To reviewers

Reviewer 1 (Reviewer's code: 02679280)

#1. Major points Material and methods p11. Please, provide at least the following information regarding the gene-panel sequencing: what is the size of the target, how much of the captured region is covered by at least 25 reads that authors used as a threshold for filtering, what was the mean coverage for the samples (including minimum and maximum). Please, provide more details for the variant filtering: were all SNVs included or were these further filtered based on their effect (synonymous/non-synonymous, splicing etc.)? Were any predictions made for the effect of the amino acid alterations (PolyPhen, SIFT, MutationTaster etc.)?

We agree with comments from the reviewer. The requested information has been added to the Material and Methods part in the manuscript (page 7,8).

In the case study, total 1,688,650 target bases were sequenced per case. With the threshold (at least 25 reads) for filtering, 99% of the sequenced region was covered in each case. Total numbers of mapped reads in the studied samples were between 20,907,271 and 37,358,579.Average base coverage depth ranged from 1,355 to 2,397. Other details are listed in the table below: (A-F: represents sample ID)

Original Sample ID	Tumor specimen labeling	Number of mapped reads	Percent reads on target	Uniformity of base coverage	Target base coverage at 25x	Target base coverage at 100x	Target base coverage at 500x	Average base coverage depth	Minimum base coverage depth	Maxmum base coverage depth
1202037A/B	A	25,656,765	99%	93%	99%	97%	88%	1,617	1	21,433
1513541Bd	B	37,358,579	99%	94%	99%	98%	93%	2,397	1	31,969
1507600	C	23,998,935	99%	91%	99%	97%	84%	1,512	1	24,332
1513541A/B	D	20,907,271	99%	94%	99%	96%	88%	1,355	1	18,982
1513373	E	26,624,353	99%	92%	99%	97%	87%	1,681	1	26,604
1513541cb	F	34,621,775	99%	93%	99%	97%	92%	2,221	1	27,939

Yes, in this commercially available genetic testing panel, we routinely use

Grantham, SIFT, and Polyphen to predict the effect of amino acid alterations. Our raw data analysis is listed below:

Gene	cDNA change	protein change	Grantham	SIFT	PolyPhen
FANCA	c.4075G>T	p.D1359Y	damaging	tolerated	possibly_damaging
BAI3	c.932C>G	p.S311W	damaging	tolerated	probably_damaging
RPS6KA2	c.1784C>T	p.T595I	damaging	deleterious	probably_damaging
TET1	c.2617G>A	p.V873I	benign	tolerated	benign
GNAS	c.2530C>T	p.R844C	damaging	deleterious	probably_damaging
SMAD4	c.1081C>T	p.R361C	damaging	deleterious	probably_damaging
CTNNB1	c.134C>T	p.S45F	damaging	deleterious	probably_damaging
PLEKHG5	c.607C>T	p.R203C	damaging	deleterious	probably_damaging
BAI3	c.1850T>C	p.L617P	damaging	deleterious	probably_damaging
TP53	c.489dupC	p.K164fs	NA	NA	NA
KIT	c.1502_1503insTGCCTA	p.S501_A502insAY	NA	NA	NA
PSIP1	c.1541G>A	p.S514N	benign	tolerated	benign
BCL9	c.3572C>G	p.A1191G	damaging	tolerated	possibly_damaging
BCL9	c.3583C>G	p.P1195A	benign	tolerated	benign
NKX2-1	c.627C>A	p.F209L	benign	tolerated	benign

#2. Tables Table 1. Please provide the overall read depth for variants having less than 10% frequency in the samples. As C to T conversion is common artefact of FFPE samples, how authors excluded the possibility that these (particularly the ones observed only in one sample) do not represent artefacts?

The overall allele frequency and read depth for each variant is listed in the table below: (A-F: represents sample ID)

a. Variant frequency

Gene	cDNA change	protein change	A	B	C	D	E	F
FANCA	c.4075G>T	p.D1359Y	48.4%	54.0%	47.5%	52.0%	37.8%	45.7%
BAI3	c.932C>G	p.S311W	50.3%	46.1%	40.9%	50.4%	51.3%	49.3%
RPS6KA2	c.1784C>T	p.T595I	48.0%	49.4%	47.7%	45.7%	51.5%	48.4%
TET1	c.2617G>A	p.V873I		10.1%	7.4%			
GNAS	c.2530C>T	p.R844C		6.3%				
SMAD4	c.1081C>T	p.R361C			6.4%			
CTNNB1	c.134C>T	p.S45F			23.6%	4.2%		
PLEKHG5	c.607C>T	p.R203C			40.7%	15.3%	35.4%	
BAI3	c.1850T>C	p.L617P					37.7%	
TP53	c.489dupC	p.K164fs					49.9%	
KIT	c.1502_1503insTGCCTA	p.S501_A502insAY						37.3%
PSIP1	c.1541G>A	p.S514N						14.0%
BCL9	c.3572C>G	p.A1191G						13.1%
BCL9	c.3583C>G	p.P1195A						13.0%
NKX2-1	c.627C>A	p.F209L						7.5%

b. Depth of reads

Gene	cDNA change	protein change	A	B	C	D	E	F
FANCA	c.4075G>T	p.D1359Y	1346	2690	1076	1167	1620	1999
BAI3	c.932C>G	p.S311W	993	3094	955	1435	1617	2287
RPS6KA2	c.1784C>T	p.T595I	1372	2704	1167	1775	1361	2709
TET1	c.2617G>A	p.V873I	2605	2153	2701	1610	2714	3015
GNAS	c.2530C>T	p.R844C	845	1231	896	803	1042	1353
SMAD4	c.1081C>T	p.R361C	2003	2911	1506	1123	836	2287
CTNNB1	c.134C>T	p.S45F	1664	2356	1811	1344	1597	2548
PLEKHG5	c.607C>T	p.R203C	498	879	305	326	520	431
BAI3	c.1850T>C	p.L617P	1481	2235	1045	1365	1469	2027
TP53	c.489dupC	p.K164fs	844	1415	518	883	559	1577
KIT	c.1502_1503insTGCCTA	p.S501_A502insAY	2284	2822	1855	1468	2670	2755
PSIP1	c.1541G>A	p.S514N	3920	3408	3418	1698	2958	3634
BCL9	c.3572C>G	p.A1191G	1215	2587	1171	1656	1751	2805
BCL9	c.3583C>G	p.P1195A	1220	2597	1174	1666	1762	2808
NKX2-1	c.627C>A	p.F209L	384	1633	300	796	450	1243

We agree with reviewer's comment that C>T/G>A changes is common artefacts of FFPE samples especially in the 1-10% allele frequency range (Wong et al. BMC Med Genomics. 2014 May 13;7:23).

There was a significant association between low coverage, reduction of available templates caused by fragmentation and high amounts of C>T/G>A changes in the 1-10% allele frequency range.

We understand that even with careful methodology optimization, more attention should be paid to the interpretation of the results. The purpose of the genetic testing in this project is to identify the major genetic events in "adenoma-carcinoma sequence". Therefore, we only discuss the biological significance of recurrent mutation.

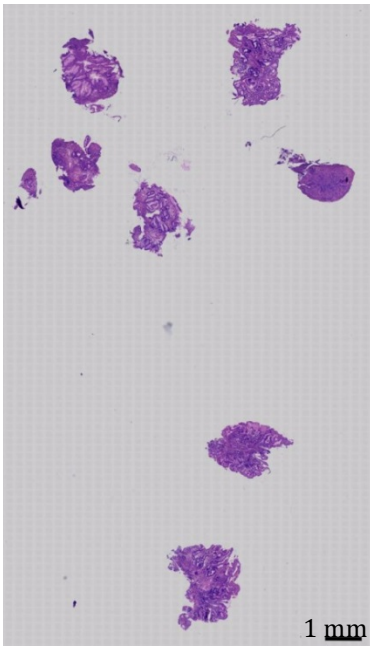
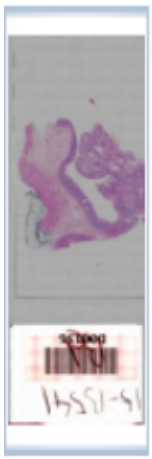
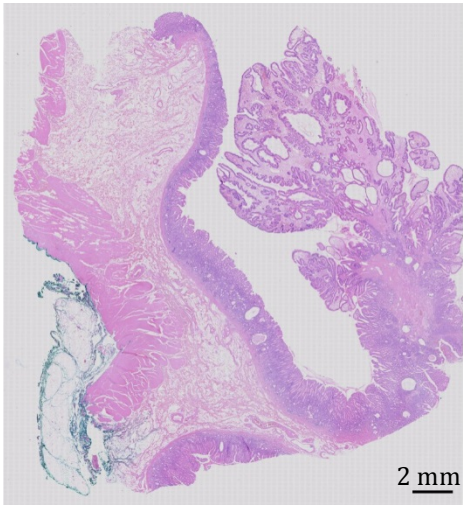
#3. Minor points Introduction p8. The phrase "genetic expression profiles" is misleading, the expression referring traditionally to RNA as a starting material. "Genetic profile" or "mutational profile" would be more suitable given the analysis done.

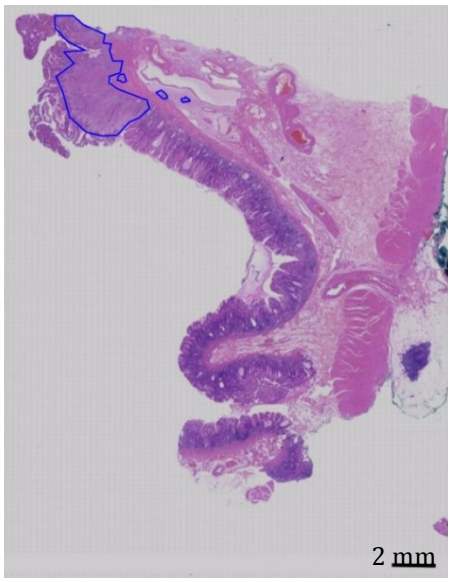
We agree with reviewer's comment and have revised the wording error (p 8).

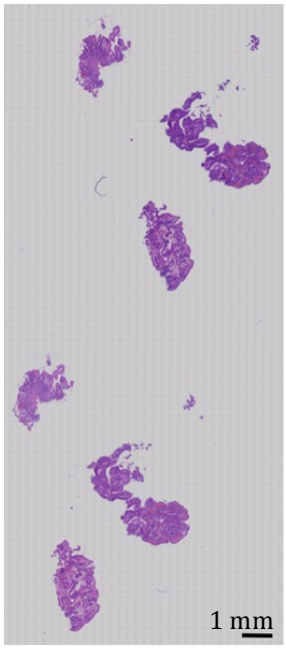
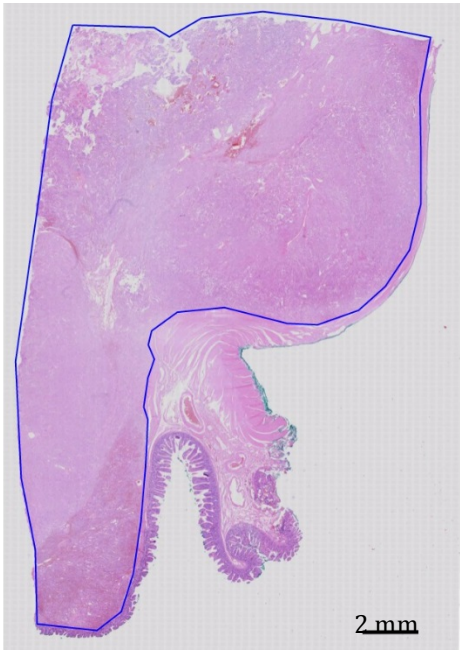
#4. Results p.11 Was samples from other organ sites (e.g. blood, buccal swab) available to further confirm the germline origin of the three observed mutations?

In this study we sequenced 6 samples, of which, 3 samples were benign. Despite the studied samples contained a variety of tumor component (0 to 75%) , frequencies of 3 variants (*FANCA* c.4075G>T, *BAI3* c.932C>G, *RPS6KA2* c.1784C>T) were detected at ~ 50%, indicating these mutation were from germline rather than somatic.

Histology of each sample in H&E staining and the portion (%) of tumor component are listed below:

Sample ID	Tumor specimen labeling	HE stain	
D00075 1202037A/B	A	 0%	
D00136 1513541Bd	B	 0%	

D00079 1507600	C	 60% (tubular adenoma: 60%; invasive cancer: 0%)	
D00077 1513541A/B	D	 5%	

D00080 1513373	E	 30%	
D00137 1513541cb	F	 75%	

#5. Discussion p.14 It would be good to mention that this FANCA D1359Y mutation has previously been described in the context of FA. p.15 I do not quite follow, how this sentence "In the past, to identify a subtype of FA needed clinical awareness and was often hampered by labor intensive conventional molecular diagnosis tools such as conventional mutation analysis, gene transfer studies or western blotting" relates to the current case report. I would recommend omitting this.

Thanks for the comments. We have deleted the part " In the past …western blotting" and corresponding reference in the revised manuscript.

#6. Tables Table 2. How does this relate to the manuscript? It is not referred in the text.

Thank you for the comments. Table 2 is removed in the revised manuscript.

Reviewer 2 (Reviewer's code: 02535507)

An excellent genetic study about gastric cancer. Some points are unclear for me:

a. Why gastric polyps were not removed, but only subjected to biopsy at the first episode of bleeding?

Thank you for the comments. The patient had multiple gastric polyps at diagnosis and endoscopic excisional biopsy was done for the largest one. Because the pathology reported benign, he was followed by annual upper endoscopy until malignant transformation was noted. Surgery was arranged when cancer diagnosis was confirmed.

b. There is a link between gastric hyperplastic polyp evolution to carcinoma and duodenal GIST development?

Thank you for the comments. Our case study showed malignancy develops in hyperplastic polyps through a dysplasia/carcinoma sequence. Given the relative low incidence (1~20%) for hyperplastic polyps to harbor foci of dysplasia, we proposed germline mutation *FANCA* D1359Y may play an important role in cancer susceptibility. In addition to genetic susceptibility, environmental factors play critical roles in the development of gastric malignancy, such as *H. pylori*, chronic use of proton pump inhibitor and smoking Our case study also showed the genetic event acquired for GIST were different from those in gastric adenocarcinoma.

c. It seems that surgery induced a sustained remission of anemia, bleeding and tumor development. In this case I cannot understand why the source of bleeding was not removed at the first episode.

Thank you for the comments. Please refer to our response for a.

Reviewer 3 (Reviewer's code: 03017407)

Theoretically, this is an interesting case. However, all along the paper, there is a tremendous confusion between germline and somatic variants (mutations). Germline mutations can predispose to cancer and this is the case of monoallelic mutations of FA genes. If mutations are low/medium penetrant, family history of cancer can be absent and, consequently, the mutation carrier affected with cancer will appear as a sporadic case. Somatic mutations are only present in pre-neoplastic and neoplastic lesions and are associated with cancer progression. Which is the evidence that the identified mutation is germline, i.e. present in normal tissues? The authors extracted DNA "from benign gastric polyp, gastric adenocarcinoma and jejunal GIST tumor". No constitutive DNA from normal tissue was analyzed? If constitutive DNA was not extracted from blood but from FFPE sections, how areas with normal cells were selected?

Thank you for the comments. Since similar questions have been raised by reviewer 1, please refer to p4 for our response.

The sentence "Massively parallel sequencing for a panel of 409 cancer-related genes in these tumors identified 3 germline mutations

(BAI3 p.S311W, FANCA pD1359Y, RPS6KA2 p.T595I) and 12 somatic mutations in 3 benign and 3 malignant tumors" is quite confusing. Assuming the presence of a germline mutation in FA gene, is a second hit in the same gene present in tumor cells? Again, about confusion between germline and somatic variants: the sentence "It was reported that patients with a monoallelic FA gene mutation are also prone to the development of colorectal cancer when an additional second hit, such as MHL gene mutation is present" is totally inappropriate. The mentioned paper (Xie et al. 2010) evaluates the functional effect of a germline mutation in MLH1 (not in FANCJ), which is a MMR gene associated with the Lynch syndrome. The MLH1 mutation was shown to impair the binding between MLH1 and FANCJ proteins, thus impairing the MMR signaling. These are just example: as I said, there is confusion all along the paper between germline (predisposition) and somatic (carcinogenic process) events. Accordingly, the paper has to be completely re-written.

Thank you for the comments.
To avoid confusion and controversy, this statement is deleted from the revised manuscript.

According to Xie's article (Cancer Prev Res (Phila). 2010 Nov; 3(11): 1409–1416), *FANCJ*, a DNA helicase, contributes to mammalian mismatch repair (MMR) *in vitro*, however, loss of *FANCJ* may not overtly limit MMR. Therefore, the authors considered that loss of MLH1 binding

to *FANCJ* could alter MMR signaling. In their research, *FANCJ*-deficient cells exhibit delayed MMR signaling, which provides time for the methylguanine methyltransferase (MGMT) enzyme to reverse DNA methylation. FANCJ deficiency alters the competition between two pathways: MGMT- prosurvival versus MMR-prodeath. From this standpoint, a second hit in the background of FANCJ deficiency is critical for tumorigenesis.

Reviewer 4 (Reviewer's code: 01214406)

this is a good case report with a followup of long duration showing sequential mutational changes from benign polyp to adenocarcinoma in stomach. the study clearly shows that like adenocarcinoma colon gastric adenocarcinoma may also follow the adenoma carcinoma sequence authors have done detailed mutational analysis of followup biopsies to prove above hypotheses but a single case is not enough to make any final conclusion. Results are clear but should be made more concise. discussion is too long,the authors should mainly concentrate on significant mutational changes that support sequential adenoma carcinoma sequence Language and formatting needs improvement

We agreed with comments from the reviewer. Discussion part has been revised to make it more concise.