Gastric adenocarcinoma and proximal polyposis of the stomach (GAPPS) is an autosomal-dominant cancer-predisposition syndrome with a significant risk of gastric, but not colorectal, adenocarcinoma. We mapped the gene to 5q22 and found loss of the wild-type allele on 5q in fundic gland polyps from affected individuals. Whole-exome and -genome sequencing failed to find causal mutations but, through Sanger sequencing, we identified point mutations in APC promoter 1B that co-segregated with disease in all six families. The mutations reduced binding of the YY1 transcription factor and impaired activity of the APC promoter 1B in luciferase assays. Analysis of blood and saliva from carriers showed allelic imbalance of APC, suggesting that these mutations lead to decreased allele-specific expression in vivo. Similar mutations in APC promoter 1B occur in rare families with familial adenomatous polyposis (FAP). Promoter 1A is methylated in GAPPS and sporadic FGPs and in normal stomach, which suggests that 1B transcripts are more important than 1A in gastric mucosa. This might explain why all known GAPPS-affected families carry promoter 1B point mutations but only rare FAP-affected families carry similar mutations. The colonic cells usually being protected by the expression of the 1A isoform. Gastric polyposis and cancer have been previously described in some FAP-affected individuals with large deletions around promoter 1B. Our finding that GAPPS-affected families carry similar mutations, the colonic cells usually being protected by the expression of the 1A isoform, Gastric polyposis and cancer have been previously described in some FAP-affected individuals with large deletions around promoter 1B. Our finding that GAPPS is caused by point mutations in the same promoter suggests that families with mutations affecting the promoter 1B are at risk of gastric adenocarcinoma, regardless of whether or not colorectal polyps are present.

Introduction

One of the key features of gastric adenocarcinoma and proximal polyposis of the stomach (GAPPS) is a carpeting of more than 100 fundic gland polyps (FGPs) in the oxyntic mucosa of the gastric body and fundus, with antral sparing, and with some FGPs showing dysplasia. This is in contrast to benign, sporadic FGPs that are fewer in number within an individual, and in which high-grade dysplasia is extremely rare. Although low-grade dysplasia has been reported in a small percentage of individuals with sporadic FGPs, there has been only one case report of
from FGPs in individuals with FAP. 21–23 are occasional reports of gastric adenocarcinoma arising
proton pump inhibitor therapy. 9–11 might be more common in individuals who have received
uals undergoing upper gastrointestinal endoscopy 6–8 and
cinomas. Sporadic FGPs are identified in ~5% of individ-
als undergoing upper gastrointestinal endoscopy 6–8 and
might be more common in individuals who have received proton pump inhibitor therapy. 9–11

The gastric antrum, pylorus, small intestine, and colon were all reported to be normal in the original GAPPS-
affected families, 1 highlighting a clear difference between this condition and familial adenomatous polyposis
(FAP [MIM: 175100]) syndrome. FAP 12–16 and attenuated FAP 17 (AFAP) are autosomal-dominant conditions charac-
terized by the development of multiple adenomatous polyps in the colorectum as well as extra-colonic manifes-
tations, due to germline coding mutations or large deletions or duplications in APC (MIM: 622731) (adenomatous
polypsis coli). Gastric FGPs have also been observed in 12%–84% of individuals with FAP, with around 50% of FAP-
and AFAP-affected individuals having more than 100 FGPs (D.W.N., unpublished data). FAP-associated
FGPs tend to be more numerous within an individual than sporadic FGPs, and carpeting of FGPs in the gastric
body and fundus has been observed in FAP and AFAP. 10, 15 Dysplasia is low grade in the majority of FAP-
associated FGPs (96%), displaying an exclusively gastric (foveolar) phenotype in 92% of the case subjects. 20 There
are occasional reports of gastric adenocarcinoma arising from FGPs in individuals with FAP. 21–23

Studies by Abraham et al. on the APC/β-catenin pathway in FGPs suggest that different mutations are present
in FAP-associated and sporadic FGPs. 24–26 Activating CTNNB1 (MIM: 116806) somatic mutations occur in
91% of sporadic FGPs, but not in any FAP-associated
FGPs, 25 whereas FAP-associated FGPs frequently harbor
somatic coding APC gene mutations. However, sporadic
FGPs with low-grade dysplasia are molecularly more
similar to FAP-associated FGPs in the type and frequency of
APC alterations than to the sporadic FGPs without
dysplasia. 26 The distribution of somatic mutations of
APC in colorectal adenomas is non-random and partly
related to the site of the germline mutation but differs
from the non-random pattern in duodenal and fundic gland polyps. 27, 28

Here we report three different point mutations in the promoter 1B of APC that are responsible for GAPPS in
all six families. We also demonstrate that these mutations are located within a Ying Yang 1 (YY1) binding
motif and reduce the expression from the promoter 1B by interrupting YY1 binding. Additionally, we report
that rare families with extensive FGPs, but with a more
classical FAP presentation in the colon, harbor one of
two point mutations in the same YY1 binding site in pro-
moter 1B.

Subjects and Methods

GAPPS-Affected Families
We identified one Australian (family 1) and five North American (families 2–6) families that meet the diagnostic criteria for
GAPPS, 1 by means of autosomal-dominant transmission of
numerous, predominantly fundic gland, gastric polyps restricted
to the body and fundus with regions of dysplasia or gastric adenocarci-
na, and no evidence of colorectal or duodenal polyposis (Figure 1). Families 1–3 were described previously, 1 but at that
time individual II-4 in family 3 was considered to be affected.
However, more detailed phenotyping found that she has fewer than 30 FGPs with no evidence of dysplasia and is therefore now considered unaffected (Figure S1). Since our previous publication, I-5 in family 3 has developed gastric cancer. In family 4, which had no case of gastric adenocarcinoma, at least one fundic gland polyp analyzed by a gastrointestinal pathologist (F.C.) contained areas of low-grade dysplasia with some focal high-grade dysplasia. All subjects included in this analysis were of European descent and provided written informed consent as well as data and blood samples under ethically approved protocols.

Figure 1. GAPPS Pedigrees
Abbreviations are as follows: POSITIVE/neg, APC promoter 1B mutation status by Sanger sequencing; WES, whole-exome sequencing; WGS, whole-genome sequencing by HiSeq or Complete Genomics; WGS-X Ten, whole-genome sequencing by X Ten. Only affected individuals and carriers are shown.
(A) Family 1. Individuals and clinical presentation are as follows: I-4, unconfirmed gastric cancer; II-2, normal endoscopy aged 72 years; II-4, normal endoscopy aged 77 years; II-6, no endoscopy; II-11, normal endoscopy aged 68 years; III-33, normal endoscopy aged 60 years; IV-24, normal endoscopy aged 42 years. Thirty or fewer fundic gland polyps were observed in these non-carriers: III-22, 17 FGPs; III-23, 15 FGPs; IV-13, 30 FGPs; III-21, 7 FGPs; III-25, 25 FGPs. In two polyps from individuals III-7 and IV-4 where parental origin could be evaluated, the wild-type haplotype was lost. Asterisk (*) indicates sample was included in the linkage analysis.
(B) Family 2. I-4, non-melanoma skin cancer, unconfirmed colon or gastric cancer; III-21, thyroid cancer (and fundic gland polyposis).
(C) Family 3. I-2, leukemia; I-3, brain cancer; I-5, antral gastric cancer, no gastric polyps; I-6, prostate cancer; I-7, lung cancer.
(D) Family 4.
(E) Family 5.
(F) Family 6. II-1, prostate cancer; II-2, ovarian cancer and unconfirmed colon cancer; II-3, unconfirmed colon cancer; III-1, unconfirmed colon cancer; IV-21, bone cancer, gastro-intestinal cancer.
Linkage Analysis
We performed linkage analysis of family 1 using 39 germline DNA samples available at the time with the Human Linkage-12 SNP panel (Illumina; 6,090 SNPs) according to the manufacturer’s protocol. We carried out linkage analyses using an affected-only Markov chain-Monte Carlo method (LNMARKERS)29 modeling GAPPs as an autosomal-dominant trait with an allele frequency of 0.001 and a penetrance of 80%. Only case subjects with florid gastric polyposis verified by endoscopy were included as affected.

Next-Generation Sequencing
We carried out whole-exome sequencing (WES) in 19 members of families 1–5 (Figure 1) using Agilent SureSelect and/or Illumina TruSeq Exome Enrichment kit, followed by 75 bp or 100 bp paired-end sequencing on Illumina HiSeq (Table S1). We obtained whole-genome sequencing (WGS) data on the Illumina HiSeq platform from four case subjects of families 1 and 2 and with Complete Genomics for additional members of family 2 (Figure 1). We also performed WGS with the Illumina X Ten platform for seven fundic gland polyps from family 1, with the matching germline DNA. We carried out sequence alignment, with BWA,30 and variant calling for SNVs via GATK 31 and qSNP 32 and small insertion and deletions via pindel.33 We validated rare variants of polyposis verified by endoscopy were included as affected.

Copy-Number Analysis
Fourteen fundic gland polyps from six members of family 1 were obtained during gastrectomies and snap frozen. “Book end” sections stained with hematoxylin and eosin (H&E) were reviewed by a gastrointestinal pathologist (M.B.) to estimate the percentage of FGP. Only sections estimated to contain at least 70% FGP were subjected to DNA extraction via the QIAamp DNA Micro kit (QIAGEN). DNA was hybridized to the Illumina Human Omni2.5-8 v1.1 SNP arrays. Copy-number changes were called with the Genome Alteration Print (GAP) tool35 and copy number, B allele frequency, and LogR ratio was visualized with Circos.36 In order to determine which allele was lost, we used 55 SNPs in the region that we could phase because they were heterozygous in individuals III-7 and IV-4 but homozygous in other family members.

Sanger Sequencing of APC Promoter 1B
The APC promoter 1B was amplified by AmpliTaq Gold polymerase (Life Technologies) using Touchdown-PCR with the following conditions: 94\(^\circ\)C for 12 min, 94\(^\circ\)C for 30 s, 68\(^\circ\)C for 30 s in the first cycle and the anneal temperature decreased 1\(^\circ\)C per cycle, 72\(^\circ\)C for 30 s; after 12 cycles conditions changed to 94\(^\circ\)C for 30 s, 55\(^\circ\)C for 30 s, 72\(^\circ\)C for 30 s for another 25 cycles, final extension of 72\(^\circ\)C for 7 min. The PCR product was sequenced using BigDye Terminator v.3.1 (Life Technologies) according to the standard protocol. Primers are listed in Table S2.

Gastric Carcinoma Samples
As part of a prospectively collected cohort of sporadic gastric adenocarcinomas, we tested a subgroup of 26 cases of gastric adenocarcinoma comprising 13 diffuse gastric cancers, 12 intestinal gastric cancers, and 1 mixed gastric cancer as classified by Lauren77 for mutations in the APC promoter 1B. For each sample, 10 \(\mu\)M sections were macrodissected to increase tumor purity and DNA extracted with the DNA QIAamp DNA Mini kit (QIAGEN). DNA was extracted from blood samples via the salt extraction method to obtain matched germline DNA.

Plasmid Generation
A 500 bp fragment (chr5: 112,042,880–112,043,379) containing part of the APC promoter 1B region (GenBank: D13981.1) was amplified from genomic DNA of affected individuals from families 1, 2, and 3. To separate c.–195A>C and c.–125delA, which co-segregate in family 1, we designed two sets of mutagenic oligonucleotides for each allele of c.–195A>C. After the first PCR step, two fragments overlapping by 20 bases were fused together by an overlapping extension PCR. DNA fragments containing c.–195A>C plus c.–125delA, c.–195A>C alone, c.–125delA alone, c.–191T>C and c.–192A>G, or wild-type sequence were subsequently cloned into the pGL3 basic luciferase reporter construct. PCR primers are listed in Table S2.

Reporter Assays
Cell lines were maintained under standard conditions, routinely tested for Mycoplasma, and identity profiled with short tandem repeat markers. AGS and MKN74 were provided by Dr. Andrew Giraud and HCT116 and RKO by Dr. Vicki Whitehall. Two gastric carcinoma cell lines, MKN74 and AGS, and two colorectal cancer cell lines, HCT116 and RKO, were transiently transfected with equimolar amounts of luciferase reporter constructs using Renilla luciferase as an internal control reporter. Luciferase activity was measured 24 hr after transfection using Dual-Glo Luciferase (Promega). Firefly luciferase activity was normalized to Renilla luciferase, and the activity of each mutant construct was measured relative to wild-type construct, which had a defined activity of 1. Expression differences were assessed by log transforming the data and performing two-way ANOVA, followed by Dunnett’s multiple comparisons test; for ease of interpretation, values were back-transformed to the original scale for the graphs.

Electrophoretic Mobility Shift Assay
Nuclear lysates were extracted from a gastric and colorectal cancer cell line (AGS and RKO, respectively) as described previously38 EMSAs were carried out with the Lightshift Chemiluminescent EMSA Kit (Thermo Scientific). Oligonucleotide sequences used in the assays are listed in Table S2. Competitor oligonucleotides were used at 10-, 30-, and 100-fold molar excess. For gel-supershift assays, 5 \(\mu\)g of rabbit polyclonal YY1 antibody (Santa Cruz cat# sc-1703; RRID: AB_2218501) was added immediately before probe addition. The rabbit pre-immune IgG (Santa Cruz cat# sc-2027; RRID: AB_737197) was used as a negative control.

Chromatin Immunoprecipitation
YY1 ChIP-qPCR (YY1; Santa Cruz cat# sc-1703; RRID: AB_2218501) assays were conducted as described previously39 with a sheared fragment size of 300 bp to 1 kb. For qPCR, 1 \(\mu\)L from 30 \(\mu\)L of immunoprecipitated DNA extract was used. Primers are listed in Table S2.

APC Allelic Imbalance Analysis
We obtained fresh blood samples in PAXgene Blood RNA tubes (QIAGEN) from two affected and one unaffected mutation carriers and from two unaffected control subjects (who did not carry the
related individuals were sequenced, we found multiple rare coding mutations were shared between affected members. The mean coverage ranged from 58- to 68-fold, no novel or rare (minor allele frequency < 0.001) missense, splice site, or coding indel mutations that co-segregated with disease in all families. Furthermore, although the mean coverage ranged from 58- to 68-fold, no novel or rare coding mutations were shared between affected members of family 1. In the smaller families in which closely related individuals were sequenced, we found multiple shared coding variants throughout the genome. We also carried out whole-genome sequencing (WGS) to a mean coverage of 17- to 75-fold in two affected members of family 1 and five affected members and three unaffected members, including a spouse, from family 2. This did not identify any novel non-coding variants in the linkage region shared between all affected case subjects from both families. However, we did find 202 putative non-coding variants shared between both sequenced members of family 1. Of these, the majority were calls of low confidence but 67 were worthy of follow up and amenable to iPLEX design. We genotyped 63 family members and 369 Australian control subjects. Sixteen variants were not validated and 18 did not co-segregate with affected status or were present in control subjects, but the remaining 33 variants did co-segregate and further defined the linkage region to 14.3 Mb (chr5: 107,080,146–121,407,036). None of these variants were in obvious regulatory regions as indicated by ENCODE, Roadmap, or FANTOM5 (see Web Resources) or in previously unannotated genes (using the latest version of GENCODE). We also found 42 non-coding variants shared between the five affected members of family 2, but none were close to the non-coding variants found in family 1 (the closest from the two families were 433 bp apart). Refiltering against dbSNP142 reduced the number of variants in family 1 and family 2 to 22 and 30, respectively (Table S3).

We isolated gDNA and RNA from cryostored normal stomach and FGP-positive samples from family 1 individuals using the TissueLyser and AllPrep DNA/RNA Mini Kit (Qiagen), including DNase treatment step to remove gDNA contamination of RNA. cDNA was transcribed with SuperScript IV Reverse Transcriptase (Life Technologies) and Sanger sequenced to look for allelic imbalance. Allelic imbalance was assessed by comparing cDNA and genomic DNA sequence traces (primer sequences in Table S2).

We isolated gDNA and RNA from cryostored normal stomach and FGP-positive samples from family 1 individuals using the TissueLyser and AllPrep DNA/RNA Mini Kit (Qiagen), including DNase treatment step to remove gDNA contamination of RNA. cDNA was transcribed with SuperScript IV Reverse Transcriptase (Life Technologies) and Sanger sequenced to look for allelic imbalance of c.−195A>C mutation (primer sequences in Table S2). Controls with no reverse transcriptase added were negative, indicating no gDNA contamination of RNA/cDNA. gDNA and RNA was also isolated from a cryostored FGP sample from individual II-3 in family 3 via the DNeasy Blood and Tissue Kit (Qiagen) and miRNeasy Mini Kit (Qiagen). cDNA was transcribed with SuperScript III Reverse Transcriptase (Life Technologies) and Sanger sequenced to look for allelic imbalance of the c.−192A>G mutation (primer sequences in Table S2).

Results

We have identified six families affected by GAPPs (Figure 1).

Linkage Analysis

Linkage analysis in 39 members of family 1 (Figure 1A) mapped the GAPPs mutation to a 46 Mb region on chromosome 5 (chr5: 75,947,905–121,407,036) with a maximum LOD score of 4.51.

Whole-Exome and -Genome Sequencing

Whole-exome sequencing (WES) in 15 affected and 4 unaffected members of families 1–5 (Table S1) failed to identify any genes in the linked region of chromosome 5 with novel or rare (minor allele frequency < 0.001) missense, splice site, or coding indel mutations that co-segregated with disease in all families. Furthermore, although the mean coverage ranged from 58- to 68-fold, no novel or rare coding mutations were shared between affected members of family 1. In the smaller families in which closely related individuals were sequenced, we found multiple
Sanger Sequencing

Protein truncating and large deletion mutations of APC cause familial adenomatous polyposis (FAP). Notwithstanding their different colonic involvement, GAPPS and FAP have similar gastric phenotypes of polyposis and predisposition to carcinoma. We therefore looked for allelic imbalance (AI) in the expression of APC in blood from affected members of family 1 as an indicator of an inherited regulatory mutation (described below). In seeking an informative polymorphism for AI analysis, we Sanger sequenced the two discrete APC promoters (Figure 3). Sequencing identified two mutations (a point mutation, chr5: 112,043,220 A/C; c.195A>C, and a 1-bp deletion, chr5: 112,043,290 A/C; c.125delA) in promoter 1B that completely co-segregated with GAPPS in all 27 affected members of family 1 from which we had blood samples (Figures 1, S3A, S3B, S3E, and S3F). These variants are in close proximity to a variant identified in family 2 by Complete Genomics (chr5: 112,043,224 T/C; c.191T>C) (Table S3). Additional sequencing of promoter 1B revealed this mutation (chr5: 112,043,224 T/C; c.191T>C) in all the affected members of families 2, 4, 5, and 6 (Figures 1, S3A, and S3C) and another (chr5: 112,043,223 A/G; c.192A>G) in both affected members of family 3 (Figures 1, S3A, and S3D). These three mutations were not reported in any public databases, including 2,598 samples from 1000 Genomes,51 nor are they present in 344 germline samples from our in-house WGS cancer projects,52–54 nor did we find any c.195A>C carriers in 2,326 Australian control samples genotyped by iPLEX. We genotyped 20/30 of the rare non-coding variants identified in family 2 in families 4–6 and found that family 4 carried the same rare variants from chr5: 108,498,647 to chr5: 115,074,561 (Table S3), suggesting that they share a common ancestor.

Re-exploration of previously reported FAP-affected families55 revealed that the single mutation reported in affected probands from eight probably related families from the same region of France, g.20377206A>T, is by more recent nomenclature a different base substitution (c.192A>T) at the same nucleotide that is mutated in family 3 (c.192A>G). Although FGPs are prominent in these families, all probands and many family members have had colectomies for florid colonic polyposis. At the same time, a family with five affected persons with profuse FGPs and a variable onset of colorectal polyposis, resulting in colectomy in all mutation carriers between the ages of 4 and 57 (Figure S4), was found to carry a c.190G>A mutation in promoter 1B.

Somatic Mutations and Methylation

We carried out additional WGS of seven FGPs without LOH (Figure S2), with their matching germline DNAs, in family 1 (using Illumina’s X Ten platform). This experiment, initiated before we found the germline promoter 1B mutations, confirmed the results of Sanger sequencing (Figure S5). In the region of linkage/LOH, we also found four somatic truncating mutations of APC—c.637C>T (p.Arg213Ter) (GenBank: NM_001127510, NP_001120982), c.1690C>T (p.Arg564Ter), c.6096_6102delTGACTCT (p.Ile2032Metfs*10), and c.4393_4394delAG (p.Ser1465Trpfs*3)—at estimated mutant allele frequencies of 31%, 9%, 8%, and 12%, respectively.

Figure 2. Genomic Copy-Number Changes of Six Gastric Fundic Gland Polyps

Whole-genome copy-number plots from copy-number analysis of each polyp represented by log R ratio and B allele frequency. Deviation of B allele frequency suggests allelic imbalances and copy-number variation at 5q. Genomic SNP data of the remaining eight polyps showing no LOH are presented in Figure S2.
respectively, in FGPs. Three of the somatic mutations were confirmed by iPLEX (Figure S6), but we could not design iPLEX primers for p.Ser1465Ilefs*3. No rare or novel germline coding mutations in the linkage/LOH region were identified by this additional X Ten WGS, nor did we find any mutations in CTNNB1 that are characteristic of sporadic FGP.s\textsuperscript{25} After microdissection to achieve >90% neoplasia in the sections used for DNA isolation, we sequenced the APC promoter 1B in 26 sporadic gastric cancer (GC) cases (including 13 diffuse, 12 intestinal, and 1 mixed-type GC) with their matching germline DNAs but found no germline or somatic mutations.

Sequencing the canonical exon 3 CTNNB1 hotspot mutations in two FGPs from individual II-4 of family 3 (a non-carrier of the familial promoter mutation) and seven sporadic FGPs from five individuals revealed hotspot mutations in both polyps from II-4 and in three of the sporadic polyps (Figure S7). This is in contrast to absence of CTNNB1 hotspot mutations in 11 FGPs from 5 carriers of promoter mutations in family 1 (data not shown).

Methylation of APC promoter 1A was demonstrated in both normal stomach and FGP’s of carriers and non-carriers within GAPPs-affected families and in sporadic FGP’s and matched normal stomach (Table S5).

Functional Analysis of 1B Promoter Mutations

Transcriptional factor binding site (TFBS) search tools\textsuperscript{56,57} predicted Yin Yang 1 (YY1) binding to the chr5:112,043,220–112,043,224 region of promoter 1B, which JASPAR\textsuperscript{58} analysis predicted would be disrupted by the three GAPPs and two FAP mutations (Figure 4A). Via chromatin immunoprecipitation (ChIP), we showed that YY1 binds this site in two gastric cancer cell lines (AGS and MKN74 cells) and two colorectal cancer lines (HCT116 and RKO) (Figures 4B and S8). Using electrophoretic mobility shift assays (EMSA), we showed that the mutant alleles of GAPPs mutations (c.−191T>C, c.−192A>G, and c.−195A>C) and the FAP mutations (c.−190G>A and c.−192A>T) disrupted protein binding to the region in both AGS and RKO cells (Figures 4C and 4D). Furthermore, EMSAs using anti-YY1 antiserum or a YY1 consensus oligonucleotide competitor suggested that the protein disrupted by the mutant alleles is likely to be YY1 (Figures 4E and S9). Luciferase reporter assays demonstrated that constructs containing c.−195A>C plus c.−125delA, c.−195A>C alone, c.−191T>C, and c.−192A>G showed significantly decreased activity compared to the wild-type construct in MKN74, AGS, and RKO cells (Figures 5 and S10). In addition, the construct containing both c.−195A>C and c.−125delA showed reduced expression in HCT116 colorectal cancer cell line. The c.−125delA variant alone only showed significantly decreased activity in RKO and HCT116 cells (Figure S10).

Allelic Imbalance

Having found the 1B promoter mutations, we looked for allelic imbalance (AI) of the 1B APC isoform. Promoter 1B transcripts are typically 100- to 1,000-fold more highly expressed than promoter 1A transcripts in blood,\textsuperscript{47} and we inferred that this might extend to transcripts derived from saliva samples because DNA from saliva is mainly derived from lymphocytes.\textsuperscript{59} We therefore used blood from family 1 and saliva from GAPPs-affected family 3 and from the FAP-affected family carrying the c.−190G>A mutation for AI analysis of carriers and non-carriers. There are no informative polymorphisms specific to APC 1B transcripts, so we genotyped a common SNP, rs448475, in 70 family members to determine which allele was in cis with mutation in each individual. We used rs448475 as the marker for allelic imbalance in family 1 because it was heterozygous in multiple affected and unaffected members and is in the 3’ UTR of both transcripts. We obtained fresh blood from two affected carriers, one unaffected mutation carrier, and two unaffected controls (who did not carry the affected haplotype) from family 1, all of whom were heterozygous for rs448475. cDNA sequencing showed AI in the three mutation carriers, with reduced expression of the G allele that was carried on the affected haplotype (Figure 6A). We sequenced seven SNPs in exons 10, 12, and 14 of APC in family 3 to look for AI; rs2229992 was the only informative SNP across both individuals II-3 and II-4. We found AI in individual II-3, who harbors the
promoter c.−192A>G variant, but not in II-4, who is a non-carrier of this mutation (Figure 6B). AI was also observed in saliva from an affected carrier, III-2, of the FAP family, using the promoter c.−190G>A mutation (Figure S11). We also looked for AI in normal stomach and FGPs from five affected carriers from family 1 and in the FGP of an affected carrier of family 3, using the c.−195A>C and c.−192G>A mutations, respectively.

Figure 4. GAPPS and FAP Mutations Alter YY1 Binding to the APC Promoter 1B
(A) Position weight matrix (PWM) of YY1 binding. c.−195A>C (g.112043220C), c.−192A>G (g.112043223G) and c.−192A>T (g.112043223T), c.−191T>C (g.112043224C), and c.−190G>A (g.112043225A) in the APC promoter 1B are predicted to disrupt YY1 binding at positions 1, 4, 5, and 6, respectively.
(B) ChIP-qPCR on the APC promoter 1B region in AGS and MKN74 cells. ChIP assays were performed with a YY1 antibody or IgG control antibody on the region of 1B containing the predicted YY1 binding site. A negative control region was used to represent nonspecific binding. One of two biological replicates is shown. Error bars denote SD.

(C and D) EMSA of promoter 1B using with a biotinylated DNA duplex representing the predicted YY1 binding region with and without (WT) GAPPS mutations c.−195A>C, c.−192A>G, and c.−191T>C (C) and FAP mutations c.−190G>A and c.−192A>T (D).
(E) EMSA-supershift using the WT DNA duplex and a polyclonal antibody against YY1 with AGS nuclear extracts. Rabbit IgG was used as a negative control. The black arrowhead denotes the YY1 supershifted complex.

Figure 5. Mutations in APC Promoter 1B Show Decreased Transcriptional Activity in Gastric Carcinoma Cell Lines AGS and MKN74 and Colorectal Cancer Cell Lines HCT116 and RKO
Error bars denote 95% confidence intervals from three independent experiments performed in triplicate. p values were determined by two-way ANOVA followed by Dunnett’s multiple comparisons test (**p < 0.01, ***p < 0.001, ****p < 0.0001 compared to wild-type pGL3-APC 1B construct) on log transformed data; for ease of interpretation, back-transformed data have been graphed.
which are transcribed. We could detect expression only of the wild-type allele and not the mutant transcript in these samples (Figure S11).

Discussion

We identified three point mutations in the promoter 1B of APC that perfectly segregate with GAPPS in all six families, including one (family 1) with 27 affected individuals (Figure 1). These mutations were missed in the WES and Hi-seq WGS we performed in these families, either because the enrichment kits did not capture the promoter 1B or because of poor coverage (4-11X), but were identified by Complete Genomics, X Ten, and Sanger sequencing. All three mutations were located in a YY1 binding motif, disrupting the affinity of YY1 for the APC promoter 1B and were found to show reduced transcriptional activity in reporter assays in gastric and colorectal cancer cell lines. In carriers from family 1, we observed allelic imbalance of APC in blood, stomach, and FGPs, suggesting that the presence of these mutations leads to decreased allele-specific expression in vivo. Allelic imbalance was also observed in an affected member of family 3 and in an affected carrier of the c.-190G>A mutation from a FAP-affected family. YY1 is a ubiquitously expressed transcription factor that has been shown to have multiple roles in oncogenesis and can act as both an activator and repressor of transcription. Here we show that it acts as a transcription factor regulating expression of APC 1B promoter transcripts.

We also observed a second hit in the majority of GAPPS FGPs, either by loss of the wild-type allele or somatic protein truncating mutations. However, despite using sections composed of 75%–100% polyp cells for DNA isolation, these events appeared to occur in only a small proportion of cells. This suggests that they are late events in the genesis of FGPs. We propose that APC haploinsufficiency is responsible for the fundic gland polyposis in GAPPS, and the second APC hit might be the driver of dysplasia.

Although the APC promoter 1B and distal enhancer elements have not been extensively studied, Hosoya et al. estimated transcription from the promoter 1B was 15-fold higher than from the promoter 1A in gastric mucosa. Consistent with this, the APC promoter 1A (but never promoter 1B) was reported to be methylated in 100% of gastric cancer cell lines, 97.5% of non-cancerous gastric mucosa, and 82.5% of primary gastric cancers. We also observed promoter 1A methylation in GAPPS and sporadic FGPs and in normal stomach. These studies suggest that 1B transcripts are more important than 1A in gastric mucosa, which might explain why all known GAPPS-affected families carry promoter 1B mutations but only rare cases of FAP (1/111; 0.9% W.D.F., unpublished data) carry similar variants, the colonic cells usually being protected by the expression of the 1A isoform. Large (11–132 kb) deletions that include all or some of the promoter 1B, as well as additional sequences, have been observed in FAP and sporadic FGPs and in normal stomach. These studies suggest that 1B transcripts are more important than 1A in gastric mucosa, which might explain why all known GAPPS-affected families carry promoter 1B mutations but only rare cases of FAP (1/111; 0.9% W.D.F., unpublished data) carry similar variants, the colonic cells usually being protected by the expression of the 1A isoform.
occurs in about 50% of individuals with AFAP or FAP. We have found that there is no difference in the presence of FGPs between individuals who have mutations within exons 2 and 7, consistent with an AFAP phenotype, versus those with mutations outside this region, consistent with a classic FAP phenotype (D.W.N., unpublished data). Point mutations in promoters have been observed in other cancer predisposition syndromes, but only once before in the APC 1B promoter. Our finding that GAPPS is caused by point mutations in the promoter 1B of APC, and the description of gastric polyposis and gastric cancer in some individuals with large deletions around the promoter 1B, would suggest that families with alterations in the promoter 1B are at risk of gastric adenocarcinoma, regardless of whether colorectal polyps are or are not present. Conversely, most observations to date indicate low risk for colonic polyposis and cancer in GAPPS, but recognition that GAPPS is due to a variant in the same gene that causes FAP suggests that GAPPS exists along the phenotypic spectrum of FAP, thus broadening the understanding of the disease spectrum caused by APC aberrations. Therefore, in small GAPPS-affected families in which there has been little opportunity to observe the family's phenotype, it might be prudent to undertake colonoscopic surveillance to characterize a family's colon phenotype with a plan for reassessment dictated by the findings. In carriers of these point mutations in the YY1 binding site of the promoter 1B, penetrance of the gastric polyposis phenotype is high, though not complete, with variable phenotype ranging from asymptomatic adults to teenagers presenting with massive symptomatic gastric polyposis. Figure 1 shows five unaffected carriers who had clean endoscopies at ages ranging from 42 to 77. However, the penetrance for gastric cancer is less clear because many of the mutation carriers in family 1 have had prophylactic gastrectomies, often in their 20s. In summary, our data show that specific point mutations in the YY1 binding site in the promoter 1B of APC are the cause of GAPPS, a new and potentially severe variant of FAP.

Accession Numbers
The genotype data are available at the Gene Expression Omnibus under accession number GEO: GSE73547.

Supplemental Data
Supplemental Data include 11 figures and 5 tables and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2016.03.001.

Acknowledgments
We thank all the individuals from the GAPPS-affected families who took part in this study, Andrew Giraud for providing the AGS and MKN74 cell lines, and Vicki Whitehall for HCT116 and RKO. The work was funded by the NHMRC, NIH (PO1-CA073992), and grant MOP-123517 from the Canadian Institutes of Health Research and the Ride to Conquer Cancer. G.J.F. acknowledges the support of an NHMRC Career Development Fellowship (GNT1045237); D.L.W., T.R.M.L., and S.L.W. are funded by Cancer Council South Australia Beat Cancer Project; G.C.-T. and A.B.S. by NHMRC Research Fellowships; and K.A.S. by the Canadian Institutes of Health Research, Michael Smith Foundation for Health Research, and BC Cancer Foundation. W.D.F. thanks Ms. N. Wong for clinical contributions. We acknowledge the support provided by the following pathology labs: Perth Pathology, St John of God Pathology, Western Diagnostics, and PathWest (Royal Perth Hospital, Fremantle Hospital, King Edward Memorial Hospital, and Sir Charles Gairdner Hospital). The authors are very grateful to staff at the Australian Red Cross Blood Services for their assistance with the collection of risk factor information and blood samples of healthy donor control subjects and thank members of the QIMR Molecular Cancer Epidemiology Laboratory for their assistance with recruitment and biospecimen processing. This study was in part (P.S.M., U.R.) supported by the Intramural Research Program of the NIH. O.F.B. has been paid consultancy fees by Amgen and Sanofi Canada.

Received: January 25, 2016
Accepted: March 2, 2016
Published: April 14, 2016

Web Resources
The URLs for data presented herein are as follows:
- ENCODE, https://www.encodeproject.org/

References


59. Thiede, C., Prange-Krex, G., Freiberg-Richter, J., Bornhäuser, M., and Ehninger, G. (2000). Buccal swabs but not mouthwash samples can be used to obtain pretransplant DNA fingerprints from recipients of allogeneic bone marrow transplants. Bone Marrow Transplant. 25, 575–577.


