**Exome sequencing of synchronously resected primary colorectal tumours and colorectal liver metastases to inform oncosurgical management**

\*P A Sutton1, P V Jithesh2, R P Jones1, J P Evans1, D Vimalachandran3, H Z Malik4, B K Park1, C E Goldring1, D H Palmer1, N R Kitteringham1

1. Institute of Translational Medicine, University of Liverpool, Liverpool L69 3GE
2. Sidra Medical and Research Centre, Doha, Qatar
3. Countess of Chester NHS Foundation Trust, Liverpool Road, Chester CH2 1UL
4. Aintree University Hospital NHS Foundation Trust, Longmoor Lane, Liverpool L9 7AL

P A Sutton and P V Jithesh contributed equally to the manuscript (joint first authors)

\*Correspondence to:

Mr Paul Sutton

Institute of Translational Medicine

University of Liverpool

Sherrington Building

Ashton Street

Liverpool

L69 3GE

Tel: 0151 7942000

Email: paulsutton01@doctors.org.uk

Word Count: 3126

Abstract

*Background*

Next generation sequencing technology has facilitated mapping of the colorectal cancer genotype and furthered our understanding of metastogenesis. The aim of this study was to investigate for conserved and different mutations in the exomes of synchronously resected primary colorectal tumour and liver metastases. This information could potentially be utilised to guide the treatment of advanced disease with the help of biological information from the primary tumour.

*Methods*

We performed exome sequencing of synchronously resected primary colorectal cancer and colorectal liver metastases as well as normal colonic mucosa and liver parenchyma, from four patients who had received neo-adjuvant chemotherapy, at a depth of 50X using the Ion Proton platform. Raw data was mapped to the reference genome prior to variant calling, annotation and downstream analysis.

*Results*

Exome sequencing identified 585 non-synonymous missense single nucleotide variants (SNVs), of which 215 (36.8%) were unique to the primary tumour, 226 (38.6%) unique to the metastasis and 81 (13.8%) present in patient matched pairs. SNVs identified in the ErbB pathway appear to be concordant between primary and metastatic tumours.

*Conclusion*

Only 13.8% of the metastatic exome can be predicted by the genotype of the primary tumour. We have demonstrated concordance of a number of SNVs in the ErbB pathway, which may inform selection of therapeutic agents in advanced colorectal cancer.

Key Words: colorectal cancer, exome, next generation sequencing, stage IV, synchronous resection

Introduction

Colorectal cancer (CRC) remains the fourth commonest cancer worldwide and the second commonest cause of cancer-related death. Like most solid tumours, CRC is a highly complex disorder arising from an interplay between a number of genomic alterations, such as point mutations, rearrangements, gene fusions or copy number alterations and environmental factors such as diet, lifestyle and disease. [1,2] A failure to identify the disease at an early stage of carcinogenesis increases the likelihood of the development of metastases, which, despite significant advances both surgically oncologically, is still associated with significantly poorer outcomes. [3]

Our understanding of the development of primary colorectal cancer is improving, with sequential genetic alterations in *APC*, *KRAS*, mismatch repair genes, *TP53* and loss of 18q heterozygosity known to be significant disease drivers. However, how these cells evolve in the process of metastagenesis and the clonal origin and genetic heterogeneity of colorectal liver metastases remains unclear. The initial concept of monoclonal evolution of a tumour, essentially the clonal expansion of dominant tumour clones, has been challenged. The self-seeding polyclonal hypothesis proposes that tumour cells enter the systemic circulation via the tumour vasculature and colonise at a different site therefore establishing a new subpopulation, whereas the mutator phenotype suggests a small population of diverse tumour cells may migrate. [4] Metastases therefore can be expected to carry similar mutations to their corresponding primary tumour, but also harbour additional mutations acquired after transformation. [5] The accrual of these genetic mutations in the metastases results in tumour heterogeneity between sites, presenting problems for treating clinicians in selecting the most appropriate treatment regimen. Inter-metastatic evolution and expansion presents yet further challenges. [6]

These problems are further confounded in the case of colorectal liver metastases with the knowledge that transcutaneous needle biopsy of the liver lesion is associated with cutaneous seeding and poorer outcomes following surgery. [7] A thorough understanding of genetic heterogeneity between tumour sites is therefore vital to the development of novel therapeutic agents such as small molecule inhibitors.[8]

The aim of this study was to compare the exomes of synchronously resected primary and metastatic colorectal tumours to examine for conserved mutations between primary and metastatic tumours in order to guide the treatment of advanced disease on easily accessible tissue from the primary tumour.

Patients and Methods

National Health Service Research Ethics Committee and Research and Development approval was obtained for this study (12/NW/0011), and the study was conducted in accordance with this approval. Following informed consent, samples were obtained from four patients undergoing synchronous resection of a primary colorectal tumour and liver metastases.

*Tissue collection*

Following delivery of the colorectal specimen the proximal staple line was incised and a linear cut made down the antimesenteric border before excising a peripheral section of tumour using forceps and a scalpel. Following delivery of the liver specimen, an incision was made through the resected surface to the liver metastasis, with care being taken not to breach the liver capsule; a peripheral sample of tumour was obtained. Macroscopically normal adjacent colonic mucosa and liver parenchyma were also obtained, with all tissue sampling performed in the operating theatre under aseptic conditions followed by immediate stabilisation in liquid nitrogen. An additional representative sample was obtained for each (primary and metastatic) tumour sampled, and formalin fixed/paraffin embedded for concurrent histopathological assessment prior to sequencing. Review of hematoxylin and eosin stained slides by a consultant histopathologist confirmed the tissue to be adenocarcinoma.

*DNA extraction, quality control and library preparation*

DNA was extracted from harvested material using the DNeasy Blood & Tissue Kit as per manufacturer’s instructions (Qiagen, Venlo, Netherlands). Library preparation and sequencing was performed using Ion TargetSeq™ exome enrichment for the Ion Proton™ System (Life Technologies, Carlsbad, USA). Genomic DNA (3.5µg) was fragmented using an ion shear enzyme, and the resulting fragmented DNA cleaned up using Agencourt® AMPure® XP beads (Beckman Coulter, Brea, USA). The size distribution was checked by running an aliquot of the sample on 2% E-Gel® (Life Technologies, Carlsbad, USA). Adapter ligation was performed using the Ion Plus Fragment Library Kit according to manufacturer’s instructions (Life Technologies, Carlsbad, USA). The resulting samples were again cleaned using Agencourt® AMPure® XP beads of 300bp in length using a 2% agarose gel and were eluted using MinElute columns (Qiagen, Venlo, Netherlands). PCR amplification (10 cycles) was subsequently performed to enrich the adaptor-ligated fragments.

*Exome sequencing*

For capture, 500ng of prepared library was concentrated using a vacuum concentrator and then hybridized with the Ion TargetSeq™ Exome Probe Pool at 47°C for 66 hours. Hybridized library fragments were isolated by magnetic capture using Dynal M-270 streptavidin-coated beads (Invitrogen, Carlsbad, USA). PCR amplification (8 cycles) was carried out to amplify the captured library which was again cleaned using Agencourt® AMPure® XP beads. An aliquot of the captured library was run on a High Sensitivity DNA Chip on an Agilent Bioanalyser (Agilent, Santa Clara, USA). Real time PCR validation was performed with pre and post capture libraries to observe the capture efficiency. The purified, exome-enriched library was then used to prepare clonally amplified templated Ion PI™ Ion Sphere™ Particles (ISPs) for sequencing on an Ion PI™ Chip to obtain the necessary data coverage. Sequencing was performed on an Ion Proton™ sequencer at Genotypic Technology’s Genomics facility (Bangalore, India).

*Variant calling and annotation*

Sequenced data were analyzed with Torrent SuiteTM v 3.6 (Life Technologies, Carlsbad, USA). Following base calling, raw reads underwent trimming of undesired base calls at the 3' end of the read, the adapter sequence, and low quality 3' ends. The resultant data were filtered of low quality base calls. The raw reads obtained were then aligned to the reference genome (HG19) using the Torrent Mapping Alignment Program for Ion Torrent Data (TMAP) software. Variants were then called and detected using Variant Caller v 4.0, an inbuilt plugin of Torrent Suite v 3.6. These variants were further annotated using Ion Reporter v 1.6 to give location (intronic/exonic/utr), gene name, protein change, function and dbSNP identifier (from the dbSNP database 137).

*Somatic variant analysis*

Data from tumorous and non-tumorous tissue were compared for each patient. The variants of each sample were compared by chromosome number and position to identify those mutations in the tumour samples which were somatic. The somatic single nucleotide variants (SNVs) identified in both primary and metastatic samples were assessed for functional consequence and biologically deleterious non-synonymous SNVs were selected. The resulting SNVs were compared to identify those present in patient-matched primary and metastatic paired samples.

*Pathway and network analysis*

Genes identified from the lists of SNVs were subjected to pathway analysis using Ingenuity Pathway Analysis (Redwood City, USA). Significant canonical pathways were identified with a threshold p value of <0.05 for significance following Benjamini-Hochberg multiple test correction.

*Validation of SNVs by Sanger sequencing*

SNVs identified as being of interest during analysis were subjected to validation by Sanger sequencing on the same tissue set. Following reconfirmation of DNA quality by agarose gel electrophoresis, samples were cleaned up with an Ultraclean® PCR Clean Up Kit (Mobio, Carlsbad, USA) and quantified using the NanodropTM spectrophotometer (Thermo Scientific, Waltham, USA). Targeted forward and reverse primers were specifically designed and manufactured for each of the SNVs of interest. Sanger sequencing was performed using a BigDye® Terminator v3.1 Cycle Sequencing Kit (Life Technologies, Carlsbad, USA) and 3130XL Genetic Analyser (Life Technologies, Carlsbad, USA), with the resultant FASTA sequence compared with both the reference genome and exome sequencing data to confirm the presence of the previously identified SNVs.

Results

*Clinical information*

Demographic and clinical information pertaining to the four patients sequenced is displayed in Table 1. All patients underwent 5-fluorouracil based neoadjuvant chemotherapy followed by synchronous resection of a colonic primary and metastatic lesions at a single institution during the course of the study. All resections were R0 with no significant complications. At the time of writing, all patients were alive with no evidence of recurrent disease.

*Sequencing statistics*

All DNA samples passed the necessary quality control steps with adequate 260:280 and 260:230. The total number of reads obtained per sample ranged from ~23-54x106. Further analysis of the sequencing statistics identified that 62.96 - 83.17% of reads were successfully on target areas of the genome, with average base depth per target of 34.68 – 79.63. The result of this coverage has ensured that a minimum of 96.58% of the target (the whole exome) has been sequenced, with 67.61 – 88.33% of targets covered to a depth of at least 20X.

*Somatic Variants*

The mapped reads were subjected to variant calling based on a number of quality control criteria designed to optimise the analysis for high frequency variants and minimal false positive calls. These include minimum coverage criteria and quality scores for the SNV on each strand and a minimum observed allele frequency. The total number of variants identified can be seen in Supplementary Table S1. The analysis workflow was designed to remove germline mutations and identify somatic SNVs that are non-synonymous, missense and biologically deleterious, resulting in 585 unique SNVs.

A total of 215 SNVs (36.8%) were present in one or more of the primary tumours but not present in any of the metastases, and 226 (38.6%) were present in the metastases but not present in any of the primary tumours. A total of 81 SNVs (13.8%) were present in at least one primary and metastatic tumour pertaining to the same patient, and the remaining 63 (10.8%) were present in a primary and metastasis of different patients. Ingenuity core analysis of direct and indirect relationships revealed the significant canonical pathways in this group to be ErbB signalling, germ cell-sertoli cell junction signalling, endothelin-1 signalling, antiproliferative role of somatostatin receptor 2, and rening-angiotensin signalling. Pathway statistics for these identified pathways can be seen in Table 2.

The ErbB signalling pathway was identified as the most significant pathway arising from the analysis of SNVs meeting these criteria. Six (*KRAS, MAP2K3, PAK2, PIK3C2G, PLCG2, RRAS*) genes within this pathway (Figure 1) contained SNVs (ratio 0.07); the likelihood of this combination of SNVs occurring due to chance alone is <5.87x10-7.

Further analysis of all (unfiltered) biological pathways identified 3 networks with scores greater than 2 pertaining to endocrine system disorders, gastrointestinal disease, inflammatory disease, cellular development, nervous system development and function, tissue development, free radical scavenging, molecular transport and cellular growth and proliferation. The focus genes and statistics pertinent to these networks can be seen in Table 3.

*Validation by Sanger sequencing*

Sanger sequencing was performed for the 6 SNVs identified in the ErbB signalling pathway, which largely correlate the sequencing data (Table 4). The SNV in the *KRAS* gene was confirmed in 1 of 2 expected samples and no others. The SNV in the *MAP2K3* gene was confirmed in 2 of 2 expected samples and two others (from the same patient). The SNV in the *PAK2* gene was not confirmed in any sample. The SNV in the *PIK3C2G* gene was confirmed in 2 of 2 expected samples and one other. The SNV in the *PLCG2* gene was confirmed in 2 of 2 expected samples and no others. Finally, the SNV in the *RRAS* gene was confirmed in 2 of 2 expected samples and no others

Discussion

In this study, a comprehensive exome analysis of four synchronously resected patient matched primary and metastatic tumours was performed. Of the 585 unique non-synonymous, missense SNVs identified, 215 (36.8%) were unique to the primary tumour, 226 (38.6%) unique to the metastasis and 81 (13.8%) present in at least one primary and metastatic tumour pertaining to the same patient. The remaining 63 (10.8%) were neither unique to a tissue type nor present in a pair of samples from the same patient. The ErbB signalling pathway was identified as the most significant pathway arising from the analysis of SNVs in paired patient samples, with 6 (*KRAS, MAP2K3, PAK2, PIK3C2G, PLCG2, RRAS*) pathway genes containing SNVs (Ratio 0.07, p=5.87x10-7).

For decades our understanding of the genetics of colorectal carcinogenesis has gone unchallenged, however a complete catalogue of oncogenic drivers has yet to be determined. The largest published study to date remains The Cancer Genome Atlas Network of 276 colorectal samples subjected to exome sequencing, 97 of which also underwent whole genome sequencing (Cancer Genome Atlas Network, 2012).[9] This paper demonstrated that colon and rectal cancers had similar patterns of genomic alteration, and that most of the 24 genes found to be significantly mutated were known cancer-related genes such as *APC, TP53, KRAS, PIK3CA and SMAD4*.

A number of smaller studies, published both before and after, have attempted to improve our understanding of the genomic aberrations in colorectal cancer with three main aims: identifying clinically actionable targets for personalised therapy, improving pathway-level understanding of carcinogenesis, and identifying novel features or mutation types. A summary of these papers can be seen in Supplementary Table S2. Kloosterman *et al* demonstrated that chromothripsis is a widespread phenomenon in colorectal primary and metastatic tumours with considerable variation between sites [10]. Xie *et al* performed whole genome sequencing of patient matched primary and metastatic colorectal tumours. [11] Comparing the data to germline DNA they characterised somatic alterations including single nucleotide variants, insertions and deletions, copy number aberrations and structural alterations. The proportion of somatic SNVs shared between primary tumour and liver metastasis were 68.4% and 52.8%, compared to 13.8% in our study. Brannon *et al* performed targeted sequencing of 230 key genes in 69 matched primary and metastatic tumours, alongside whole genome sequencing of 4 patients. [12] Overall, 434 distinct non-synonymous somatic mutations were identified, with 79% concordance between primary and metastatic tumour. Lee *et al* performed exome sequencing of 15 patient matched micro-satellite stable primary and metastatic tumours. [13] In total, 1079 and 4366 mutations were identified in the primary and metastatic tumours respectively. 60.8% of those mutations in the primary tumour were non-synonymous compared to 54.9% in the metastases. [14] A more recent study reported a global concordance rate between primary and metastatic tumours of 78% despite considerable variation in neoadjuvant treatment, suggesting the mutations are inherent rather than induced by therapy. [15] Our data are complementary to these publications, and the finding of concordance in the SNVs within the ErbB family remains novel. The discrepancy in the proportion of mutations concordant between primary and metastatic tumours may possibly be explained by the biopsy of clones/subclones within a tumour, but also the inclusion of other genetic aberrations (for example indels) included in the reported figures.

The epidermal growth factor receptor (EGFR) family consists of four members, ErbB1-4. These receptors comprise a glycosylated extracellular domain and an intracellular domain with a juxtamembrane segment, protein kinase domain and carboxy terminal tail. ErbB1, ErbB3 and ErbB4 all have seven biologically occurring ligands (EGF, EPG, TGFα, AR, BTC, HB-EGFEPR), whereas ErbB2 has none. [16] Downstream signalling of ErbB activation is by three integrated pathways: the phosphatidylinositol 3-kinase/Akt (PKB) pathway, the Ras/Raf/Mek/ERK pathway and the phospholipase C pathway. These mutations result in aberrant activity of downstream pathways. [17] Given their importance in a number of cell functions, including regulation of apoptosis, cell cycle progression, cytoskeletal rearrangement, differentiation and development, and their biological implication in colorectal cancer, the ErbB family and downstream pathways have become a commonly utilised drug target (Supplementary Table 3). [18,19]

The efficacy of cetuximab in relation to *KRAS* status, and concordance of the primary and metastatic genotype in this regard has been established. Given the inter-relation of the *RAS/RAF* pathway with other downstream effectors of ErbB signalling, our study raises the question of whether the concordance seen in the case of *KRAS* can be extrapolated to other members of the ErbB pathways. As our understanding of the role of these genes in treatment response improves, this concordance could be exploited to create personalised treatment regimens. Such a move would have the potential for improved treatment efficacy as well as a reduction in unwanted off-target effects.

Our study offers a rare insight into the biology of synchronously resected patient matched primary colorectal tumour and colorectal liver metastases. Whilst such samples are challenging to acquire, the use of synchronously resected samples as described in this study, means that several potential confounding factors can be excluded from the analysis. These include exposure of the secondary tumour to chemotherapeutic agents that might show preferential cytotoxicity within the tumour mass; the influence of the initial surgical procedure to remove the primary lesion, which may result in stress responses within the secondary tumour site; the spontaneous evolution of the secondary tumour in the period between the two surgical procedures. Such factors could have a profound effect on the gentotypic characteristics of the metastasis, given that the scheduling of the two surgical interventions might involve a lengthy time period.

We opted to validate the SNVs of interest in the same tissue set but using an alternative technology, specifically Sanger sequencing. The preferred alternative would have been validation on a larger, different tissue set using either the same or a different platform. The precious nature of these difficult to acquire samples meant that this was not feasible. The incorporation of metachronous resections into the tissue set would have increased numbers (and cost) at the expense of losing the novelty and scientific value we have achieved with this study. Given this approach to validation the applicability of these findings to other matched sample sets remains unanswered, limiting this study to one which is clinically relevant and novel, but descriptive only.

The final limitation is the inherent heterogeneity in the study population. Differences existed in terms of patient age, gender, TNM stage, burden and distribution of metastases, response to treatment and the presence of adverse pathological features. It is advantageous that all of the primary tumours were confined to the colon (rather than including rectal cancers also), although there were primary tumours from both the right and left colon. It is now widely accepted that right and left colonic tumours are to be regarded as different biological entities, particularly where genomic aberrations are concerned. [20]

In conclusion our study reports that approximately one third of mutations identified in this study are present exclusively in the primary tumour, one third are present exclusively in the metastases and one third are present in tumours at both sites. In this cohort, only 13.8% of mutations were present in patient matched samples, i.e. only 13.8% of the metastatic genotype can be predicted by the genotype of the primary tumour. SNVs identified in the ErbB pathway appear to be concordant between primary and metastatic tumours. As our understanding of the relevance of this pathway (and genetic variants within) to response to treatment with chemotherapy and targeted agents grows, treatment decisions may increasingly be made on the basis of other genetic aberrations beyond the currently used *KRAS* status. This genotypic information can be acquired endoscopically from the primary tumour given its concordance with the metastatic genotype.

Conflict of Interest Statement

The authors have no conflict of interest to declare.

Acknowledgements

This work was funded by Cancer Research (UK) with sequencing performed at Genotypic Technology Ltd, Bangalore, India.

References

[1] Fearon, E. Molecular genetics of colorectal cancer. *Annual Review of Pathology.* **6,** 479-507 (2011).

[2] Huxley, R. *et al*. The impact of dietary and lifestyle risk factors on risk of colorectal cancer: a quantitative overview of the epidemiological evidence. *International Journal of Cancer*. **125,** 171-180 (2009).

[3] Cancer Research UK 2014 *Bowel Cancers Statistics* [on-line]. Available at: http://www.cancerresearchuk.org/cancer-info/cancerstats/types/bowel [Accessed 2 September 2016].

[4] Navin, N. and Hicks, J. Tracing the tumor lineage. *Molecular Oncology.* **4**, 267-283 (2010).

[5] Siegmund, K., Marjoram, P., Tavare, S. and Shibata, D. Many colorectal cancers are “flat” clonal expansions. *Cell Cycle.* **8**, 2187-2193 (2009).

[6] Gerlinger, M. *et al*. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *New England journal of medicine.* **366**, 883-892 (2012).

[7] Jones, O., Rees, M., John, T., Bygrave, S. and Plant, G. Biopsy of resectable colorectal liver  metastases causes tumour dissemination and adversely affects survival after liver resection. *British Journal of Surgery*. **92,** 1165-1168 (2005).

[8] Klein, C. Parallel progression of primary tumours and metastases. *Nature Reviews Cancer.* **9,** 302-312 (2009).

[9] Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature.* **487**, 330-337 (2012).

[10] Kloosterman, W. *et al*. Chromothripsis is a common mechanism driving genomic rearrangements in primary and metastatic colorectal cancer. *Genome Biology*. **12**: R103 (2011).

[11] Xie, T. *et al*. Patterns of somatic alterations between matched primary and metastatic colorectal tumors characterized by whole-genome sequencing. *Genomics*. **104**, 234-241 (2014).

[12] Brannon, A.R. *et al*. Comparative sequencing analysis reveals high genomic concordance between matched primary and metastatic colorectal cancer lesions. *Genome biology.* **15,** 1-8 (2014).

[13] Lee, S.Y. *et al*. Comparative genomic analysis of primary and synchronous metastatic colorectal cancers. *PloS One*. **9**, e90459 (2014).

[14] Greenman, C. *et al*. Patterns of somatic mutation in human cancer genomes. *Nature*. **446**, 153–158 (2007).

[15] Vignot, S. *et al*. Comparative analysis of primary tumour and matched metastases in colorectal cancer patients: evaluation of concordance between genomic and transcriptional profiles. *Eur J Cancer*. **51**, 791-9 (2015).

[16] Roskoski, R. The ErbB/HER family of protein-tyrosine kinases and cancer. *Pharmacological Research*. **79**, 34-74 (2014).

[17] Miettinen, P. *et al*. Epithelial immaturity and multiorgan failure in mice lacking epidermal growth factor receptor. *Nature.*  **376**, 337-341 (1995).

[18] Manning, G., Whyte, D., Martinez, R., Hunter, T. and Sudarsanam, S. The protein kinase complement of the human genome. *Science.* **298**, 1912-1934 (2002).

[19] Yao, Y. *et al*. Proliferation of colorectal cancer is promoted by two signalling transduction expression patterns: ErbB2/ErbB3/AKT and MET/ErbB3/MAPK. *PLoS One*. **8**, e78086 (2013).

[20] Zhu, H. *et al*. Screening for differentially expressed genes between left- and right-sided colon carcinoma by microarray analysis. *Oncology Letters*.  **6**, 353-358 (2013).

Legends for Figures and Tables

Table 1 - Demographic and clinical features of the four patients from whom tumours were sequenced. IrMdG – Irinotecan and modified de Gramont (5-fluoruracil and folinic acid).

Table 2 – Pathway analysis for SNVs present in at least one primary and metastatic tumour pertaining to the same patient. P-value is adjusted using Benjamini-Hochberg correction. Ratio identifies the number of focus genes identified (numerator) in the pathway (denominator).

Table 3 - Network analysis for SNVs present in at least one primary and metastatic tumour pertaining to the same patient. The score is derived from a p-value and indicates the probability of the focus genes in a network being found together due to random chance. Focus molecules is the total number of SNVs of interest identified in the respective network.

Table 4 – Summary of the Sanger sequencing for 6 SNVs of interest in 8 samples (matched primary and metastatic tumours from 4 patients). The highlighted fields were the location of SNVs identified through Ion Proton sequencing, with the content of the fields representing the identified base(s) on Sanger sequencing. Those marked undetermined failed due to inability to detect the amplicon.

Figure 1 – The ErbB signalling pathway was identified as the most statistically significant pathway arising from analysis of the SNVs present in at least one primary and metastatic tumour pertaining to the same patient. The focus genes in which the SNVs have arisen are highlighted in purple. Image from Ingenuity Pathway Analysis (Qiagen).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Patient 1** | **Patient 2** | **Patient 3** | **Patient 4** |
| Age at diagnosis | 67 | 70 | 59 | 75 |
| Gender | M | M | F | M |
| Site of primary | Ascending | Sigmoid | Sigmoid | Sigmoid |
| T stage | 4 | 3 | 3 | 3 |
| N stage | 0 | 0 | 0 | 2 |
| Metastatic burden | 16 - Bilobar | 7 - Bilobar | 3 - Bilobar | 4 - Unilobar |
| Neoadjuvant agents | IrMdG | IrMdG | Oxaliplatin, capecitabine, cetuximab | IrMdG |
| Procedure performed | Right hemicolectomy and multiple segmentectomies | Sigmoid colectomy and multiple segmentectomies | Sigmoid colectomy and multiple segmentectomies | Sigmoid colectomy and left hemi-hepatectomy |
| Pathological features (primary) | Adenocarcinoma  Well differentiated  No lympovascular invasion | Adenocarcinoma  Moderately differentiated  No lympovascular invasion | Adenocarcinoma  Moderately differentiated  No lympovascular invasion | Adenocarcinoma  Poorly differentiated  Lymphoid invasion  No vascular invasion |
| Pathological features (liver) | Adenocarcinoma  Well differentiated | Adenocarcinoma  Moderately differentiated | Adenocarcinoma  Moderately differentiated | Adenocarcinoma  Poorly differentiated |
| Tumour regression grade | 1 | 1 | 3 | 4 |

Table 1

|  |  |  |
| --- | --- | --- |
| **Pathway** | **p-value** | **Ratio** |
| ErbB signalling | 5.87E-07 | 6/85 (0.071) |
| Germ cell-sertoli cell junction signalling | 1.40E-06 | 7/156 (0.045) |
| Endothelin-1 signalling | 2.20E-06 | 7/167 (0.042) |
| Antiproliferative role of somatostatin receptor 2 | 2.37E-06 | 5/60 (0.083) |
| Renin angiotensin signalling | 2.41E-06 | 6/108 (0.056) |

Table 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ID** | **Molecules in Network** | **Score** | **Focus Molecules** | **Top Diseases and Functions** |
| 1 | A2M,ABLIM,ADCY7,Akt,ARPC1A,ATP9A,BRAF,C1QTNF9,COL18A1,Cyclin A,DUSP3,ERK,ERK1/2,FSH,GNLY,GNRH,IgG,KRAS,KRT18,LGALS3,Lh,MAP2K3,NCOA1,NFkB (complex),PAK2,PDXK,PLCG2,PLIN3,POP5,PRKX,PRSS1,PRSS3,PTPRN,TLK1,USPL1 | 23 | 13 | Endocrine System Disorders, Gastrointestinal Disease, Inflammatory Disease |
| 2 | ACAA1,ACSS2,ANAPC5,CCND1,CDC27,CDK5RAP2,CINP,CREB1,CTNNB1,DBN1,DPEP1,EPHB3,ERG,GFRA3,LAMB1,MAGI1,MAML1,MGEA5,MINK1,MYO5A,NME1,NOS1,NRP2,PABPC1,PPARA,RB1,STARD4,SVIL,TCF7,TUBA4A,UTRN,VENTX,WNT2,WNT11,ZMIZ2 | 16 | 10 | Cellular Development, Nervous System Development and Function, Tissue Development |
| 3 | ACTN2,AGTR1,AKAP13,BNIP3L,CYP3A4,DIABLO,DRD5,EPB41L3,G3BP2,GCH1,KCNJ12,LTB,MAP3K7,NOD2,NR3C1,NRIP1,PABPC3,PAWR,PLD1,PPP1R13L,PTCH1,RELA,RGS2,RIPK2,RRAS,SDHA,SHH,SNX1,SPTBN1,TADA3,TP53BP2,TRAF5,TRAF3IP2,WDR26,YWHAH | 12 | 8 | Free Radical Scavenging, Molecular Transport, Cellular Growth and Proliferation |

Table 3

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Patient 1** | | **Patient 2** | | **Patient 3** | | **Patient 4** | |
| **Annotation** | **Primary** | **Metastasis** | **Primary** | **Metastasis** | **Primary** | **Metastasis** | **Primary** | **Metastasis** |
| KRAS\_12\_25398284\_C/A | C | Undetermined | Undetermined | C | C | C | C | C and A |
| MAP2K3\_17\_21207834\_C/T | C and T | C and T | C and T | C and T | C | C | C | C |
| PAK2\_3\_196529982\_A/G | A | A | A | A | A | A | A | Undetermined |
| PIK3C2G\_12\_18650667\_A/T | A and T | A and T | A | A and T | A | A | A | A |
| PLCG2\_16\_81888178\_C/T | C and T | C and T | Undetermined | C | C | C | C | C |
| RRAS\_19\_50140130\_G/A | G and A | G and A | G | Undetermined | G | G | G | G |

Table 4

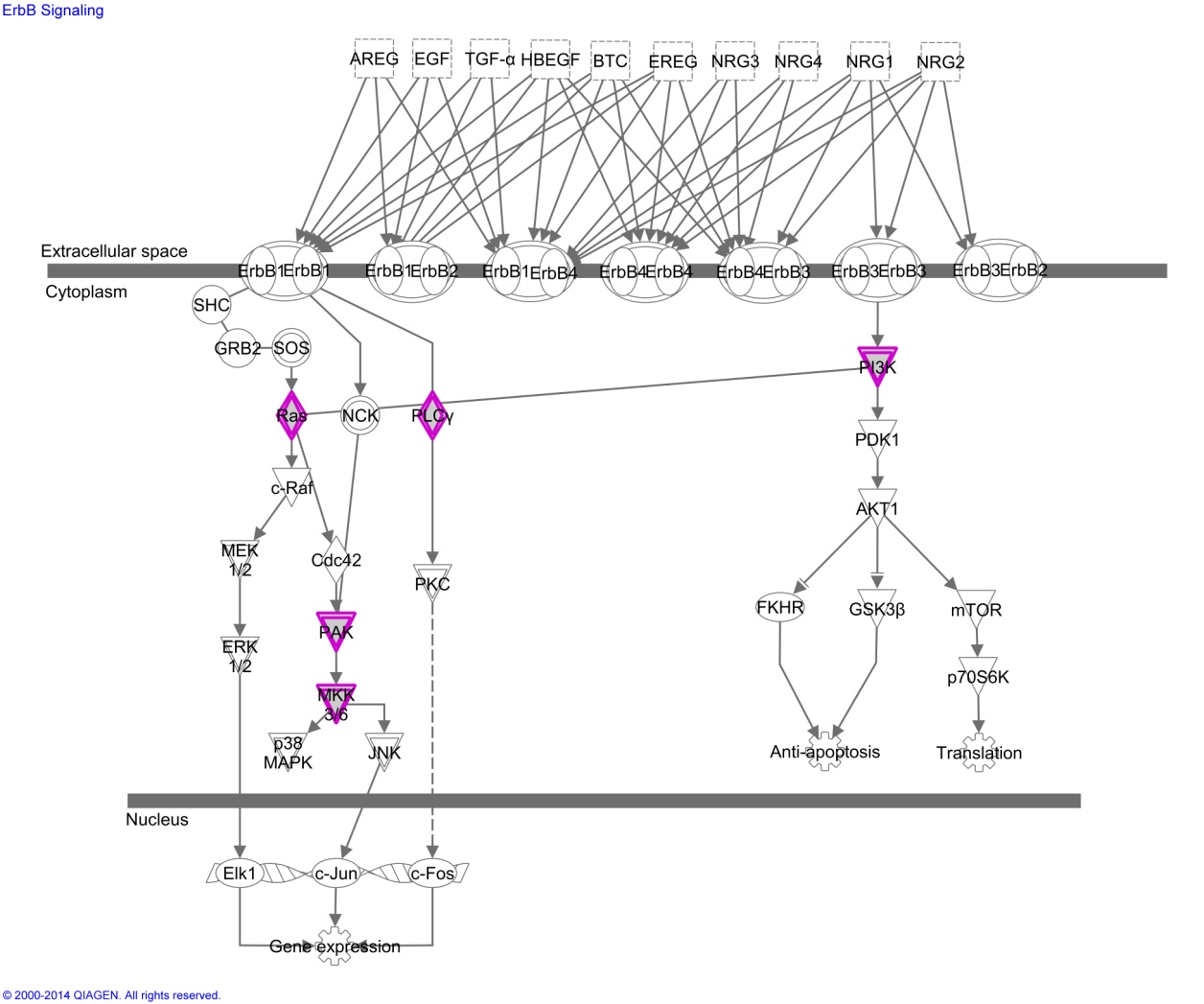


Figure 1