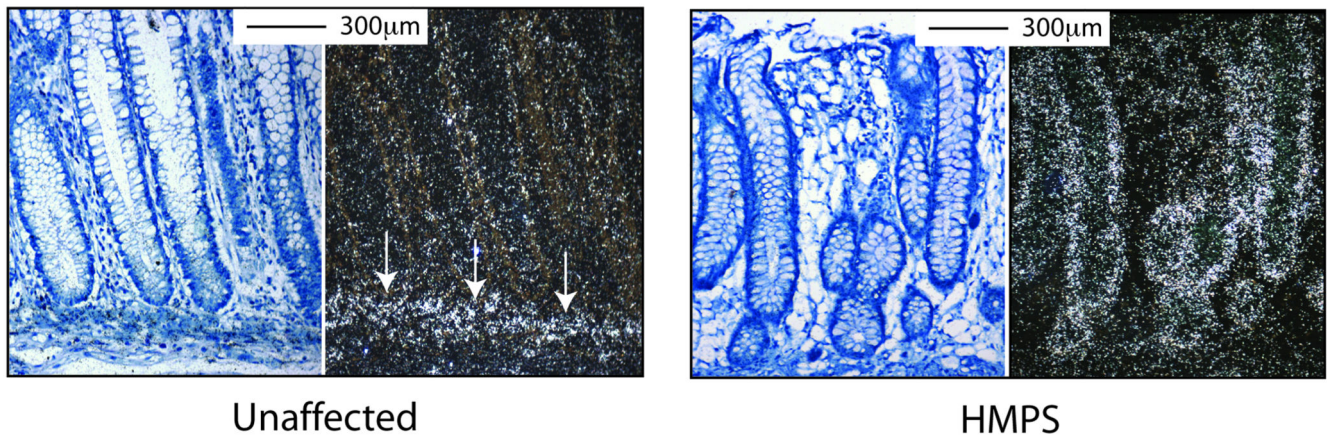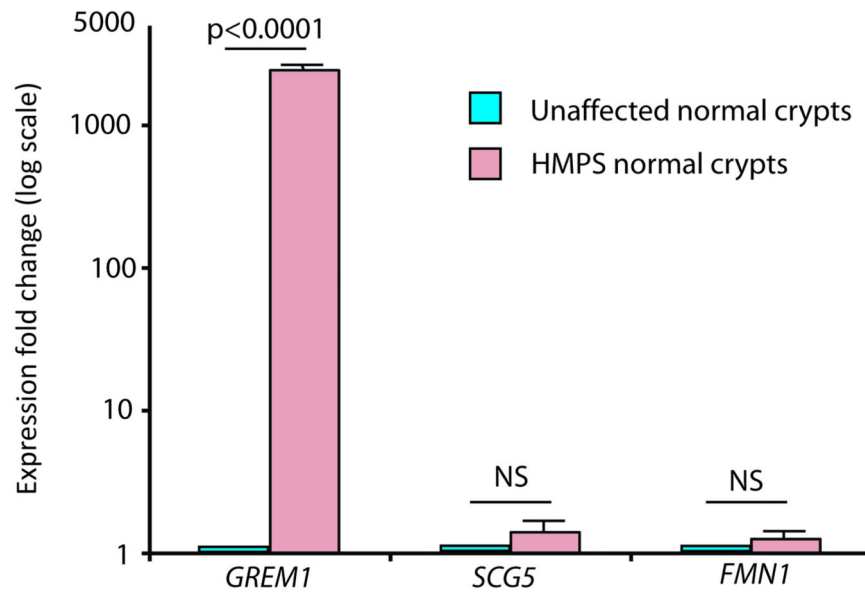
Figure 2. Expression of GREM1, SCG5 and FMN1
A, expression levels of each mRNA by qRT-PCR analysis of multiple individual crypts of normal morphology from HMPS patients relative to controls. Error bars show standard deviation. Note the massively increased expression of *GREM1*. Immunohistochemistry additionally provided no evidence of differences in SCG5 expression between HMPS patients and controls (not shown); reliable antibodies do not exist for GREM1 and FMN1. B, Representative light- and dark-field mRNA *in situ* hybridisation of *GREM1* mRNA in colorectum of normal appearance. *GREM1* is restricted to ISEMFs and the subcryptal smooth muscle layer in control individuals (left), but is found throughout the crypt epithelium in HMPS patients (right).
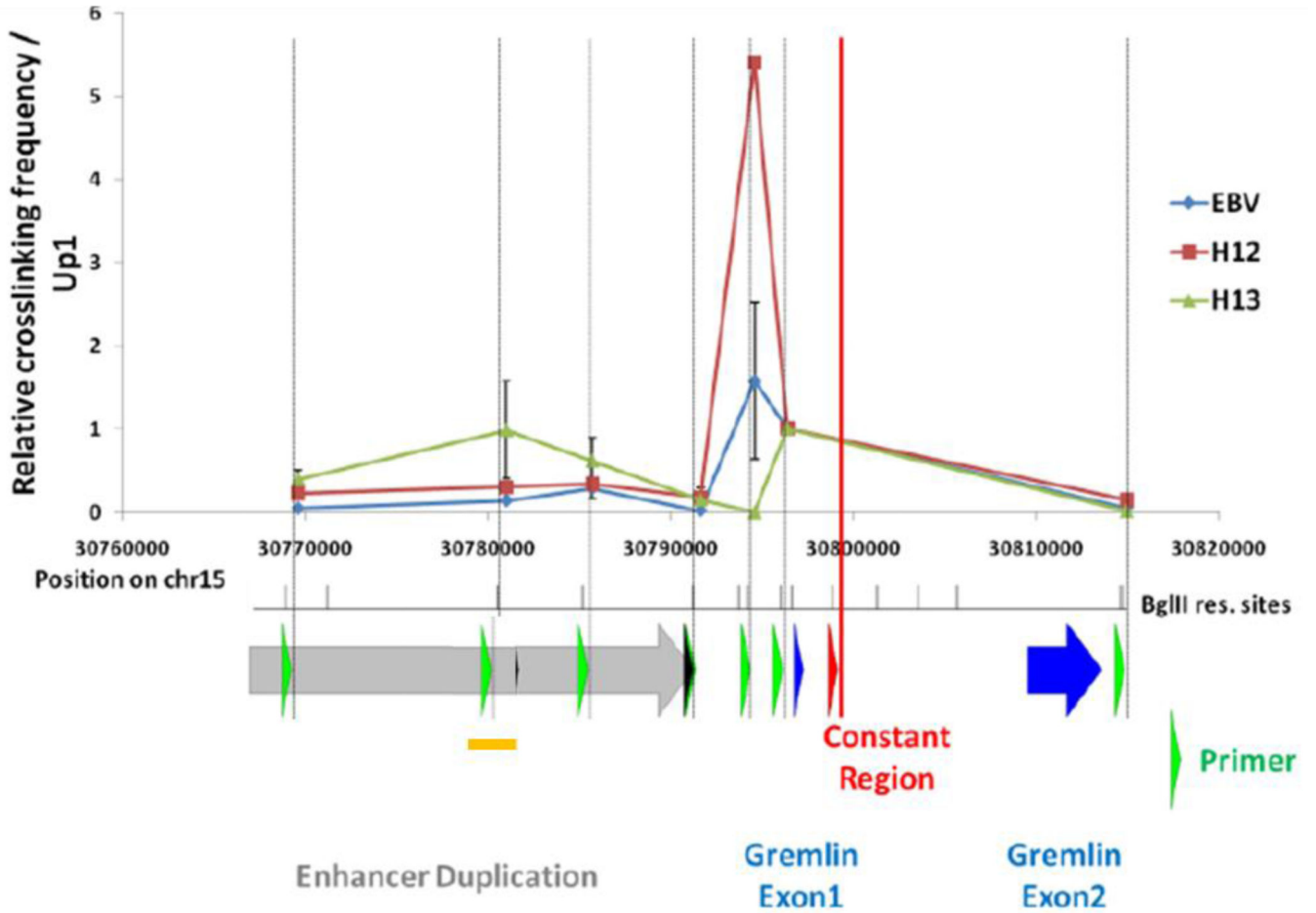
Figure 3. 3C analysis of a region upstream of GREM1 in somatic cell hybrids carrying the duplicated and "wildtype" copies of chromosome 15

The genomic region of GREM1 is depicted below the x-axis, with the telomeric limit of the duplication shown in grey. The 3kb region within the duplication on which we focussed is shown as an orange bar. The 3C constant region and its primer/probe location (red) are located on a BglII fragment containing the transcription start site and exon1 of *GREM1*. Variable human-specific primers (green) are located at the 3′- end of selected fragments. The relative cross-linking frequencies of each upstream fragment derived from qRT-PCR were normalised to BAC control and calculated as the ratio of individual site cross-linking frequencies to the cross-linking frequency of first upstream adjacent BglII fragment. The duplication hybrid cell line, H13, showed a significant and specific increase in the interaction of the 3kb region with the *GREM1* promoter. The H12 line, containing the "wildtype" human duplication, also showed some interaction between these sequences compared with the comparison EBV cell lines that do not express *GREM1*.