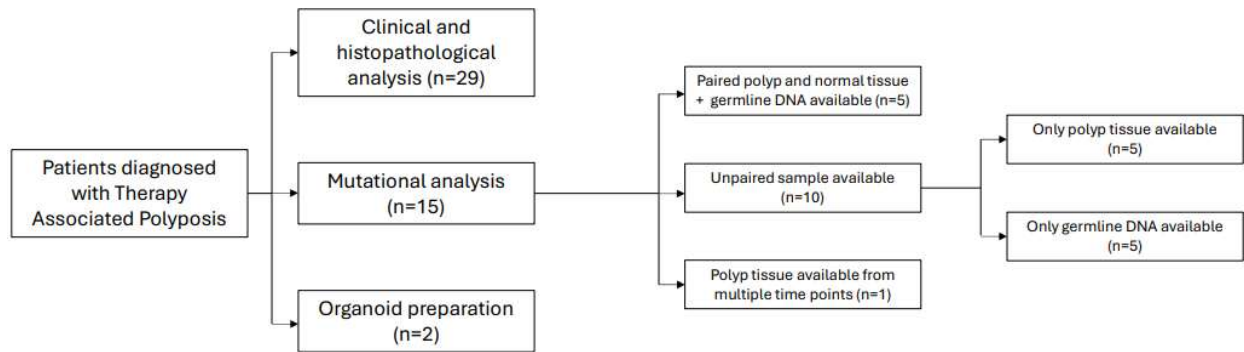
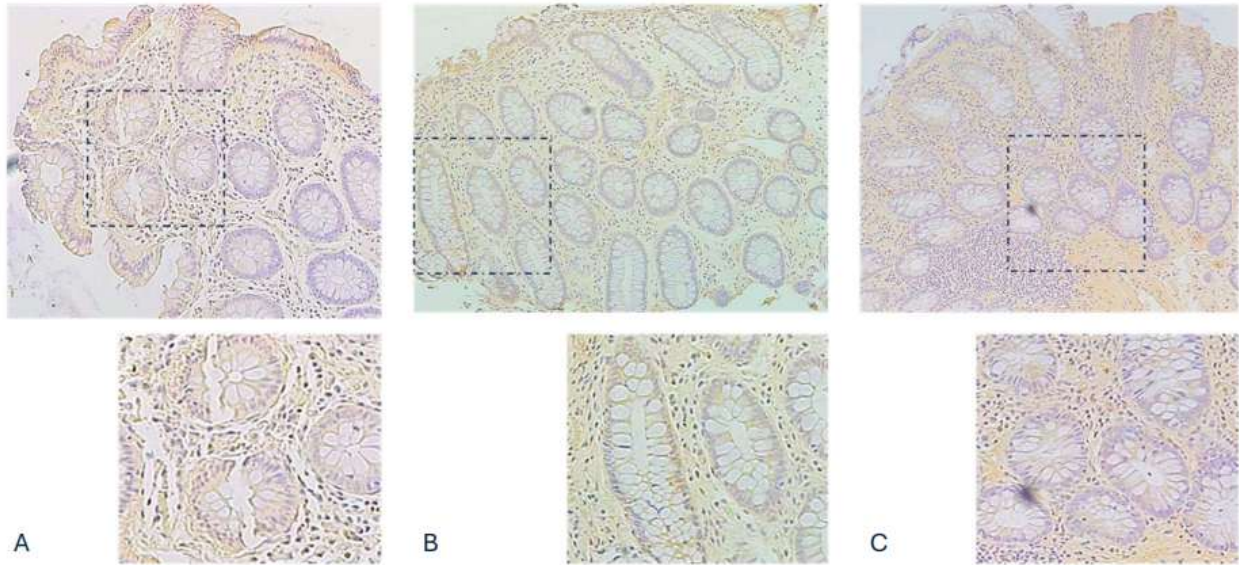


Supplemental Material

Supplemental Figures



Supplemental Figure 1. Consort diagram of the study population included in the study.



Supplemental Figure 2. Representative images of anti-GREM1 immunohistochemistry of colonic polyps from (A) a patient with therapy associated polyposis, B) a patient with known GREM1 duplication, C) a patient with familial adenomatous polyposis.

PID	Total number of colonoscopies	Total number of polyps reported in first colonoscopy recorded in EMR	Range of total number of polyps across all colonoscopies
1	8	Colonoscopy prior to 2015 (report not retrieved): Notes mention intramucosal carcinoma + multiple polyps 2015: 15 polyps	1 to 18
2	6	2012: 17 polyps	1 to 17
3	21	2009: >30-40 polyps	1 to >30-40
4	2	2021: 9 polyps	4 to 7
5	4	2004: 5 polyps	4 to 5
6	8	2015: 1 polyp	2 to 6
7	2	2021: 2 polyps	9 to 12
8	4	2018: 2 polyps	2 to 17
9	5	2017: 6 polyps	2 to >5
10	4	2017: 8 polyps	6 to 14
11	3	2019: 2 polyps	1 to 2
12	4	2021: 10 polyps	2 to 6
13	2	2009: 4 polyps	2 to 4
14	4	2020: 12+ polyps	12 to 18
15	2	2011: 7 polyps	4 to 6
16	2	2008: few polyps (number of polyps not available in EMR)	1 to multiple
17	8	2018: 2+ polyps	2 to multiple
18	5	2011: 50 polyps	1 to multiple
19	3	2016: few polyps (number of polyps not available in EMR)	2 to 6
20	3	2015: 12 polyps	5 to 33
21	2	2020: 4 polyps	5 to 6
22	2	2020: multiple polyps (number of polyps not available in EMR)	10 to multiple
23	2	2014: 9 polyps	2 to 9
24	5	2018: 13 polyps	2 to 13
25	6	2011: 18 polyps	1 to 18
26	5	2012: 4 polyps	1 to 20
27	6	2004: multiple polyps (number of polyps not available in EMR)	1 to multiple
28	5	2014: 2 polyps	1 to 3
29	3	2007: 13 polyps	3 to 13

Supplemental Table 1. Table summarizing total number of colonoscopies, polyp burden at first colonoscopy and range of number of polyps across surveillance for 29 patients with TAP.

PID	Location of CRC	Molecular subtype of CRC	Polyps showed cancer
1	Records unavailable	Unknown for Intramucosal carcinoma	Yes
6	Splenic flexure of the colon	Adenocarcinoma: absent expression of MSH6; MSI-H by PCR	Information not available
12	Splenic flexure	Adenocarcinoma with mucinous features arising from an adenomatous polyp, MMR normal	Information not available
22	Ascending colon	Adenocarcinoma: BRAF mutant, MMR deficient with loss of MLH1 and PMS2 with hyper methylation of MLH1	Information not available
28	Transverse colon	Information on molecular subtype unavailable	No

Supplemental Figure 2. Table summarizing the site and molecular subtype of CRC diagnosed in 5 patients with TAP.

Supplemental Methods:

Sex as a biological variable:

Both sexes, male and female, were included in the study.

Statistics:

A p value of 0.05 was considered significant. Due to small sample size, we didn't conduct any correlation tests.

Study approval:

The Institutional Review Board at Dana Farber Cancer Institute, Harvard Medical School, Boston, MA approved the present study in humans. Informed written consent was obtained from all patients.

Methods:

Patients with Therapy Associated Polyposis enrolled at Dana Farber Cancer Institute, Brigham and Women's Hospital and Massachusetts General Hospital between 2014- 2023 were identified and included in the study. Polyposis was defined as the lifetime incidence of 10 or more gastrointestinal polyps, and individuals diagnosed with CYAC at age 30 or between ages 31–45, provided their first gastrointestinal polyp was identified ten years after initial CYAC treatment, were included in the study. Patients with personal or family diagnoses of pathogenic germline variants in genes linked to inherited colorectal cancer susceptibility were excluded.

Baseline characteristics of the population including age, sex, race, ethnicity, mean age of primary cancer, type of primary cancer, treatment received for primary cancer, including site of

radiotherapy and type of chemotherapy received, mean age of development of first polyp, and detailed histology of polyps observed, were annotated retrospectively from electronic health records. Additionally, history of secondary cancers, other systemic side effects of exposure to radiotherapy and chemotherapy and history of cancers in first degree relatives and second-degree relatives were recorded. We investigated the overexpression of GREM1 protein using Immunohistochemistry with Anti GREM1 antibodies on formalin fixed paraffin embedded (FFPE) tissue slides obtained for 2 (28.6%) out 10 patients (34.5%) with family history of colonic polyposis in first degree relatives among the 29 patients with TAP. No cell lines/animal samples were used in the study.

To identify genomic alterations associated with TAP, we conducted Whole Genome Sequencing on polyp samples for 15 patients (51.7%) with TAP. Samples for DNA isolation were collected retrospectively from archived frozen blocks containing polyp tissue for 9 patients, including 2 patients with polyp tissue from more than 1 site. Fresh polyp samples along with normal tissue were collected for 1 patient during colonoscopy. DNA isolation from FFPE samples was performed using the QIAamp DNA FFPE tissue kit while QIAamp DNA Micro Kit was used to extract DNA from the fresh polyp sample.

Mutect2 was used to detect point mutations and indels in every sample (GATK 4.5.0.0), after aligning to the hg38 (GCA_000001405.15_GRCh38) reference genome. Only samples with a median coverage of at least 10 were evaluated. For samples without a paired normal sample (both WGS and WES cases), only variants with an allele frequency of less than 25% were regarded as possible somatic variants. Furthermore, these samples were further filtered for variants which had more than 5 tumor ALT read count and this list was refined more by retaining those mutations that are not present in gnomAD (V2.1) or KAVIAR (version 160204-Public) with an allele frequency greater than 0.01. Additionally, clustered mutations were also removed. To filter out the remaining

variants caused by technical errors, the intersection of mutations between all the samples was also not included in the final call set.

Single base substitution (Cosmic V3 SBS, filtered for the colorectal cancer signatures, based on: <https://www.nature.com/articles/s41586-020-1943-3>) signatures were extracted using the deconstructSigs and MutationalPatterns R packages (version 1.9.0, and 3.16.0 respectively), the short indel and doublet base substitution (ID and DBS) signatures were extracted using the ICAMS R package (version 2.3.12). Treatment-associated signatures have been also extracted alongside the colorectal cancer signatures, and included: SBS11, SBS25, SBS18, SBS36, SBS87, SBS31) These identify the linear combination of pre-defined signatures that best reconstructs a tumor's mutational profile.

In case of whole-genome samples copy number, changes were evaluated using depth- and segmentation-based callers and then merged into a single high-confidence catalogue. CNVnator (v0.4.1; bin size = 100 bp) generated read-depth segments with associated quality metrics, including the fraction of zero-mapping reads (q0) and CNVkit (v0.9.10) output $\log_2$ copy-ratio segments for each sample. Both call sets were imported into R 4.3 (GenomicRanges, data.table) and low-confidence regions and segments shorter than 1 kb were excluded. Because CNVkit exhibits finer break-point resolution, its segments were first symmetrically padded by ± 20 bp to allow for caller-specific rounding differences, then intersecting regions of the same CNV type (gain or loss) were identified. To identify and exclude GREM1 duplications CNVnator and delly was used and the reads were manually inspected using IGV. For visualization the normalized $\log_2$ copy-ratio profiles coverage using 3kb bins and default CNVkit parameters were plotted.

Tissue specimen from an unidentifiable patient with TAP was collected post endoscopy under approval (14-408) by the Internal Review Board of the Dana Farber Cancer Institute, Boston,

Massachusetts, USA. For generation of organoids from the patient-derived and mouse-derived tissues, the tissues were treated with EDTA and then resuspended in 30–50 μ l of Matrigel (BD Bioscience) and plated in 24-well plates. WNT/R-spondin/Noggin (WRN) containing DMEM/F12 with HEPES (Sigma-Aldrich) containing 20% FBS, 1% penicillin/streptomycin and 50 ng/ml recombinant mouse EGF (Life Technologies) was used for culturing colon organoids. For the first 2–3 days after seeding, the media was also supplemented with 10 mM ROCK inhibitor Y-27632 (Sigma Aldrich) and 10 mM SB431542 (Sigma Aldrich), an inhibitor for the transforming growth factor (TGF)- β type I receptor to avoid anoikis. For passage, colon organoids were dispersed by trypsin-EDTA and transferred to fresh Matrigel. Passage was performed every 3–4 days with a 1:3–1:5 split ratio. For human colon organoid culture, the previous media was supplemented with antibiotics 100 μ g/ml Primocin (Invivogen), 100 μ g/ml Normocin (Invivogen); serum-free supplements 1 $\times$ B27 (Thermo Fisher (Gibco)), 1X N2 (Thermo Fisher (Gibco)); chemical supplements 10 mM Nicotinamide (Sigma), 500mM N-acetylcysteine (Sigma), hormone 50 mM [Leu15]-Gastrin (Sigma), growth factor 100 μ g/ml FGF10 (recombinant human) (Thermo Fisher) and 500nM A-83-01 (Sigma), which is an inhibitor of the TGF- β Receptors ALK4, 5, and 7.

Data availability:

Data on clinical characteristics of patients with TAP can be accessed upon request from the corresponding author. Genomic data is available in public repositories dbGAP study ID 67888.