

OPEN

Systematic mutation analysis in rare colorectal cancer presenting ovarian metastases

Sungjin Park^{1,2,10}, Hee Kyung Ahn^{3,10}, Dae Ho Lee^{1,4,5}, YunJae Jung⁶, Joo-Won Jeong⁷, Seungyoon Nam^{1,2,3,8*} & Won-Suk Lee^{1,9*}

Although colorectal cancer is one of the most lethal cancer types in the world, its metastasis to the ovary is rare, compared to metastasis to other organs. Consequently, the genomic basis for colon-to-ovary metastasis remains unstudied, due to limited available patients, and thus there have been no attempts to construct individual-specific networks. Due to its rarity, the small sample size makes common mutations difficult to find. To overcome this problem, we herein attempted to apply a biological connectivity map called a sample-specific network (SSN), to reveal common biological functions in three samples. Our three samples were compared to a clinical dataset contained in The Cancer Genome Atlas (TCGA) Colorectal Adenocarcinoma (COAD), showing different mutational spectra, compared to matched samples based on age, gender, microsatellite instability (MSI) status, and tumor, node, metastasis (TNM) stage. The SSNs for the three samples revealed significant correlations of the mutation statuses of several apoptosis genes, in contrast to the TCGA-matched samples. Further analysis of a targeted-gene panel sequencing dataset for colon-to-ovary metastasis of primary tumor samples also confirmed significant correlations of the mutational statuses among apoptosis genes. In summary, using SSN, we successfully identified a common function (apoptosis) among our three patients having colon-to-ovary metastasis, despite no common mutations in the three patients. Such computational analyses could facilitate productive study of rare cancers and other diseases.

Colorectal cancer (CRC) is the third-most common cancer in the world, with an estimated 43.7 new cases, per 100,000 men and women, per year, and the fourth-leading lethal cancer, with 881,000 deaths in 2018¹. During the past decade, clinicians have witnessed remarkable progress in the treatment of metastatic colon cancer, especially upon understanding its molecular basis, leading to effective combinations of chemotherapy regimens of oxaliplatin, irinotecan, cetuximab, and bevacizumab². These novel agents have yielded both higher tumor response rates and prolonged overall survival.

Despite this substantial improvement in overall survival, however, CRC patients with rare colon-to-ovary metastases still have poorer prognoses and chemotherapy responses, compared to other metastatic sites^{3–6}. It was also reported that, remarkably, the non-ovarian metastases tumor response rate to 5-FU-based chemotherapy was 40%, but only 5% for ovarian metastases⁷. Although Lee *et al.*⁶ reported that ovarian metastatectomy improved survival⁶, better understanding of the underlying mechanism(s) for resistance to nonsurgical therapy is urgently needed.

The frequency of ovarian metastasis from a primary CRC is low, at approximately 3.4% of women diagnosed with a malignant colorectal cancer⁸, compared to higher occurring metastases to the liver and the lung⁹. Due to this rarity, limited sample sizes (and thus low statistical power) represent a major hurdle for genomic analyses of

¹College of Medicine, Gachon University, Incheon, Korea. ²Gachon Institute of Genome Medicine and Science, Gachon University Gil Medical Center, Incheon, Korea. ³Department of Medical Oncology, Gachon University Gil Medical Center, Incheon, Korea. ⁴Gachon Advanced Institute of Health Sciences & Technology, Gachon University, Incheon, Korea. ⁵Department of Internal Medicine, Gachon University Gil Medical Center, Incheon, Korea. ⁶Department of Microbiology, College of Medicine, Gachon University, Incheon, Korea. ⁷Department of Anatomy and Neurobiology, College of Medicine, Kyung Hee University, Seoul, Korea. ⁸Department of Life Sciences, Gachon University, Seongnam, Gyeonggi-do, Korea. ⁹Department of Surgery, Gachon University Gil Medical Center, Incheon, Korea. ¹⁰These authors contributed equally: Sungjin Park and Hee Kyung Ahn. *email: nams@gachon.ac.kr; lws@gilhospital.com

Patient #	Gender	Age	Survival status (survival time/ months)	TNM stage	KRAS	EGFR status	Microsatellite instability	Primary tumor size(cm)	Primary tumor treatment	Ovary metastasis size(cm)	Metastasis treatment
5	F	54	Alive (11)	IV	WILD	+	Stable	5	Anterior resection	4	FOLFIRI/FOLFOX/Cetuximab/Radiation
8	F	52	Alive (48)	IV	WILD	+	Stable	3.5	Low anterior resection	10	FOLFIRI/FOLFOX/Cetuximab/Bevacizumab/
9	F	51	Alive (10)	IV	WILD	+	Stable	3	Anterior resection	1	FOLFIRI/FOLFOX/Bevacizumab/lapatinib

Table 1. Clinico-pathological data of the three patients.

CRC-ovarian metastasis. Therefore, it is important to apply a proper strategy to analyze mutational phenomena, using a systematic view, not only by comparing mutational occurrences.

In light of these challenges we performed a systematic study to improve the genetic characterization of ovarian metastases. We employed a novel bioinformatics approach to investigate genetic characteristics of CRC-ovarian metastatic tumors, by applying a sample-specific network (SSN) method to analyze unbiased, whole-exome sequencing (WES). A SSN allows creation of an individual-specific molecular network, via genomic profiling of a single sample¹⁰. We then compared the genetic landscape between primary colon tumors and their colon-to-ovary metastases, to attempt to reveal potential biomarkers for this rare condition.

Results

Demographic and clinical manifestations. Simultaneous ovarian metastases from three Korean female CRC patients, aged 52.33 ± 1.52 years old (range of 51 to 54 years old), were collected. All three patients possessed wild type *KRAS* and microsatellite stability (MSS), and one patient showed *HER2* amplification, as determined by pathologic examination. These and other clinical details are described in Table 1 and the Materials and Methods.

Pathway mutations in CRC tumors. We first examined the mutational profiles of our patients, to subsequently built SSNs, using a method to determine how mutational co-occurrences of each patient's ovarian CRC metastasis differed in samples matched for age, gender, and tumor stage¹⁰. Such data, of five samples selected (henceforth, "TCGA-matched samples"), was retrieved from The Cancer Genome Atlas (TCGA), allowing us to build tumor mutational co-occurrence networks. We first built such networks from mutation data of the TCGA-matched samples, using the Pearson correlation coefficient as an association index. This network, henceforth referred to as a "reference network", contained molecular-level mutational co-occurrences among genes¹⁰. Then, we added each patient's primary tumor mutational profile to the mutation dataset of the TCGA-matched samples, generating a "perturbed network", i.e., a network derived from a "reference network," by adding a sample of interest¹⁰. Then, we used a statistical test to identify differential "edges" between the reference network and the perturbed network¹⁰. Here, a "node" represents a gene, while an edge represents a relationship, between two nodes, in the perturbed network. The value of an edge represents a quantitative strength of how two nodes are related. The entire framework of our method is illustrated in Fig. 1.

We identified mutational profiles of members of the WNT beta-catenin signaling pathway, including MAPK, PI3K, TGF- β , and p53 pathways, when comparing our three patients to the TCGA-matched samples (Fig. 2). In Fig. 2, along with mutational occurrences of our three patients, we juxtaposed frequencies of genetic alterations of a non-hypermuted group (defined as a mutation rate of <8.24 per 10^6 base pairs)¹¹ from the TCGA Colorectal Adenocarcinoma (COAD) report in 2012¹¹. While only one of our three patient samples had a somatic *APC* mutation, the TCGA-matched samples all showed mutated *APC* (81% of the non-hypermuted group had *APC* alterations) in the WNT/beta-catenin signaling pathway (Fig. 2). Although all three of our patients had *TP53* mutations, only 60% of the TCGA-matched samples had these (59% of non-hypermuted group showed *TP53* alteration) (Fig. 2). For *PIK3CA* (a member of the PI3K signaling pathway), one of the three, and 40% of the TCGA-matched samples, had mutations, although only 15% were found altered in the non-hypermuted group (Fig. 2). The DNA double-strand break repair enzyme gene, *ATM*, was not mutated among our patients, although 20% of the TCGA-matched samples did possess *ATM* mutations, as did 7% of the non-hypermuted samples¹¹ (Fig. 2).

Differential correlation changes, of mutation statuses, among genes, in our three patients. The apoptosis signaling pathway was a common functional context among our three patient ovarian primary tumors, when we constructed three perturbed networks, and then merged these into one network (Fig. 3a). These assessments revealed that apoptosis-related gene correlation coefficients of mutational statuses (between two neighboring genes), were changed significantly in three of our patients, #5, #8, and #9. In all our patients, the neurofilament peptide gene *NEFH* (not mutated in the network of the TCGA-matched samples) correlated with genes belonging to the apoptosis process (Fig. 3b). For patient #5, *IL13RA2* and *ROCK1* (a Rho kinase) negatively correlated with *NEFH*, in terms of their mutation statuses. *ROCK1* was also a crosstalk gene between the apoptosis and mitotic spindle pathways. For patient #8, *GSN* (gelsolin), which facilitates crosstalk between the apoptosis and coagulation pathways, negatively correlated with *NEFH* (Fig. 3b). For patient #9, *ERBB2*, *BCL2L10*, *PDCD4*, and *PACRG*, all showed negative correlations with *NEFH* (Fig. 3b). *ERBB2*, an oncogene, and *BCL2L10*, belonging to the apoptosis pathway, positively correlated with each other and *PDCD4*, of the estrogen response late pathway (Fig. 3b). Since patient #9 had *HER2* amplification, its crosstalk genes (*PDCD4*, *BCL2L10*, *NEFH*) should be later examined for correlation with *ERBB2*, and possible clinical associations. Changes in gene correlation are described in Table 2.

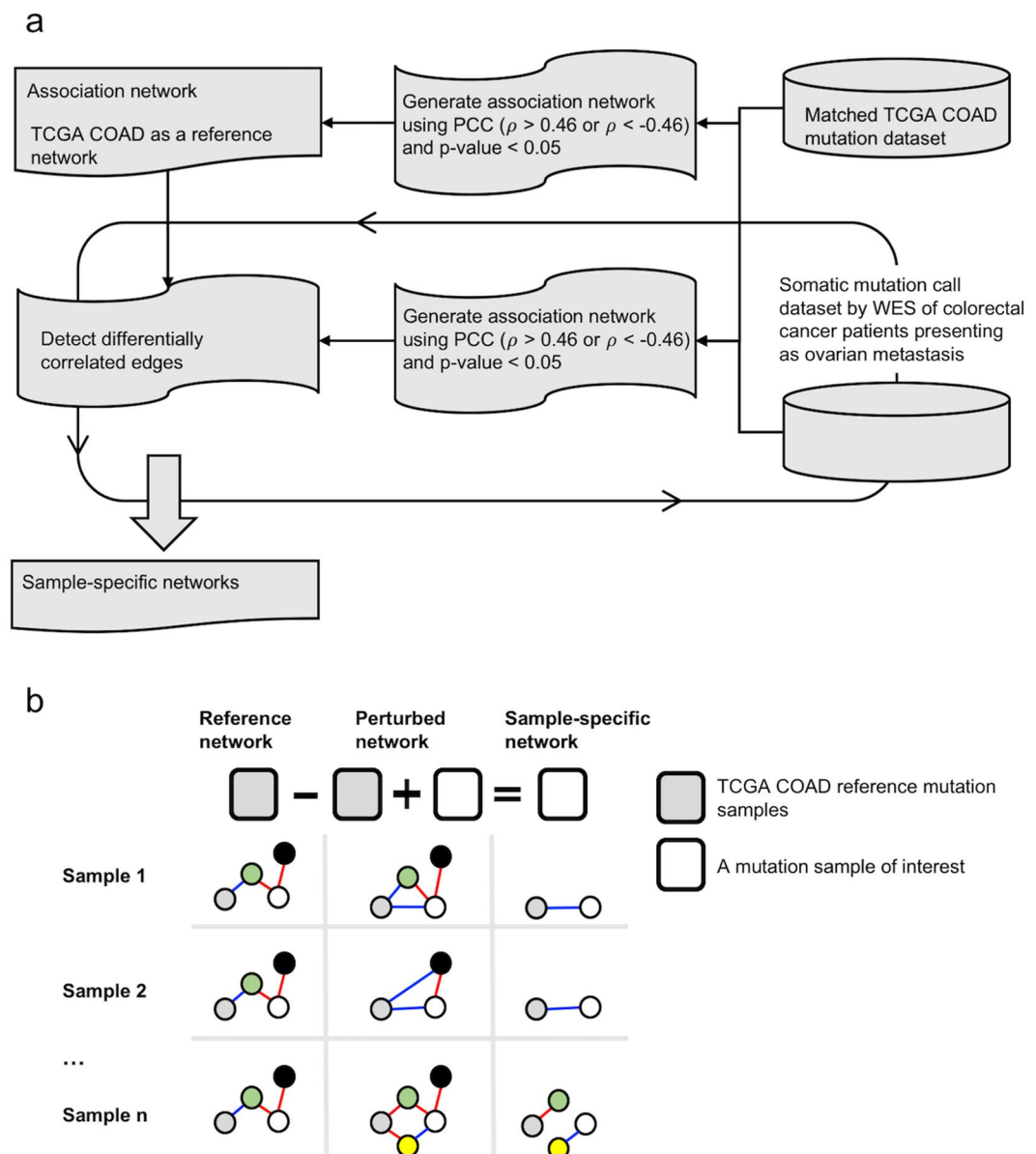


Figure 1. Framework of building sample-specific networks. **(a)** We built an association network (a “reference network”), using mutation information from TCGA-matched samples. Each of our samples was compared to the reference network, to build another association network of the sample, i.e., a “perturbed network”. We then used a Pearson correlation-based statistical test to identify differentially correlated edges between the “reference network” and the “perturbed network”. We applied this process to all patients’ somatic mutation sample datasets. **(b)** A conceptual illustration of the building of the sample-specific network. ρ is a correlation coefficient.

We next assessed whether the functional contexts (apoptosis, mitotic spindle, and spermatogenesis), in our dataset, were likewise present in an independent dataset. To that end, we analyzed a targeted panel sequencing dataset of 26 colorectal ovarian metastases¹², identifying significant correlation changes in 23 of the 26 samples, compared to the TCGA-matched samples. Those three contexts were revealed again (HALLMARK_APOPTOSIS in ten patients, HALLMARK_MITOTIC_SPINDLE in four patients, and HALLMARK_SPERMATOGENESIS in one patient) in the targeted panel sequencing dataset (Supplementary Fig. S1a). Moreover, there were significant correlation changes of mutational statuses in apoptosis genes, in ten of the 26 samples (HALLMARK_APOPTOSIS in Supplementary Fig. S1a), with apoptosis nodes also correlating with *NEFH*, forming six different network topological patterns (Supplementary Fig. S1b). *NEFH* also negatively correlated with *ERBB2*, in five samples (patterns 1 and 2 in Supplementary Fig. S1b). In three samples, *NEFH* negatively correlated with the metastasis-related gene, *MMP2* (patterns 3 and 4 in Supplementary Fig. S1b), and in two samples, *NEFH* negatively correlated with *SMAD7* (patterns 5 and 6 in Supplementary Fig. S1b). As a result, we revealed that our dataset, and the independent dataset, shared common functional contexts, with *NEFH* implicated in both datasets.

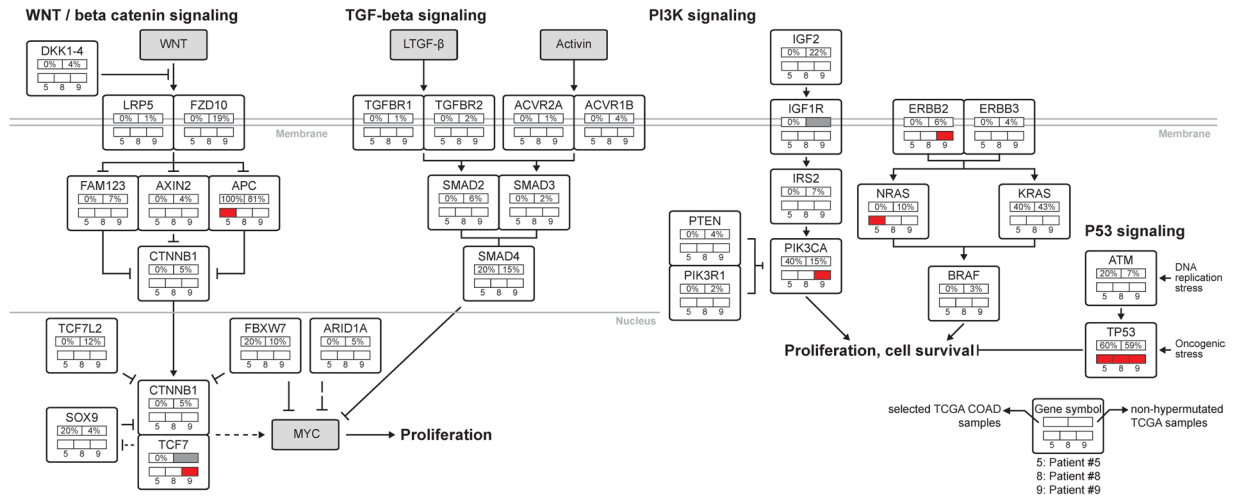


Figure 2. Diversity and frequency of genetic changes in our patients. *APC* was reported highly altered in the TCGA COAD report¹¹ and we sought similar mutation ratios in the TCGA-matched samples. However, in our patients, only patient #5 had an *APC* mutation. All our patients had mutated *TP53*, which was mutated in only 60% of the TCGA-matched samples. Each gene has two box rows. In the upper box row, the left cell indicates a percentage of mutation observed from the TCGA-matched samples, and the right cell represents a percentage of alterations reported in TCGA COAD. A gray-filled box means a missing value. In the lower box row, three cells represent mutation statuses of patients #5, #8, and #9, from left to right (red cells = mutation; white cells = wild type).

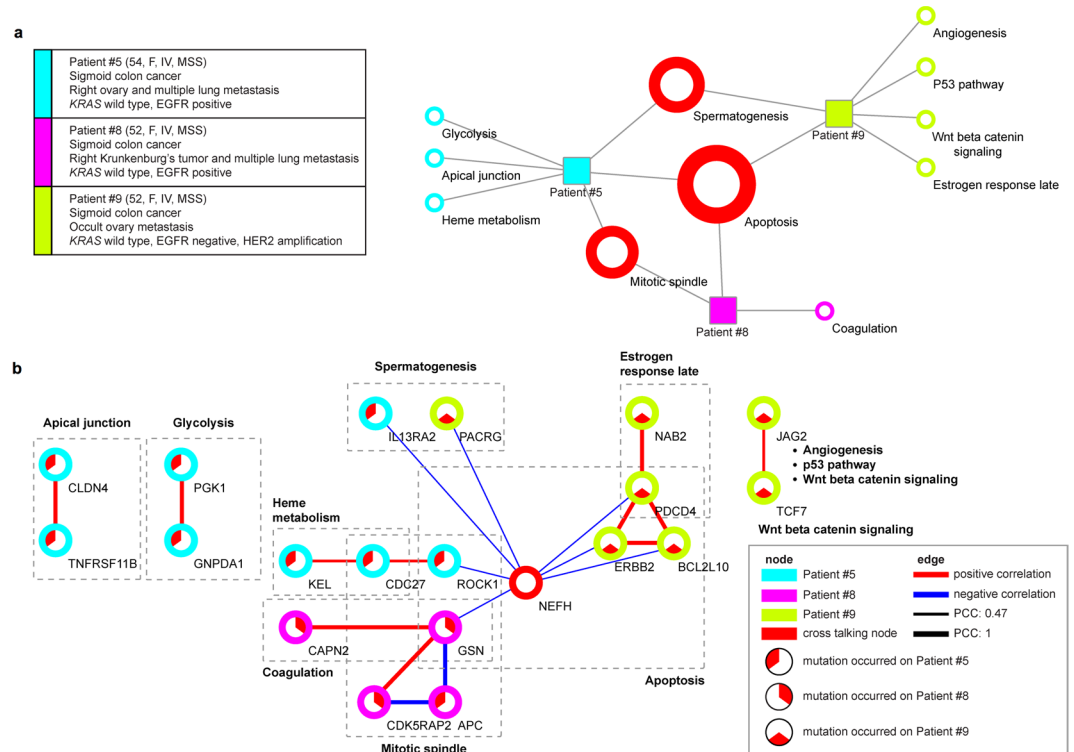


Figure 3. Inspection of sample-specific networks across our three colorectal cancer patients presenting ovarian metastases, compared to the TCGA-matched samples. (a) Network view showing the hallmark gene sets, as their correlation coefficients changed significantly among our colon-to-ovary cancer patients, compared to the TCGA-matched samples. (b) Gene-gene interaction diagram showing the mutational spectra of each patient. Colored circles indicate patients and crosstalk nodes. Colored edges indicate positive (red) and negative (blue) correlations, and their respective thicknesses indicate amounts of correlation coefficients. PCC: Pearson correlation coefficient.

Interaction (Genes A-B)	Hallmark gene sets	W_{AB}^a	ΔW_{AB}^b	p-value of interaction
JAG2-TCF7	WNT BETA CATENIN SIGNALING	0	0.65	0.001
PACRG-NEFH	SPERMATOGENESIS	0	-0.47	0.018
BCL2L10-PDCD4	APOPTOSIS	0	1.00	5.73E-07
BCL2L10-ERBB2	APOPTOSIS	0	1.00	5.73E-07
BCL2L10-NEFH	APOPTOSIS	0	-0.47	0.018
PDCD4-NAB2	ESTROGEN RESPONSE LATE	0	1.00	5.73E-07
PDCD4-ERBB2	APOPTOSIS	0	1.00	5.73E-07
PDCD4-NEFH	APOPTOSIS	0	-0.47	0.018
ERBB2-NEFH	APOPTOSIS	0	-0.47	0.018
CDK5RAP2-GSN	MITOTIC SPINDLE	0	1.00	5.73E-07
APC-GSN	MITOTIC SPINDLE	0	-1.00	5.73E-07
APC-CDK5RAP2	MITOTIC SPINDLE	0	-1.00	5.73E-07
GSN-CAPN2	COAGULATION	0	1.00	5.73E-07
GSN-NEFH	APOPTOSIS	0	-0.47	0.018
CLDN4-TNFRSF11B	APICAL JUNCTION	0	1.00	5.73E-07
PGK1-GNPDA1	GLYCOLYSIS	0	1.00	5.73E-07
IL13RA2-NEFH	SPERMATOGENESIS	0	-0.47	0.018
KEL-CDC27	HEME METABOLISM	0	0.65	0.001
CDC27-ROCK1	MITOTIC SPINDLE	0	0.65	0.001
ROCK1-NEFH	APOPTOSIS	0	-0.47	0.018

Table 2. Correlation coefficient changes in sample-specific networks. W_{AB} and ΔW_{AB} are described in detail in the Methods section. We also added the relevant hallmark gene sets for an interaction of two genes. The p-value was obtained from the Z-value that showed a ΔW_{AB} of an edge was significant statistically. ^aPearson correlation coefficient between two gene pair in a “reference network”. ^bDegree of Pearson correlation coefficient change between two gene pair in a “perturbed network”.

Comparison of mutations of primary colon and metastasized ovarian tumors. In our three patients, both primary and metastatic ovarian tumor tissues were compared, showing that all patients shared *TP53* mutations in both their primary and metastasized tumors (Supplementary Table S1). Consequently, we built two mutational co-occurrence networks for the primary CRC and metastasized ovarian tumors, to compare their topological configurations. We hence observed a network structural similarity between the two networks, based on single nucleotide variations (SNVs, Supplementary Table S1). Overall, significant genes, and their correlations, were preserved in the two mutational co-occurrence networks, although there were changes of neighboring partners or correlational statuses (Supplementary Fig. S2). For example, the gene pairs *APC-CDC27*, and *CDC27-ROCK1* were positively connected in the CRC network, while these became negatively connected in the metastatic network.

Discussion

In this study, we studied the mutational landscape of three patient samples of rare ovarian colorectal (CRC) metastases, as compared to their primary CRC tumors. The mutational co-occurrences of our three samples showed different mutational co-occurrences, both in age-/tumor stage-matched samples in the TCGA, and in a non-hypermutated group, a TCGA Colorectal ADenocarcinoma (COAD) report from 2012¹¹. All three patients had *TP53* mutations, which were present in only 60% of the TCGA-matched and non-hypermutated patient samples. For the CRC-causing gene *APC*, in the WNT/beta-catenin signaling pathway, all the TCGA-matched, and 81% of non-hypermutated, samples, had mutations, but only one of our three ovarian metastasis-suffering patients had somatic *APC* mutations.

We also detected significant correlation changes of mutations statuses of genes, belonging to apoptosis, in all three SSNs derived from our patients' samples, when we applied a statistical method to identify significant differential correlation changes. Another oft-present gene, *NEFH*, encodes a heavy polypeptide that is assembled into neurofilaments¹³, the main cytoskeletal components of a mature neuron¹⁴. Aberrations in neurofilaments have been implicated in numerous neurological diseases^{15–17}, although *NEFH* has been little studied in cancer^{18,19}, where it was observed in human autonomic nerve tumors and central neurocytomas^{20–22}. Hypermethylated *NEFH* may also confer poor prognosis in breast cancer²³, and its dysregulation has been observed in lung, prostate, and renal cancers^{24–26}. Our three patients showed negative correlations in *NEFH* with six genes, *PDCD4*, *ERBB2*, *BCL2L10*, *ROCK1*, *IL3RA2*, and *GSN*, in the hallmark apoptosis gene set (Fig. 3b). *NEFH*, however, had no relation with these six genes in the reference network.

In the independent targeted-gene panel sequencing analysis (Supplementary Fig. S1b), *NEFH* mutation statuses directly associated with the oncogene *ERBB2* mutation statuses, which was consistent with our dataset. *NEFH* mutation statuses also associated with *CTNNB1* mutation statuses, a member of the WNT oncogenic signaling cascade, through *ERBB2* and *SMAD7* (Supplementary Fig. 1b; patterns 2 and 6). *NEFH* also associated with

APC (another WNT pathway member, upstream of *CTNNB1*), through *GSN*, in our dataset analysis (Fig. 3b). Thus, these findings might implicate *NEFH* in WNT signaling in cancer^{19,27,28}, even as *NEFH*'s role in cancer is yet to be discovered. Recent studies^{19,27,28} have reported *NEFH* protein to be linked to Akt/beta-catenin signaling^{27,28}, as well as a tumor suppressor role¹⁹. Due to mutation status changes of the WNT genes (*APC* and *CTNNB1*), *NEFH* might lose its tumor suppressor role, leading to apoptosis dysregulation.

Metastasis is a complex, multi-step process that leads to the accumulation of genomic alterations in primary tumor cells, during malignancy progression, from the primary tumor, to localized invasion, to metastasis^{29,30}. Genomic changes, and characteristics of metastatic-stage tumors, are complex and only partially understood. Chemotherapy-naïve primary tumors, and their paired metastases, frequently share many altered genes at both sites, especially in liver metastasis³¹. However, emerging evidence suggests that treatment of colon cancer patients with anti-EGFR agents increases intra-tumor heterogeneity, leading to the emergence of clones with different genetic alterations³². To assess this, we used chemo-naïve synchronous colon cancer-to-ovary metastasis samples. Also, considering well-conserved network configurations between co-occurrence networks of the primary CRC and the metastasized ovarian networks (Supplementary Fig. S2), our systematic approach discovered that the paired primary CRC and metastasized ovarian tumors share many altered genes and less heterogeneity.

Colon cancer with ovarian metastasis is less responsive to chemotherapy, compared to non-ovarian metastases^{5,6}. Likewise, we observed no tumor response of ovarian metastases⁵, whereas the tumor regression rate for non-ovarian metastatic lesions was over 65%, in a series of twenty-two patients with colon cancer ovarian metastases⁵. Based on these reports, the ovary has been regarded as a “sanctuary” organ for resistance to systemic chemotherapy. Patient #5 has now survived over four years. Thus, the underlying mechanism of chemoresistance needs to be determined, including several tumor features such as volume, cystic properties, and mucin-richness, as these may contribute to poor sensitivity to chemotherapy⁶.

The observation of HER2 (*ERBB2*) amplification in one tumor can have major impact on treatment strategies in one patient, and for the patient #9, after 1 cycle of lapatinib, the patient was switched to palliative care, due to disease progression. Although the majority of metastatic CRC patients had HER2 (*ERBB2*) amplification³³, our study only considered mutations, not amplifications, and moreover HER2 amplification in one of our patients might be insufficient to explain biology in ovarian metastasis.

In summary, although ovarian metastasis from the colon is rare⁹, precluding (due to insufficient sample sizes) most statistical methods that merely identify common mutations, we successfully applied network biology to this disease. The result of that analysis was that these patients all shared a common biological function, i.e., a perturbed network involving the process of apoptosis. We strongly assert that such approaches will allow analyses of similarly uncommon maladies, facilitating increased understanding of rare disease pathogenesis.

Materials and Methods

Patient samples. Three patients who were pathologically diagnosed with colon cancer, with ovary metastases, were included in this study. All subjects were recruited from Gachon Gil Medical Center (Gachon University, South Korea), and the study was approved by the Institutional Review Board of Gachon University Gil Medical Center (IRB#GIRBA2216). Written informed consents were obtained from the patients, in accordance with institutional guidelines. All methods were performed in accordance with the relevant guidelines and regulation. These patients were chemo-naïve at the time of primary colon surgery, and were diagnosed with ovary metastasis during the first operation.

Patient #5. A 54-year old woman, presenting with ongoing lower abdominal pain, had obstructing distal sigmoid colon cancer, with synchronous right ovary metastasis and multiple lung metastases. The patient's initial cancer embryonic antigen (CEA) level was 85.31 ng/mL, and she underwent palliative anterior resection, with bilateral oophorectomy, when the primary tumor and ovary metastases were collected and freshly frozen. The pathologic features were as follows: wild type *KRAS*, T4a lesion, lymphovascular invasion, perineural invasion, and 9 of 20 lymph node metastases with negative resection margins. The patient underwent palliative chemotherapy with FOLFIRI (leucovorin calcium, 5-fluorouracil, and irinotecan) and cetuximab. Currently, the patient is in therapeutic remission for stable lung metastasis.

Patient #8. A 52-year-old female presented with a 3.5-cm sigmoid colon cancer with 10-cm sized right Krukenburg's tumor, and multiple lung metastases. The patient underwent palliative low anterior resection, with total hysterectomy and bilateral oophorectomy (tissue procurement of the primary tumor and ovarian metastasis). The pathologic features were as follows: adenocarcinoma, wild type *KRAS*, moderate differentiation, direct invasion of the uterine myometrium (T4b), lymphovascular invasion, perineural invasion, and 20 positively metastatic of 60 lymph nodes dissected. Thirteen months later, the patient developed a small recurrent tumor in her lumbar spine, and underwent radiation therapy (3600 cGy from 300 cGy x 12 fractions), and 12 cycles of oxaliplatin-based chemotherapy (FOLFOX6) thereafter. Shortly after chemotherapy, lung metastasis progressed. The patient still has slowly progressive, multiple lung metastases refractory to cetuximab and bevacizumab. The patient is still alive, four years after the initial surgery.

Patient #9. A 51-year old female presented with obstructing sigmoid colon cancer, with occult ovarian metastasis, and underwent anterior resection, with intraoperative bowel lavage with simultaneous bilateral oophorectomy. Pathologic examination revealed metastases in both ovaries, with wild-type *KRAS* and HER2 amplification. After 12 cycles of oxaliplatin-based chemotherapy, the regimen was switched to bevacizumab plus irinotecan-based chemotherapy. Shortly after 1 cycle of lapatinib, the patient was switched to palliative care, due to disease progression.

Clinicopathological features of the three patients are listed in Table 1, and surgical specimens from the three above-described patients, their corresponding abdominopelvic CT scans, and histologies, are shown in Fig. 4.

Whole-exome sequencing. We used exome sequencing, at 200X coverage, to profile somatic mutations in primary colon cancers, and synchronous ovary metastases.

DNA preparation. Genomic DNA was extracted from tumors and normal tissues, using a QIAamp DNA mini kit (Qiagen Inc, Valencia, CA, USA), according to the manufacturer's protocol. DNA quality and quantity were assessed using a Nanodrop spectrometer (Nanodrop Technologies, Wilmington, DE, USA) and a Qubit fluorometer (Life Technologies, Invitrogen, Carlsbad, CA, USA).

Experimental DNA sequencing. Whole-exome sequencing was performed using SureSelect Human All Exon V5 51 Mb (Agilent, Santa Clara, CA, USA), following the manufacturer's standard protocol. A paired-end DNA sequencing library was prepared through genomic DNA shearing, use a Covaris™ (Woburn, MA, USA) sonicator, followed by peak detection, end repair, poly A-tailing, paired-end adaptor ligation, and amplification. After hybridization of the library with bait sequences for 24 hours, the captured library was purified and amplified with an index barcode tag, and the library quality and quantity were determined. Sequencing of the exome library was carried out using the 100-bp paired-end mode of the HiSeq SBS kit (Illumina, San Diego, CA).

Whole-exome sequencing processing and alignment. Sequence reads in FASTQ format were mapped to the human assembly UCSC hg19, using the Burrows-Wheeler Aligner (BWA, v0.7.7)³⁴ with 'mem' and seed value parameters '-k 45', to create SAM files with correct mate pair information, with a read group tag that included the sample name. We then used Picard (v1.92) to convert SAM files to compressed BAM files, and sorted the BAM files by chromosome coordinates. The Genome Analysis Toolkit (v2.3.9Lite)³⁵ was used to locally realign the BAM files at intervals that may have insertion/deletion (indel) alignment errors. Somatic mutations and indels were "called" with Mutect³⁶ and GATK Somatic Indel Detector (Broad Institute), respectively. Single nucleotide variants and indels were annotated using snpEff (v4.1a)³⁷ to classify variants as synonymous, non-synonymous, missense, or frameshift point mutations and frameshift indels. We provide all variants detected in primary CRC tumors (single nucleotide variants in Supplementary Table S3 and indels in Supplementary Table S4) and metastasized ovarian tumors (single nucleotide variants in Supplementary Table S5 and indels in Supplementary Table S6).

TCGA sample selection for reference network construction. *Dataset from The Cancer Atlas Genome (TCGA).* We downloaded a TCGA primary colon cancer¹¹ mutation dataset from the UCSC Cancer Genomics Browser³⁸ (last accessed on 02-21-2018), namely, the version of TCGA COAD mutation database 2015-02-24. From this mutation dataset, we selected five samples under the following conditions; female, age at initial pathologic diagnosis between 46 and 59 years-old, MSI, and tumor stage IV (IVA and IVB included). The age range of the five samples was approximately close to that of our three ovarian metastasis-bearing patients. These samples were referred to as "TCGA-matched samples" throughout the manuscript. Supplementary Table S2 lists the clinicomolecular characteristics of the TCGA-matched samples selected for the reference network construction.

Mutational co-occurrence network. We calculated a correlation coefficient between mutation statuses of patients for each gene pair (genes A and B) to build mutational co-occurrence profiles, using Pearson's correlation coefficient (PCC).

$$\text{Correlation coefficient}_{AB} = \frac{\text{cov}(A, B)}{\sigma_A \sigma_B}$$

where *cov* is a covariance, and σ is a standard deviation. We then constructed a reference network of mutational co-occurrences in the TCGA-matched samples, with Pearson correlation coefficient (PCC) cutoff criteria; $|\text{PCC}| > 0.46$ and $p\text{-value} < 0.05$, according to Green *et al.*³⁹. Correlation states refer to positive, negative, and noncorrelation. A $\text{PCC} > 0.46$ was considered a positive correlation, and a $\text{PCC} < -0.46$ a negative correlation. A PCC between -0.46 and 0.46 was considered noncorrelation. Genes of interest in calculating PCCs focused on the hallmark gene sets provided by MIT MSigDB⁴⁰.

Detecting significant correlation changes between two genes. For SSN construction for our individual samples, Lie *et al.*¹⁰, showed that the statistical significance of an edge, (correlation of mutation statuses) between two genes (equivalently, nodes; say, nodes A and B), could be verified by the Z-test, and the Z-value could be defined as below:

$$Z = \frac{\Delta W_{AB}}{\frac{1 - W_{AB}^2}{n - 1}}$$

where W_{AB} is a PCC value (for the TCGA-matched samples) of an edge between nodes A and B; X_{AB} is a PCC value between nodes A and B, when the TCGA-matched samples plus our individual samples were combined; ΔW_{AB} is X_{AB} minus W_{AB} ; and n is a number of samples used to build the reference network. Edges in associated networks were selected from those correlation values that changed significantly in a statistical test. We only considered gene pairs with one type of association change (ΔW_{AB}), gain-of-correlation, which indicates noncorrelation of W_{AB} changed to either positive or negative correlation, of X_{AB} . We did not consider the case where correlation direction (e.g., positive, negative) of W_{AB} was not changed in X_{AB} , or was changed to noncorrelation.

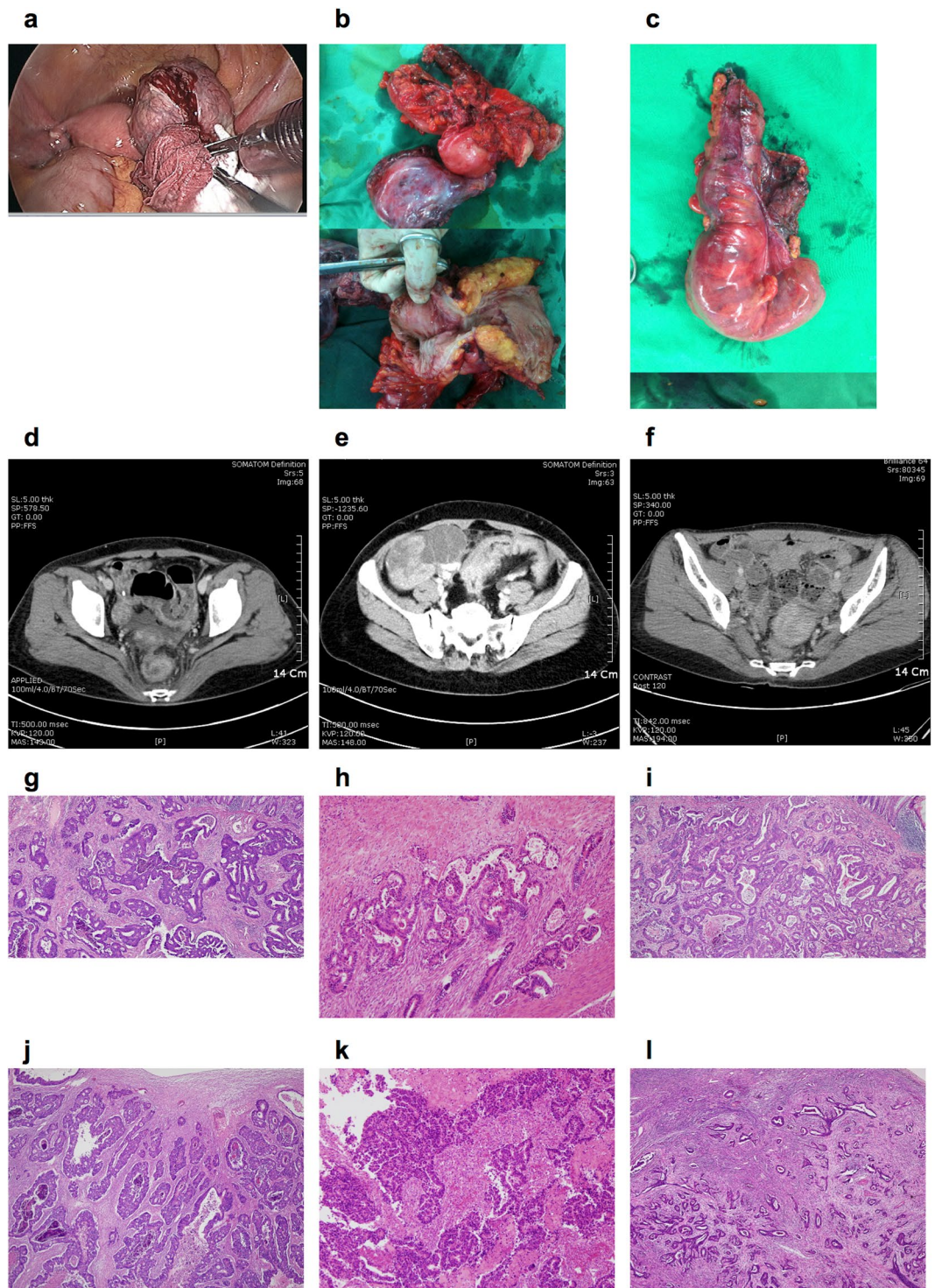


Figure 4. Macroscopic examinations of the excised tumor specimens. (a) Patient #5 underwent palliative anterior resection, with oophorectomy, when freshly frozen primary tumor and ovary metastasis tissue was collected. (b) Patient, #8 underwent palliative low anterior resection, with total hysterectomy and bilateral oophorectomy. (c) Patient #9 underwent anterior resection, with intraoperative bowel lavage and simultaneous bilateral oophorectomy. Abdominal enhanced computed tomography of patients (d) #5, (e) #8, and (f) #9. Histological analysis of the excised colon tumor specimens of (g) patients #5, (h) #8, and (i) #9. Histological analysis of the excised ovary tumor specimens of patients (j) #5, (k) #8, and (l) #9.

Targeted gene sequencing dataset of colorectal cancer patients metastasized to ovaries. To evaluate the consistency of the functional contexts of our dataset, we evaluated another colorectal cancer dataset metastasized to ovaries¹². This dataset was a targeted gene sequencing dataset using its own custom-made

115-gene panel. The entire gene list, for the panel, and list of mutations in primary COAD tumors and metastases to ovaries, are available in the supplementary materials of their report¹². From that data set, we used the mutation profiles of the primary COAD tumors. Because this dataset was from the targeted gene panel, *NEFH* was not listed in the panel. Thus, we queried the mutational occurrence ratio of *NEFH*, in TCGA COAD to cBioPortal⁴¹, and the occurrence ratio of *NEFH* aberrations was 1.5% (among 594 samples, 1 amplification, 1 deep deletion, 1 truncating mutation, and 6 missense mutations, last accessed DEC 31, 2018). However, the majority of TCGA COAD primary tumor tissue samples did not have mutations, and we assumed that there was no mutation on *NEFH* in this targeted sequencing dataset.

Data availability

NCBI BioProject accession: PRJNA489923.

Received: 12 October 2018; Accepted: 24 October 2019;

Published online: 18 November 2019

References

1. Siegel, R., Naishadham, D. & Jemal, A. Cancer statistics, 2013. *CA Cancer J Clin* **63**, 11–30, <https://doi.org/10.3322/caac.21166> (2013).
2. Patel, J. D. *et al.* Clinical Cancer Advances 2013: Annual Report on Progress Against Cancer from the American Society of Clinical Oncology. *J Clin Oncol* **32**, 129–160, <https://doi.org/10.1200/JCO.2013.53.7076> (2014).
3. Tan, K. L., Tan, W. S., Lim, J. F. & Eu, K. W. Krukenberg tumors of colorectal origin: a dismal outcome—experience of a tertiary center. *Int J Colorectal Dis* **25**, 233–238, <https://doi.org/10.1007/s00384-009-0796-x> (2010).
4. Ojo, J. *et al.* Krukenberg tumors from colorectal cancer: presentation, treatment and outcomes. *Am Surg* **77**, 1381–1385 (2011).
5. Goere, D. *et al.* The differential response to chemotherapy of ovarian metastases from colorectal carcinoma. *Eur J Surg Oncol* **34**, 1335–1339, <https://doi.org/10.1016/j.ejso.2008.03.010> (2008).
6. Lee, S. J. *et al.* Survival benefit from ovarian metastectomy in colorectal cancer patients with ovarian metastasis: a retrospective analysis. *Cancer Chemother Pharmacol* **66**, 229–235, <https://doi.org/10.1007/s00280-009-1150-2> (2010).
7. Taylor, A. E., Nicolson, V. M. & Cunningham, D. Ovarian metastases from primary gastrointestinal malignancies: the Royal Marsden Hospital experience and implications for adjuvant treatment. *Br J Cancer* **71**, 92–96 (1995).
8. Lewis, M. R., Deavers, M. T., Silva, E. G. & Malpica, A. Ovarian involvement by metastatic colorectal adenocarcinoma: still a diagnostic challenge. *Am J Surg Pathol* **30**, 177–184 (2006).
9. Patanaphan, V. & Salazar, O. M. Colorectal cancer: metastatic patterns and prognosis. *South Med J* **86**, 38–41 (1993).
10. Liu, X., Wang, Y., Ji, H., Aihara, K. & Chen, L. Personalized characterization of diseases using sample-specific networks. *Nucleic Acids Res* **44**, e164, <https://doi.org/10.1093/nar/gkw772> (2016).
11. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337, <https://doi.org/10.1038/nature11252> (2012).
12. Crobach, S. *et al.* Somatic mutation profiles in primary colorectal cancers and matching ovarian metastases: Identification of driver and passenger mutations. *J Pathol Clin Res* **2**, 166–174, <https://doi.org/10.1002/cjp2.45> (2016).
13. Liem, R. K., Yen, S. H., Salomon, G. D. & Shelanski, M. L. Intermediate filaments in nervous tissues. *J Cell Biol* **79**, 637–645 (1978).
14. Steinert, P. M. & Roop, D. R. Molecular and cellular biology of intermediate filaments. *Annu Rev Biochem* **57**, 593–625, <https://doi.org/10.1146/annurev.bi.57.070188.003113> (1988).
15. Mersiyanova, I. V. *et al.* A new variant of Charcot-Marie-Tooth disease type 2 is probably the result of a mutation in the neurofilament-light gene. *Am J Hum Genet* **67**, 37–46, <https://doi.org/10.1086/302962> (2000).
16. Skvortsova, V. *et al.* Analysis of heavy neurofilament subunit gene polymorphism in Russian patients with sporadic motor neuron disease (MND). *Eur J Hum Genet* **12**, 241–244, <https://doi.org/10.1038/sj.ejhg.5201144> (2004).
17. Kudo, L. C. *et al.* Integrative gene-tissue microarray-based approach for identification of human disease biomarkers: application to amyotrophic lateral sclerosis. *Hum Mol Genet* **19**, 3233–3253, <https://doi.org/10.1093/hmg/ddq232> (2010).
18. Alholle, A. *et al.* Functional epigenetic approach identifies frequently methylated genes in Ewing sarcoma. *Epigenetics* **8**, 1198–1204, <https://doi.org/10.4161/epi.26266> (2013).
19. Revill, K. *et al.* Genome-wide methylation analysis and epigenetic unmasking identify tumor suppressor genes in hepatocellular carcinoma. *Gastroenterology* **145**(1424–1435), e1421–1425, <https://doi.org/10.1053/j.gastro.2013.08.055> (2013).
20. Mena, H., Morrison, A. L., Jones, R. V. & Gyure, K. A. Central neurocytomas express photoreceptor differentiation. *Cancer* **91**, 136–143 (2001).
21. Miettinen, M., Shekitka, K. M. & Sobin, L. H. Schwannomas in the colon and rectum: a clinicopathologic and immunohistochemical study of 20 cases. *Am J Surg Pathol* **25**, 846–855 (2001).
22. You, H. *et al.* Immunohistochemical study of central neurocytoma, subependymoma, and subependymal giant cell astrocytoma. *J Neurooncol* **74**, 1–8, <https://doi.org/10.1007/s11060-004-2354-2> (2005).
23. Calmon, M. F. *et al.* Epigenetic silencing of neurofilament genes promotes an aggressive phenotype in breast cancer. *Epigenetics* **10**, 622–632, <https://doi.org/10.1080/15592294.2015.1050173> (2015).
24. Schleicher, R. L. *et al.* Neurofilament heavy chain-like messenger RNA and protein are present in benign prostate and down-regulated in prostatic carcinoma. *Cancer Res* **57**, 3532–3536 (1997).
25. Bobos, M., Hytiroglou, P., Kostopoulos, I., Karkavelas, G. & Papadimitriou, C. S. Immunohistochemical distinction between merkel cell carcinoma and small cell carcinoma of the lung. *Am J Dermatopathol* **28**, 99–104, <https://doi.org/10.1097/01.dad.0000183701.67366.c7> (2006).
26. Dubrowskaja, N. *et al.* Neurofilament Heavy polypeptide CpG island methylation associates with prognosis of renal cell carcinoma and prediction of antivasular endothelial growth factor therapy response. *Cancer Med* **3**, 300–309, <https://doi.org/10.1002/cam4.181> (2014).
27. Kim, M. S. *et al.* Neurofilament heavy polypeptide regulates the Akt-beta-catenin pathway in human esophageal squamous cell carcinoma. *PLoS One* **5**, e9003, <https://doi.org/10.1371/journal.pone.0009003> (2010).
28. Hummon, A. B. *et al.* Systems-wide RNAi analysis of CASP8AP2/FLASH shows transcriptional deregulation of the replication-dependent histone genes and extensive effects on the transcriptome of colorectal cancer cells. *Mol Cancer* **11**, 1, <https://doi.org/10.1186/1476-4598-11-1> (2012).
29. Fearon, E. R., Hamilton, S. R. & Vogelstein, B. Clonal analysis of human colorectal tumors. *Science* **238**, 193–197 (1987).
30. Vogelstein, B. *et al.* Genetic alterations during colorectal-tumor development. *N Engl J Med* **319**, 525–532, <https://doi.org/10.1056/NEJM198809013190901> (1988).
31. Munoz-Bellvis, L. *et al.* Unique genetic profile of sporadic colorectal cancer liver metastasis versus primary tumors as defined by high-density single-nucleotide polymorphism arrays. *Mod Pathol* **25**, 590–601, <https://doi.org/10.1038/modpathol.2011.195> (2012).

32. Adua, D. *et al.* Heterogeneity in the colorectal primary tumor and the synchronous resected liver metastases prior to and after treatment with an anti-EGFR monoclonal antibody. *Mol Clin Oncol* **7**, 113–120, <https://doi.org/10.3892/mco.2017.1270> (2017).
33. Ross, J. S. *et al.* Targeting HER2 in colorectal cancer: The landscape of amplification and short variant mutations in ERBB2 and ERBB3. *Cancer* **124**, 1358–1373, <https://doi.org/10.1002/cncr.31125> (2018).
34. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760, <https://doi.org/10.1093/bioinformatics/btp324> (2009).
35. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**, 491–498, <https://doi.org/10.1038/ng.806> (2011).
36. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80–92, <https://doi.org/10.4161/fly.19695> (2012).
37. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* **31**, 213–219, <https://doi.org/10.1038/nbt.2514> (2013).
38. Zhu, J. *et al.* The UCSC Cancer Genomics Browser. *Nat Methods* **6**, 239–240, <https://doi.org/10.1038/nmeth0409-239> (2009).
39. Green, R. A. *et al.* A high-resolution *C. elegans* essential gene network based on phenotypic profiling of a complex tissue. *Cell* **145**, 470–482, <https://doi.org/10.1016/j.cell.2011.03.037> (2011).
40. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102**, 15545–15550, <https://doi.org/10.1073/pnas.0506580102> (2005).
41. Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**, 401–404, <https://doi.org/10.1158/2159-8290.CD-12-0095> (2012).

Acknowledgements

This work was supported by funding from the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education NRF-2017R1D1A1B03035453 to W.-S. Lee. and FRD2013-34 to H.K. Ahn.

Author contributions

W.-S.L. and S.N. contributed to the conception and design of the study, and supervised this study. S.P., H.K.A., D.H.L., J.W.J. and Y.J.J. contributed to the development of methodology, acquisition, analysis, and interpretation of data. S.P. and W.-S.L. wrote the manuscript. HKA and SN helped to edit the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-019-53182-6>.

Correspondence and requests for materials should be addressed to S.N. or W.-S.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019