Human Molecular Genetics, 2008, Vol. 17, No. 24 ***3864–3875***
doi:10.1093/hmg/ddn286
Advance Access published on September 9, 2008

Putative direct and indirect Wnt targets identified through consistent gene expression changes in *APC*-mutant intestinal adenomas from humans and mice

Stefania Segditsas[1,*,†], Oliver Sieber[1,†], Maesha Deheragoda[2], Phil East[3], Andrew Rowan[1], Rosemary Jeffery[4], Emma Nye[5], Susan Clark[6], Bradley Spencer-Dene[5], Gordon Stamp[5], Richard Poulsom[4], Nirosha Suraweera[7], Andrew Silver[7], Mohammad Ilyas[8] and Ian Tomlinson[1]

[1]Molecular and Population Genetics Laboratory, [2]Histopathology Laboratory, [3]Bioinformatics and Biostatistics, [4]In Situ Hybridisation Service and [5]Experimental Pathology Laboratory, London Research Institute, Cancer Research UK, 44, Lincoln's Inn Fields, London WC2A 3PX, UK, [6]Polyposis Registry, St Mark's Hospital, Watford Road, Harrow HA1 3UJ, UK, [7]Colon Cancer Genetics, Institute of Cell and Molecular Science, Bart's and the London Medical School, Whitechapel, London, UK and [8]Department of Histopathology, University of Nottingham, Queen's Medical Centre, Nottingham, UK

Received August 13, 2008; Revised and Accepted September 6, 2008

In order to identify new genes with differential expression in early intestinal tumours, we performed mRNA (messenger ribonucleic acid) expression profiling of 16 human and 63 mouse adenomas. All individuals had germline *APC* mutations to ensure that tumorigenesis was driven by 'second hits' at *APC*. Using stringent filtering to identify changes consistent between humans and mice, we identified 60 genes up-regulated and 151 down-regulated in tumours. For 22 selected genes—including known Wnt targets—expression differences were confirmed by qRT–PCR (quantitative reverse transcription polymerase chain reaction). Most, but not all, differences were also present in colorectal carcinomas. *In situ* analysis showed a complex picture. Expression of up-regulated genes in adenomas was usually uniform/diffuse (e.g. *ITGA6*) or prominent in the tumour core (e.g. *LGR5*); in normal tissue, these genes were expressed at crypt bases or the transit amplifying zone. Down-regulated genes were often undetectable in adenomas, but in normal tissue were expressed in mesenchyme (e.g. *GREM1/2*) or differentiated cells towards crypt tops (e.g. *SGK1*). *In silico* analysis of TCF4-binding motifs showed that some of our genes were probably direct Wnt targets. Previous studies, mostly focused on human tumours, showed partial overlap with our 'expression signature', but 37 genes were unique to our study, including *TACSTD2, SEMA3F, HOXA9* and *IER3* (up-regulated), and *TAGLN, GREM1, GREM2, MAB21L2* and *RARRES2* (down-regulated). Combined analysis of our and published human data identified additional genes differentially expressed in adenomas, including decreased BMPs (bone morphogenetic proteins) and increased *BUB1/BUB1B*. Several of the newly identified, differentially expressed genes represent potential diagnostic or therapeutic targets for intestinal tumours.

*To whom correspondence should be addressed. Tel: +44 2072692884; Fax: +44 2072693093; Email: stefania.segditsas@cancer.org.uk
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

INTRODUCTION

To the limit of its current resolution, genetic evidence suggests that most early colorectal adenomas in humans, and intestinal adenomas in the equivalent mouse models, have developed solely as a result of bi-allelic inactivation of the *APC* gene. Gene expression changes in these early tumours can therefore be taken as candidate Wnt signalling targets, being driven directly or indirectly by the increase in beta-catenin-mediated signalling that loss of APC function causes.

The normal intestinal crypt is subject to a gradient of extrinsic Wnt signalling, with a maximum at the crypt base where the stem cells almost certainly reside, and a minimum at the top of the crypt or villus where fully differentiated cells are shed into the gut lumen. Three categories of putative Wnt target gene have been suggested (1). The first type of gene is expressed primarily in the proliferating or transit amplifying cells of the normal crypt, and includes established Wnt targets such as *MYC*. The second type of gene—e.g. *MMP7* and the cryptdins—is a marker of Paneth cell differentiation in the small bowel. The third type of gene principally shows expression in the bases of normal intestinal crypts, perhaps in the crypt stem cells. There are probably several reasons for these differences in expression patterns, including the actions of other signalling pathways such as the Eph/Efn, Bmp and Delta/Notch systems. Although colorectal tumours generally show over-expression of Wnt targets, it is not clear whether their molecular phenotypes most closely resemble normal cells with type 1, 2 or 3 gene expression.

Understanding the full transcriptional programme of Wnt signalling and its downstream effects is important. Proven transcriptional Wnt targets include *MYC, CCND1, CD44* and *AXIN2*. Genes with reduced levels of expression may also be found in regions of high physiological Wnt signalling or in colorectal tumours. Many of these down-regulated transcripts are likely to be genes involved in the functioning of differentiated cells. Examples include genes/proteins involved in the absorption of nutrients (e.g. lipoprotein lipases and alkaline phosphatases) or in instructing cell differentiation (e.g. *HATH1* that is involved in cell fate direction towards the secretory lineage).

There has been a handful of previous studies that aimed to identify the transcriptional programme driven by Wnt signalling. Most of these studies have been based on messenger ribonucleic acid (mRNA) expression microarray analysis. The methods and systems used have been very variable, and have included inactivation of Wnt in colorectal cancer (CRC) cell lines, activation of Wnt in cell lines from normal intestinal epithelium and study of colorectal tumours (benign and malignant) from humans. Most studies have been focused on human tumours. Perhaps unsurprisingly, there has been considerable variation between studies in their detection of known Wnt targets and in their identification of new targets. Thus, despite some notable successes, it has sometimes proved difficult to distinguish signal from noise. Reichling *et al*. (2) undertook transcriptional profiling of 14 Apc^{Min} mouse adenomas and found 114 differentially expressed genes, excluding known Wnt targets. When analysis of CRCs was added, differential expression of genes such as *Igfbp5, Lcn2, Ly6d, N4wbp4 (PMEPA1), S100c* and *Sox4* was found. Van der Flier *et al*. (1) expressed dominant-negative TCF4 or TCF1 in CRC cell lines DLD1 and LS174T. Changes were confirmed in 32 human adenomas and 25 CRCs. Consistently down-regulated genes (putative positive Wnt targets) included *AXIN2*, *MYC*, *ASCL2*, *LGR5/GPR49*, *HIG2*, *BMP4* and *SOX4*. Kaiser *et al*. (3) compared gene expression profiles in embryonic colon and colorectal tumours and found significant resemblances between them. Sabates-Bellver *et al*. (4) undertook transcriptome profiling of human colorectal adenomas from 32 patients. After micro-enrichment for epithelium, 319 possible positive and negative Wnt targets were identified by expression profiling. Of these, 144 genes displayed significantly altered expression on verification. The *KIAA1199* gene was particularly highly expressed (60-fold increased). Gaspar *et al*. (5) studied 42 polyps from 13 patients (eight *APC*- and five MYH-mutant) with three normal samples for comparison. Duodenal adenomas from three Apc^{1638N} mice and two wild-type mice were also analysed. Approximately 10% of 18 000 genes were differentially expressed in humans and, after analysis of the mouse samples, a 166-gene signature (100 up, 66 down) was derived, yet this included only a few known Wnt targets. Immunohistochemistry was used to validate CD44, ANXA1, CCNA2 and MARCKSL1.

We aimed to identify mRNAs that were consistently differentially expressed between early adenomas and morphologically normal tissue from both humans and mice. We therefore collected a much larger sample of mouse tumours for analysis than previous studies had done. In order to utilize a relatively homogeneous system and to ensure that tumorigenesis was driven by 'second hits' at *APC*, we focused on individuals with known germline *APC* mutations, either humans with familial adenomatous polyposis (FAP) or mice with the $Apc^{Min(R850X)}$ mutation. Adenomas were early lesions (<5 mm diameter, tubular morphology, mildly dysplastic), and a small number of sporadic human CRCs was analysed alongside them. Using stringent criteria, we identified a number of genes that are differentially expressed in early colorectal tumours and represent strong candidates as direct or indirect targets of Wnt signalling. We discuss these data in the light of previous studies.

RESULTS AND DISCUSSION

Expression profiling and verification

Using a stringent significance threshold of $P < 0.005$ after Benjamini–Hochberg false discovery rate (FDR) correction, we identified 220 genes that were significantly up-regulated and 548 that were significantly down-regulated by at least 2-fold when comparing normal tissue and large-bowel adenomas from human FAP patients. If multiple probes from one gene were present on the array, we scored that gene as differentially expressed if any probe fulfilled our criteria, although in no case were two different probes from the same gene significantly altered in opposite directions. The 768 genes are shown in Supplementary Material, Table S1. The list included several genes that have been described as direct or indirect targets of Wnt signalling, including *CCND1*, *CD44*, *CLU*, *EPH* receptors, *LGR5/GPR49*, *MMP7*, *MYC* and *SFRP1*. Absent from the list, because they were not significantly differentially expressed, were known Wnt targets such as

AXIN, conductin (*AXIN2*) and c-Jun. The most strongly up-regulated gene was *CDH3* (P-cadherin), which was over-expressed ~35-fold.

We then compared the list of human genes with genes that passed the same thresholds in three series of large and small bowel adenomas from Min mice on the C57BL/6J background (see Materials and Methods). In line with our stringent selection process, initial filtering required >95% of all mouse probes from all three series of polyps to show changes in the same direction as the human changes. Evidently, this strategy will have excluded some genuine direct or indirect Wnt targets, especially those-specific to each species or to different regions of the bowel, or those with sub-optimal probes on the arrays. Other genes, such as *CD44*, showed ~10% of discordant probes, presumably as a result of isoform-specific expression in the two species. *TCF4* and *CLU* (clusterin) showed an unusual pattern, with universal up-regulation in mice, but down-regulation in all human tumours.

We then excluded all genes for which no mouse probe showed significantly changed expression at $P < 0.005$. In practice, this selection largely excluded genes with few probes on the arrays and for which there was consequently low confidence in the expression changes being consistent in humans and mice. This reduced the list to a 'gene expression signature' of 60 up-regulated and 151 down-regulated genes (Table 1). This gene set included the up-regulated targets *CCND1*, *EPHB2*, *EPHB3*, *LGR5*, *MMP7* and *MYC* and the down-regulated *SFRP1*. The most up-regulated gene was *TACSTD2* (18-fold in humans, 82-fold in mouse) and the most down-regulated were *CCL19* (25-fold in humans) and *Cnn1* (33-fold in mouse). The genes performed no evident common function (apart from being in the Wnt pathway in some cases), although some related genes were present in the list, an example being *GREM1* and *GREM2*.

In order to validate the expression array data, we then selected for quantitative reverse transcription polymerase chain reaction (qRT–PCR) analysis a mixture of 22 differentially expressed genes showing up- or down-regulation (described later), including some known Wnt targets and some potentially novel targets. These genes were studied in additional sets of up to 19 FAP large bowel adenomas, seven sporadic human CRCs, 14 Min mouse small bowel adenomas, and 19 Min mouse large bowel adenomas. Actual sample numbers studied for each gene varied according to the quantity of mRNA extracted from each lesion (see Materials and Methods). *Pre hoc*, our criteria were to reject any gene for which any of the sample sets showed an expression change in the opposite direction to that found by microarray analysis and also to exclude any gene that did not show a significant expression difference (at $P < 0.05$ to reflect fewer tests) in any sample type. In practice, almost all genes passed these thresholds. We successfully validated the array data for the following up-regulated genes: *MMP7*, *LGR5*, *MYC*, *CCND1*, *ITGA6*, *PHLDA1*, *CXCL3*, *IER3*, *TGIF*, *SEMA3F*, *SOX4*, *TACSTD2* and *HOXA9*. However, *ETS2* was only over-expressed in adenomas, and actually significantly under-expressed in two of the three CRCs investigated. For down-regulated genes, we validated under-expression of *SGK1*, *GNA11*, *SFRP1*, *RARRES2*, *HPGD* and *MAB21L2*. However, whilst *GREM1* and *GREM2* were under-expressed in adenomas, these two genes showed increased expression in CRCs. In summary, all targets were successfully verified in adenomas, but in a small number of cases, gene expression levels changed on progression to carcinoma.

In situ analysis of gene expression

We then undertook *in situ* hybridization (ISH) of mRNAs using normal intestinal epithelium and adenomas from FAP patients and Min mice, in order to determine the topographical expression patterns of a subset of the consistently changed genes. Although it is difficult to make truly quantitative comparisons between gene expression levels using ISH, it is possible to make semi-quantitative comparisons using the beta-actin control probe, to assess the gradient of expression of any gene along the crypt axis, and to assess the cell types in which genes are principally expressed. In some cases, expression levels were low and reliable ISH data could not be obtained. However, clear results were obtained in the following cases.

We first analysed some established Wnt targets. *MYC* and *EPHB2* were strongly expressed in the transit amplifying zone of normal crypts (type 1 genes), and throughout the epithelium of adenomas. *MMP7* was strongly expressed only at the bases of normal crypts, but strongly or moderately diffusely in adenoma epithelium; expression appeared to be particularly strong in Paneth cells, but was not restricted to this lineage and *MMP7* could therefore be classed as of mixed type 2/3. *CD44* was moderately/strongly expressed at the crypt bases (and in villus cores of normal small bowel epithelium), and strongly in adenomas (type 3). CCND1 (assessed using immunohistochemistry) was principally expressed in the transit amplifying zone of normal crypts and widely in adenomas, with the strongest expression near the crypt tops.

LGR5 encodes an orphan G-protein coupled receptor and was originally identified as a Wnt target in CRC cell lines. It was shown to be up-regulated in hepatocellular carcinomas carrying beta-catenin mutations (6) and in colorectal and ovarian tumours (7). It has been reported as being expressed-specifically in crypt base columnar (CBC) cells, between the Paneth cells, by Barker *et al*. (8). On this basis, coupled with lineage tracing, *LGR5* has been proposed as a crypt stem (or progenitor) cell marker. We too found *LGR5* to be a type 3 gene, with expression restricted to the crypt bases, especially the CBC cells, in both large and small bowel. In adenomas, *LGR5* was generally present in patches of variable size that comprised about two-thirds of the total area of the lesion (Fig. 1A), more extensive than reported by Barker *et al*. (8). In general, these areas were close to the adenoma core, consistent with retention of some normal crypt hierarchy, some differentiation towards the surface of early adenomas, and a 'bottom-up' model of adenoma pathogenesis.

SGK1 encodes a kinase involved in the control of ion transport. It was expressed at the tops of normal crypts (Fig. 1B) and on villi, but was absent from all adenomas and carcinomas. Naishiro *et al*. (9) had found decreased expression of *SGK1* mRNA in a rat intestinal epithelial cell line that expressed a stable form of beta-catenin. Immunohistochemistry by Naishiro *et al*. (9) did not confirm the mRNA expression data, but in our hands, SGK1 antibodies have proved unsuitable for use in fixed material. Interestingly, *SGK1* expression

Table 1. Early intestinal adenoma 'expression signature'

Direction	Human Fold change	*P*-value	Mouse Fold change	*P*-value	Gene symbol	Gaspar *et al.*	Van der Flier *et al.* adenomas	Sabates-Bellver *et al.*	Hatzis *et al.*	Our Wnt
Up	17.65	0.0001	82.22	0.0010	TACSTD2*	N	N	N	Y	N
Up	10.91	0.0000	11.05	0.0000	DUSP4	N	N	N	Y	
Up	8.62	0.0000	36.12	0.0010	MMP7*	Y	N	N	N	
Up	7.60	0.0025	1.19	0.0010	SERPINB5	N	N	N	N	
Up	7.40	0.0000	1.95	0.0010	GDF15	N	N	N	Y	
Up	7.37	0.0007	49.64	0.0010	CXCL3*	Y	N	N	Y	N
Up	5.92	0.0001	4.56	0.0010	GPR49/LGR5*	N	Y	Y	N	Y
Up	5.89	0.0000	28.15	0.0010	CXCL1	Y	N	Y	N	
Up	5.84	0.0005	1.16	0.0010	FABP6	N	N	N	N	
Up	4.99	0.0000	1.29	0.0010	SLC12A2	Y	N	N	Y	
Up	4.33	0.0000	2.29	0.0010	ITGA6*	Y	N	N	Y	N
Up	3.99	0.0000	3.03	0.0010	EPHB3	Y	Y	N	Y	
Up	3.91	0.0000	7.72	0.0001	FLJ20315	N	Y	N	N	
Up	3.88	0.0005	1.29	0.0010	HYAL1	N	N	N	N	
Up	3.47	0.0003	1.32	0.0010	EPHB2	N	Y	N	Y	
Up	3.10	0.0034	2.22	0.0005	DDX31	N	N	N	N	
Up	2.99	0.0000	4.47	0.0010	SEMA3F*	N	N	N	Y	N
Up	2.95	0.0025	1.65	0.0010	HOXA9*	N	N	N	N	
Up	2.90	0.0004	8.43	0.0010	PHLDA1*	Y	Y	Y	Y	N
Up	2.82	0.0001	7.91	0.0010	SOX4*	Y	Y	N	N	
Up	2.79	0.0006	2.28	0.0010	ETS2*	Y	Y	N	Y	N
Up	2.74	0.0001	9.95	0.0005	C9orf26	N	N	N	N	
Up	2.73	0.0004	1.34	0.0010	CCT2	N	N	N	N	
Up	2.72	0.0047	1.42	0.0041	FLJ10458	N	N	N	N	
Up	2.63	0.0003	2.39	0.0010	S100A11	N	N	Y	N	
Up	2.63	0.0002	2.03	0.0001	GTF2IRD1	Y	N	N	N	
Up	2.53	0.0008	4.64	0.0024	TRIM29	N	N	N	N	
Up	2.48	0.0006	1.07	0.0010	APG-1	N	N	N	N	
Up	2.46	0.0027	2.95	0.0010	MYC*	N	Y	N	Y	
Up	2.45	0.0004	1.71	0.0010	CGI-07	N	N	N	N	
Up	2.41	0.0003	2.95	0.0010	TGIF*	N	Y	N	Y	
Up	2.41	0.0000	2.02	0.0012	SH3BP4	N	N	N	Y	
Up	2.37	0.0004	1.59	0.0010	ABCE1	Y	N	N	N	
Up	2.37	0.0002	1.98	0.0010	SLC39A8	N	N	N	Y	
Up	2.34	0.0040	2.02	0.0010	ZNF22	N	N	N	N	
Up	2.33	0.0018	1.71	0.0001	FAIM	N	N	N	N	
Up	2.31	0.0002	1.13	0.0010	RGNEF	N	N	N	N	
Up	2.29	0.0002	3.05	0.0010	CCND1*	Y	N	N	Y	
Up	2.25	0.0025	9.25	0.0010	IER3*	N	N	N	N	
Up	2.22	0.0004	2.14	0.0002	SOX9	N	Y	N	N	
Up	2.16	0.0019	2.21	0.0010	MEP50	N	N	N	N	
Up	2.16	0.0046	1.51	0.0010	CDC2	Y	N	N	N	
Up	2.14	0.0050	2.14	0.0132	TOPK	N	N	N	N	
Up	2.14	0.0003	3.00	0.0260	RHOBTB3	N	Y	N	N	
Up	2.12	0.0000	3.37	0.0000	PLEK2	N	N	N	N	
Up	2.11	0.0003	1.43	0.0010	BTG3	N	N	N	N	
Up	2.11	0.0005	1.82	0.0010	SLC39A6	N	N	N	N	
Up	2.10	0.0032	1.40	0.0010	HSPD1	N	N	N	N	
Up	2.10	0.0017	1.25	0.0010	TRIP6	Y	N	N	N	
Up	2.09	0.0022	1.77	0.0010	DTYMK	N	N	N	N	
Up	2.07	0.0048	1.37	0.0010	NPM1	N	N	N	N	
Up	2.07	0.0011	1.04	0.0010	GCSH	N	N	N	N	
Up	2.06	0.0034	1.15	0.0010	IPO7	Y	N	N	N	
Up	2.04	0.0020	1.45	0.0010	RRS1	N	N	N	N	
Up	2.04	0.0005	1.50	0.0010	CKAP2	N	N	N	N	
Up	2.02	0.0025	1.07	0.0010	RCN2	N	N	N	N	
Up	2.02	0.0005	2.44	0.0001	AASDHPPT	N	N	N	N	
Up	2.02	0.0032	1.02	0.0010	PMPCB	N	N	N	N	
Up	2.01	0.0018	4.13	0.0010	TM4SF12	N	N	N	N	
Up	2.00	0.0001	3.07	0.0011	DUSP14	N	N	N	N	
Down	24.75	0.0000	1.60	0.0010	CCL19	Y		N	N	
Down	16.47	0.0000	1.79	0.0010	CCL21	N		N	N	
Down	12.82	0.0000	1.69	0.0010	FY	N		N	N	
Down	11.47	0.0004	33.20	0.0010	CNN1	N		Y	N	
Down	10.25	0.0000	3.49	0.0010	CXCL12	N		N	N	

Continued

Table 1. *Continued*

Direction	Human Fold change	*P*-value	Mouse Fold change	*P*-value	Gene symbol	Gaspar *et al*.	Van der Flier *et al*. adenomas	Sabates-Bellver *et al*.	Hatzis *et al*.	Our Wnt
Down	9.31	0.0000	17.72	0.0010	TNA	N		N	N	
Down	8.97	0.0001	1.98	0.0010	CXCL13	N		N	Y	
Down	8.92	0.0001	2.81	0.0010	GCG	N		N	Y	
Down	8.59	0.0001	1.25	0.0010	MAP4K1	N		N	N	
Down	7.92	0.0000	2.63	0.0010	RTN1	N		Y	N	
Down	7.81	0.0014	1.02	0.0010	CCL22	N		N	N	
Down	7.67	0.0001	10.94	0.0010	TAGLN*	N		N	N	
Down	7.66	0.0001	8.91	0.0010	ADH1B	N		N	N	
Down	7.06	0.0001	1.24	0.0008	PPP1R16B	Y		N	Y	
Down	6.96	0.0000	12.96	0.0010	TNXB	N		N	N	
Down	6.69	0.0000	7.60	0.0010	GREM2*	N		N	N	Too low
Down	6.65	0.0000	1.18	0.0010	RASGRP2	N		N	N	
Down	5.90	0.0000	6.70	0.0010	MYLK	Y		Y	N	
Down	5.85	0.0042	1.30	0.0003	CRYBA2	N		N	N	
Down	5.81	0.0000	4.65	0.0010	ABCG2	Y		N	N	
Down	5.59	0.0000	1.02	0.0010	CCR7	N		N	N	
Down	5.36	0.0001	8.46	0.0010	CA1	Y		N	Y	
Down	5.28	0.0001	3.35	0.0010	HLA-DQA1 /// HLA-DQA2	N		N	N	
Down	5.14	0.0001	1.52	0.0010	ADAMDEC1	N		N	N	
Down	5.11	0.0000	1.32	0.0010	CD22	N		N	N	
Down	5.06	0.0000	1.11	0.0463	CA7	N		N	N	
Down	4.83	0.0000	2.83	0.0067	SCGN	N		N	N	
Down	4.82	0.0001	1.92	0.0010	RHOH	N		Y	N	
Down	4.76	0.0001	5.75	0.0010	CHGA	Y		N	N	
Down	4.74	0.0001	1.14	0.0010	SFRP1*	Y		Y	N	
Down	4.72	0.0034	3.73	0.0010	MAB21L2*	N		N	N	
Down	4.71	0.0001	1.67	0.0010	INSM1	Y		N	N	
Down	4.69	0.0000	1.30	0.0307	P2RX5	N		N	N	
Down	4.69	0.0000	1.24	0.0010	LSP1	N		N	N	
Down	4.69	0.0000	2.51	0.0010	RARRES2*	N		N	N	N
Down	4.64	0.0019	1.28	0.0018	BRDG1	N		Y	N	
Down	4.62	0.0001	1.27	0.0218	KIAA0657	N		N	N	
Down	4.60	0.0003	2.01	0.0010	BCHE	N		N	N	
Down	4.35	0.0001	1.85	0.0010	FOLR2	N		N	N	
Down	4.31	0.0000	5.56	0.0019	PDE2A	N		N	N	
Down	4.29	0.0001	1.28	0.0010	CD7	N		N	N	
Down	4.10	0.0037	1.45	0.0010	ANPEP	N		N	N	
Down	4.07	0.0000	1.85	0.0010	DSIPI	N		N	N	
Down	4.04	0.0003	3.15	0.0010	TNFRSF17	N		Y	N	
Down	4.01	0.0001	1.29	0.0010	RAC2	N		N	N	
Down	4.00	0.0001	1.44	0.0010	NR3C1	N		Y	N	
Down	3.99	0.0003	1.06	0.0010	GREM1*	N		N	N	Too low
Down	3.96	0.0001	1.35	0.0010	XLKD1	N		Y	N	
Down	3.96	0.0012	1.66	0.0010	AIF1	N		N	N	
Down	3.92	0.0006	1.91	0.0010	LTB	N		N	N	
Down	3.84	0.0000	1.77	0.0010	CD79A	Y		N	N	
Down	3.82	0.0000	1.70	0.0010	POU2AF1	N		N	Y	
Down	3.81	0.0010	1.30	0.0007	DKFZP434C171	N		N	N	
Down	3.81	0.0000	33.73	0.0010	CES1	N		N	N	
Down	3.81	0.0000	1.86	0.0010	HSPB2	N		N	N	
Down	3.80	0.0000	1.68	0.0010	CD48	N		Y	N	
Down	3.76	0.0014	1.46	0.0010	PCSK2	N		N	N	
Down	3.72	0.0001	1.34	0.0010	KLF2	N		N	Y	
Down	3.68	0.0001	1.28	0.0003	OLFML3	N		N	N	
Down	3.65	0.0003	2.21	0.0010	CD79B	N		N	N	
Down	3.49	0.0020	5.22	0.0010	PTPRR	N		Y	N	
Down	3.47	0.0009	1.54	0.0010	SELL	N		N	N	
Down	3.42	0.0000	2.22	0.0010	HLA-DRB1	Y		N	N	
Down	3.41	0.0022	5.30	0.0010	CCL5	N		Y	N	
Down	3.40	0.0018	1.71	0.0010	LRMP	N		N	N	
Down	3.35	0.0009	1.55	0.0005	TTID	N		N	N	
Down	3.31	0.0005	2.45	0.0010	KCTD12	N		Y	N	
Down	3.29	0.0001	12.81	0.0015	ITIH5	N		Y	N	

Continued

Table 1. *Continued*

Direction	Human Fold change	*P*-value	Mouse Fold change	*P*-value	Gene symbol	Gaspar *et al.*	Van der Flier *et al.* adenomas	Sabates-Bellver *et al.*	Hatzis *et al.*	Our Wnt
Down	3.28	0.0002	2.63	0.0010	COL13A1	N		N	N	
Down	3.22	0.0003	1.70	0.0010	GIMAP4	N		N	N	
Down	3.17	0.0007	1.14	0.0010	WASPIP	Y		N	N	
Down	3.15	0.0001	1.32	0.0005	MAP1B	N		N	Y	
Down	3.14	0.0004	11.87	0.0001	TRPM6	N		N	N	
Down	3.13	0.0003	1.10	0.0010	AKAP12	N		Y	N	
Down	3.07	0.0008	3.79	0.0010	FXYD1	N		Y	N	
Down	3.04	0.0000	5.83	0.0010	PCK1	N		N	N	
Down	3.03	0.0004	5.68	0.0007	RODH	N		N	N	
Down	3.02	0.0025	1.25	0.0010	MS4A1	N		N	N	
Down	2.99	0.0000	2.62	0.0010	CD74	N		N	N	
Down	2.98	0.0007	2.63	0.0010	CEACAM1	Y		N	N	
Down	2.96	0.0035	22.47	0.0010	CA4	N		N	N	
Down	2.95	0.0000	22.17	0.0001	DPYD	N		N	N	
Down	2.94	0.0004	1.41	0.0019	SLAMF8	N		N	Y	
Down	2.94	0.0000	2.22	0.0010	HLA-DRB3	N		N	N	
Down	2.91	0.0004	1.60	0.0010	GMFG	N		N	N	
Down	2.84	0.0009	2.64	0.0010	TCF21	N		N	N	
Down	2.84	0.0008	1.77	0.0010	L1CAM	N		Y	N	
Down	2.80	0.0004	1.78	0.0010	EDN3	Y		N	N	
Down	2.79	0.0023	8.91	0.0010	ADH1C /// ADH1A /// ADH1B	N		N	N	
Down	2.79	0.0001	1.87	0.0010	ZBTB16	N		N	Y	
Down	2.76	0.0000	1.07	0.0010	PALM2-AKAP2	N		N	N	
Down	2.76	0.0001	2.52	0.0010	EDG2	N		Y	Y	
Down	2.75	0.0026	6.75	0.0005	LOC63928	N		Y	N	
Down	2.75	0.0000	1.16	0.0010	ITGB2	N		N	N	
Down	2.74	0.0036	4.29	0.0000	OSR2	N		N	N	
Down	2.73	0.0003	1.04	0.0010	SEP06	N		N	N	
Down	2.72	0.0019	1.35	0.0105	BC008967	N		N	N	
Down	2.70	0.0000	2.54	0.0010	GNA11*	Y		N	N	N
Down	2.69	0.0000	1.82	0.0010	IL6R	Y		Y	Y	
Down	2.69	0.0025	2.02	0.0010	PLTP	N		N	N	
Down	2.67	0.0005	4.34	0.0010	CCL8	N		Y	N	
Down	2.66	0.0001	2.02	0.0010	SLC4A4	Y		N	N	
Down	2.61	0.0047	1.57	0.0010	PTPRCAP	N		N	N	
Down	2.61	0.0019	1.20	0.0010	NNMT	N		N	N	
Down	2.61	0.0004	1.41	0.0010	IL7R	N		N	N	
Down	2.60	0.0017	1.69	0.0010	CD8A	N		N	N	
Down	2.59	0.0000	1.80	0.0010	RASSF2	N		N	N	
Down	2.59	0.0004	1.33	0.0010	CD19	N		N	N	
Down	2.59	0.0006	1.55	0.0152	HOM-TES-103	N		N	N	
Down	2.58	0.0025	10.77	0.0010	HPGD*	Y		Y	Y	
Down	2.57	0.0001	5.71	0.0010	UPP1	N		N	N	
Down	2.52	0.0001	8.17	0.0001	MGC4172	Y		Y	N	
Down	2.51	0.0007	2.98	0.0010	BTK	N		N	N	
Down	2.49	0.0020	3.15	0.0010	HLA-DQB2	N		N	N	
Down	2.49	0.0032	1.64	0.0010	LAT	N		N	N	
Down	2.46	0.0050	3.42	0.0010	MFAP5	N		N	N	
Down	2.46	0.0020	2.71	0.0010	SGK1*	Y		N	N	N
Down	2.45	0.0003	1.13	0.0006	SAMSN1	N		N	Y	
Down	2.44	0.0000	1.65	0.0010	FCGRT	N		N	N	
Down	2.43	0.0021	1.66	0.0010	PER1	N		N	Y	
Down	2.41	0.0011	1.30	0.0010	PPAP2B	N		N	Y	
Down	2.41	0.0046	2.03	0.0010	RECK	N		N	N	
Down	2.41	0.0042	1.07	0.0010	C1QA	N		N	N	
Down	2.39	0.0001	4.99	0.0010	HLA-DMB	N		N	N	
Down	2.38	0.0046	6.46	0.0025	CLDN15	N		N	N	
Down	2.37	0.0042	11.61	0.0010	KCNMB1	N		N	N	
Down	2.36	0.0001	6.55	0.0031	BENE	N		N	N	
Down	2.36	0.0000	1.94	0.0010	AQP1	Y		N	N	
Down	2.34	0.0002	1.21	0.0010	ITK	N		N	N	
Down	2.30	0.0001	1.24	0.0010	CD2	N		N	N	

Continued

Table 1. *Continued*

Direction	Human Fold change	*P*-value	Mouse Fold change	*P*-value	Gene symbol	Gaspar *et al*.	Van der Flier *et al*. adenomas	Sabates-Bellver *et al*.	Hatzis *et al*.	Our Wnt
Down	2.30	0.0022	2.36	0.0010	CD3Z	Y		N	N	
Down	2.29	0.0001	1.58	0.0008	BACH2	N		N	Y	
Down	2.21	0.0001	3.42	0.0010	NR5A2	Y		Y	N	
Down	2.19	0.0040	1.15	0.0010	CD160	N		N	N	
Down	2.19	0.0028	8.32	0.0010	PCSK1N	N		N	N	
Down	2.17	0.0021	2.14	0.0235	CHRNA3	N		N	N	
Down	2.16	0.0009	1.68	0.0010	HCK	N		N	N	
Down	2.14	0.0042	3.28	0.0010	–	N		N	N	
Down	2.13	0.0003	1.29	0.0010	MXI1	Y		N	Y	
Down	2.10	0.0020	1.91	0.0000	PDLIM2	N		N	N	
Down	2.08	0.0047	1.34	0.0010	CSF1R	N		N	N	
Down	2.06	0.0046	1.48	0.0010	HOXD4	Y		N	N	
Down	2.05	0.0013	1.88	0.0010	CXCL9	N		N	Y	
Down	2.04	0.0012	9.81	0.0005	FLJ21934	N		N	N	
Down	2.03	0.0011	5.44	0.0010	SLC26A2	Y		Y	N	
Down	2.03	0.0006	17.87	0.0004	CLIC5	Y		N	N	
Down	2.02	0.0047	2.58	0.0061	FLJ11000	N		N	N	
Down	2.02	0.0012	12.55	0.0010	CHGB	N		N	N	
Down	2.02	0.0022	1.60	0.0010	EZH1	N		N	N	
Down	2.00	0.0022	1.10	0.0010	GZMA	N		N	N	
Down	2.00	0.0011	1.05	0.0010	KCNH2	N		N	N	

The table shows genes with consistent expression across all of our human and mouse series. Genes investigated further are indicated by asterisk mark. Data from the 'signatures' of the three other similar studies are shown for each gene in our list (Y = present in other signature, N = not present). The promoter TCF-binding data of Hatzis *et al*. are also shown similarly, and finally the results of our experiment to determine whether the genes are direct Wnt targets.

has actually been reported as being increased in other cancer types (10,11) and recently, Dehner *et al*. (12) found *SGK1* to be up-regulated in the CRC cell line HCT116 when Wnt signalling was increased by knock-down of *APC*. However, we have found that *SGK1* expression in colorectal carcinomas is very low, consistent with its role in ion transport which suggests that higher expression is expected at the top of the crypt and on the villus. Intestinal stem cells, adenomas and CRCs would not be expected to express *SGK1*, consistent with our data.

GNA11 is a G-protein of essentially uncertain function in the gut. It is probably involved in transmembrane signalling and may act as an activator of phospholipase C. Down-regulation of the gene has been demonstrated in breast cancers (13). *GNA11* showed moderate expression in normal epithelium, without any clear gradient or obvious restriction of expression. Expression was universally down-regulated in adenomas (Fig. 1C).

ETS2 is a transcription factor that recognizes the core consensus deoxyribonucleic acid (DNA) sequence GGAA/T upstream of target genes. *ETS2* was expressed at the base and in the lower third of the transit amplifying zone of normal crypts. Adenomas investigated by ISH showed generally increased expression, with a subtle falling-off of levels towards the tops of crypts (Fig. 1D). Interestingly, Sussan *et al*. (14) showed that trisomy 21 repressed intestinal tumorigenesis largely through *ETS2* over-expression. In line with this effect, *ETS2* may play a role in promoting apoptosis (15), and intestinal adenomas exhibit increased rates of apoptosis (our unpublished data). *ETS2* is over-expressed in several types of cancer, although out of the three sporadic colorectal carcinomas we investigated by qRT–PCR, only one showed 3.5-fold up-regulation of the gene, while the other two showed decreased expression (1.6 and 9.0-fold respectively), suggesting that its functional effects may be context-dependent.

GREM1 and *GREM2* were expressed at relatively low levels in the sub-epithelial myofibroblasts and muscularis mucosa of the normal intestinal epithelium. Expression was not detectable in adenomas. These proteins are secreted bone morphogenetic protein (BMP) antagonists that are 55% identical to each other within a 173 amino acid core. An SNP near *GREM1* is associated with increased risk of CRC in humans (16). The pattern of Gremlin expression in normal tissue (Fig. 1E) partially resembles that detected previously by Kosinski *et al*. (17) who compared gene expression in the tops and bottoms of crypts. It is also strongly suggestive of a BMP gradient along the crypt, with levels highest at the top and lowest near the stem cells at the bottom. However, we and others (18,19) have found increased Gremlin expression in cancers, apparently at odds with decreased expression in our adenomas. It is possible that these observations reflect a decreased ratio of muscularis mucosa to epithelium in benign lesions, but a tendency for malignancies to acquire their own Gremlin expression—and perhaps a mesenchyme-like phenotype—in order to maintain a stem cell-like state independent of the surrounding tissues.

ITGA6 is a member of the integrin alpha subunit family involved in cell adhesion and signalling. Up-regulation of the gene has been identified in oesophageal adenocarcinomas (20) and was found to be necessary for survival of a highly tumorigenic sub-population of breast cancer cells, through roles in adhesion and/or signalling (21). *ITGA6* mRNA

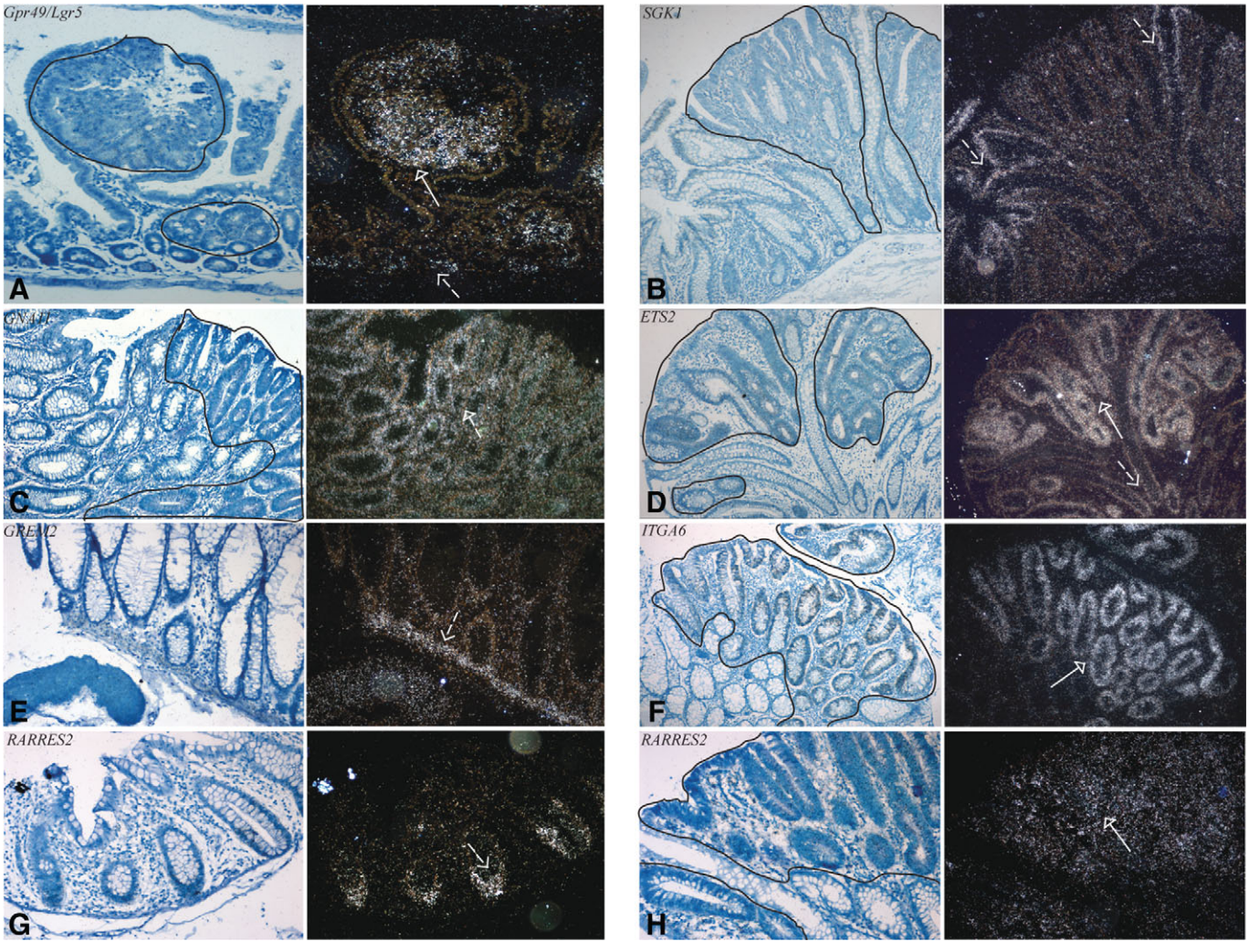

Figure 1. *In situ* hybridization analysis of selected genes showing differential expression in *APC*-mutant adenomas and normal tissue. Representative images (human adenomas unless otherwise stated) of ISH (×20 magnification) for *Gpr49/Lgr5* (mouse adenoma shown, **A**), *SGK1* (**B**), *GNA11* (**C**), *ETS2* (**D**), *GREM2* (normal tissue only, **E**), *ITGA6* (**F**) and *RARRES2* (normal tissue, **G** and adenoma, **H**). Left panels of each pair are Giemsa-stained sections. Right panels are dark-field images highlighting autoradiographic silver. Solid lines on Giemsa-stained images delineate tumour tissue. White solid arrows on dark field images illustrate gene expression in tumour tissue. Dashed arrows indicate gene expression in normal tissue.

expression was essentially absent in normal intestine, but there was strong up-regulation throughout adenomatous epithelium (Fig. 1F).

RARRES2 showed an unusual change between normal tissue and adenomas. In the normal crypt, expression was concentrated around the base (Fig. 1G). Although there was a net decrease in *RARRES2* expression in adenomas, it became more widespread, involving the whole epithelium at a low level (Fig. 1H). *RARRES2* is a retinoic acid receptor, although its function is poorly described.

Comparison and pooled analysis with other studies

We compared our data with those of three relatively large studies that have been published (1,4,5), although one of these studies (1) only reported genes that were up-regulated when Wnt was activated. Of our 60 + 151 'gene signature', 19% of genes were present in the Gaspar data set (1859 genes with FDR < 0.005), 18% in the van der Flier data set (160 up-regulated genes with FDR < 0.05) and 14% in the Sabates-Bellver data (genes with >4-fold expression difference and FDR < 0.01). By way of comparison, 16% of the genes reported by van der Flier *et al*. were also present in the list of Gaspar *et al*. The following up-regulated genes were present in our list and at least two of the other lists: *GPR49/LGR5*, *EPHB3*, *PHLDA1*, *SOX4* and *ETS2*. A further 17 up-regulated genes, including *MYC* and *CCND1*, were present in our list and in one other list. For our down-regulated genes, 44 (including *GNA11* and *SGK1*) were present in at least one of the other two studies, and seven (*MYLK, SFRP1, IL6R, HPGD, MGC4172, NR5A2* and *SLC26A2*) were listed in all three studies. Thirty-seven genes were unique to our list, including the following that were confirmed by qRT–PCR: *TACSTD2, SEMA3F, HOXA9* and *IER3* (all up), and *TAGLN, GREM1, GREM2, MAB21L2* and *RARRES2* (all down).

Since our sequential filtering strategy favours-specificity at potentially great cost in sensitivity, we also undertook a pooled analysis of our human data and the publicly available results of Sabates-Bellver *et al*. (4) who used an Affymetrix platform which was compatible with our analysis. A combined analysis of our data from human FAP adenomas with those of Sabates-Bellver was performed, with the *caveat* that *APC* mutation status was not characterized in the Sabates-Bellver sample set. The 1676 genes with a >2-fold difference in expression, using a FDR threshold of 0.001 and after correction for variation between experiments, are shown in Supplementary Material, Table S3. The combined analysis confirmed most of the putative Wnt targets reported in this study and previously. In addition, the following targets were of note and were supported in both studies. *CDH3*, highly expressed in our human data but not universally so in mouse tumours and not reported in all other studies, was strongly supported by the combined analysis. Similarly, *KIAA1199* reported as a possible Wnt target by Sabates-Bellver *et al*., was also over-expressed in our study and supported by the combined analysis. BMP expression, with the exception of *BMP4*, was generally decreased, despite the decrease in expression of the BMP antagonists *GREM1* and *GREM2*. This may favour a stem cell phenotype, but appears to be a paradoxical result, and may be secondary to changes in the Gremlins. Hypoxia pathway components, such as carbonic anhydrases, were also generally decreased. Other previously unreported, but potentially interesting targets included the spindle assembly checkpoint genes *BUB1* and *BUB1B*, which were over-expressed; given that APC has a proposed (though controversial) role in chromosomal instability through chromosome mis-segregation, it is possible that BUB1/BUB1B are over-expressed in response to mitotic checkpoint activation (22).

Differential expression and TCF-binding

In order to test whether some of our differentially expressed genes were direct Wnt targets, we searched using Fuzznuc from Emboss for the presence of putative TCF4-binding sites (WWCAAWG) in the 10 kb upstream of the translation start site of each transcript in humans (data from https://www.ensembl.org/). We examined whether genes with significant expression differences in human tumours tended to have more upstream TCF4 sequences than genes without such differences.-specifically, in a logistic regression analysis, we set as the dependent variable whether a gene was significantly differentially expressed and used as independent variables: (i) distance from the translation start site in 500 bp bins, (ii) mean number of TCF4-binding motifs in each bin over all genes, and (iii) local GC content in that bin. The mean number of TCF4 sequences was significantly higher ($P = 0.001$) for significantly differentially expressed genes (mean = 0.68 per gene) than other genes (mean = 0.62), even after stratifying for local GC content. There was overall no effect of distance from translation start site on motif frequency, although it was notable that the first 500 bp upstream, which was GC-rich and hence least likely to have TCF4-binding motifs that occurred by chance, showed the highest difference between significantly and non-significantly differentially expressed genes (means 0.46 and 0.39 binding sites per gene respectively, $P < 0.0001$, Wilcoxon test).

We tested whether some-specific members of our Wnt 'signature' were likely to be direct targets. We used the CRC cell line RKO in which Wnt signalling is not active and, owing to reported issues of toxicity if beta-catenin is over-expressed in these cells, we activated Wnt signalling using LiCl, an inhibitor of GSK3-beta. New protein synthesis was inhibited using cycloheximide. We found that *GREM1* and *GREM2* were already expressed at very low levels prior to Wnt activation, consistent with their origins from non-epithelial cells. *LGR5* showed 2.1-fold increased expression, consistent with it being a direct positive Wnt target. However, all of the other genes tested (*CXCL3*, *ETS2*, *GNA11*, *ITGA6*, *PHLDA1*, *SEMA3F*, *SGK1*, *TACSTD2*) showed less than the conventional threshold of 1.5-fold change in mRNA levels, suggesting that they are not direct positive or negative Wnt targets.

We then determined whether our 'Wnt signature' genes had been detected as possible targets of TCF4/beta-catenin binding using chromatin immunoprecipitation assays in the study of Hatzis *et al*. (23). We examined the top 2148 genes reported by Hatzis *et al*., containing a total of 6868 high-confidence binding sites. Of the 21 most strongly over-expressed genes in our human tumours (Table 1), 11 had high-confidence TCF4 -binding sites (*TACSTD2*, *DUSP4*, *GDF15*, *CXCL3*, *SLC12A2*, *ITGA6*, *EPHB3*, *EPHB2*, *SEMA3F*, *PHLDA1* and *ETS2*). In all, 16/60 over-expressed genes, but only 18/151 under-expressed genes, were associated with binding sites ($P = 0.0086$, χ^2_1 test). Whilst the possibility of direct negative Wnt targets cannot be excluded, this result confirms that Wnt activation generally leads to increased direct target expression. It also suggests that some positive Wnt target genes—such as *TACSTD2* and *SEMA3F*, which were not over-expressed when we activated Wnt in the absence of new protein synthesis—may contain TCF4-binding sites that are not involved in causing increased gene expression *in vivo*.

Concluding remarks

Using high-stringency filtering, we have identified a set of 60 up-regulated and 151 down-regulated genes that show consistent expression differences between intestinal adenomas and normal tissue from humans and mice. This list includes known Wnt targets. The remaining 'novel' genes are potential Wnt targets, mostly of the indirect type, but many are likely to be indirectly influenced by Wnt or by the cellular composition of adenomas compared with normal epithelium. All genes tested that were differentially expressed by microarray analysis were confirmed by qRT–PCR. However, ISH showed that the actual localization of the differentially expressed mRNAs was highly variable. *In situ* analysis is therefore very important when examining genes with differential expression in early intestinal tumours.

It is notable that the several studies to identify putative Wnt targets using mRNA expression analysis in colorectal tumours have found only limited overlap. Inevitably, some 'targets' are false-positives, but the differences also reflect the use of different samples and mechanisms of Wnt activation, and different methods of analysis. For example, we deliberately

excluded some differentially expressed genes for which mRNA splicing appeared to be altered between normal and tumour tissues, since comparison between the levels of these genes is extremely difficult. Overall, our study seems comparable with the three other most similar studies in the field (1,4,5) in that there was considerable, but incomplete, overlap for a core set of ~20–30 genes that are generally accepted as differentially expressed under Wnt-ON and Wnt-OFF conditions. Many other real targets were probably present in our list, the stringency of our filtering being reflected in the fact that only 37 of our 215 differentially expressed genes were not detected by any of the other three studies. This assertion was confirmed by our pooled analysis with the Sabates-Bellver data (4), in which <25% of the genes in our 60 + 151 'gene signature' genes were rendered absent from the combined data using a filter of FDR < 0.001.

In summary, the identification of genes with expression changes in early colorectal adenomas helps to identify the mechanisms of colorectal tumorigenesis and to provide potential targets for prevention and therapy. It is essential that a multi-stage approach is taken to this work, both for validation and so that different strands of evidence can be incorporated, including expression profiling and the discovery of TCF4/beta-catenin-binding sites. In this study, we have reported our own expression profiling data and some new early and consistent changes in human and mouse intestinal adenomas. We have also undertaken comparative and pooled analyses that between them have enabled us to propose new direct or indirect Wnt targets with some confidence. Several of the newly identified, differentially expressed genes represent potential diagnostic or therapeutic targets for intestinal tumours.

MATERIALS AND METHODS

Microarray expression analysis

Initially, three 'classical' FAP patients carrying germline mutations (Q163X, S1190X and delATnt2396) in the *APC* gene were identified. Sixteen snap-frozen adenomas (all with previously characterized somatic *APC* mutations) were collected from these patients; all lesions were mildly dysplastic adenomas of <0.5 cm diameter derived from different sites within the large bowel. Eight snap-frozen samples of normal bowel were collected from the same three patients and two additional FAP cases. mRNA was extracted from each sample with the GenElute mammalian total RNA miniprep kit (Sigma, RTN70) All subsequent steps in the microarray expression analysis were carried out by the Cancer Research UK GeneChip Microarray Service. Briefly, after quality control, the RNAs were reverse-transcribed to double stranded complementary DNA (cDNAs) and biotin-labelled complementary RNA (cRNA) was amplified by *in vitro* transcription with T7 RNA polymerase. cRNAs were then hybridized to HGU133A Affymetrix Chips. Full protocols are available at http://bioinformatics.picr.man.ac.uk/mbcf/protocols.jsp

For $Apc^{Min/+}$ (Min) mice, three series of small- or large-bowel adenomas and normal intestinal epithelium were studied. All samples were removed at autopsy and snap-frozen. For series 1, mRNA was extracted and cRNA prepared from 24 small-bowel adenomas and 24 normal samples from six mice as described earlier; samples from each mouse were pooled, before hybridization to 12 Affymetrix MOE430_2 arrays using the manufacturer's protocols. For series 2, tissue was homogenized using the FastRNA Kit (Q-Biogene) and total RNA prepared using the RNEasy kit (Qiagen). RNAs from up to five tumours or normal epithelium from eight individual mice were pooled before hybridization to the Affymetrix MG_U74Av2 arrays using the manufacturer's protocols; this resulted in the following pools being analysed – two tumours and two normal from the large bowel, and four tumours and two normal from the small bowel. For series 3, samples were prepared as for series 1; RNAs from nine small bowel adenomas and from six normal epithelial samples were each hybridized to Affymetrix MOE430_2 arrays.

Quantitative reverse transcription polymerase chain reaction

In order to provide further evidence of the changes identified by expression microarrays, mRNA was extracted from snap-frozen human FAP samples (on average, eight normal samples, 10 adenoma samples and four colorectal carcinoma samples were investigated for each gene) and $Apc^{Min/+}$ mouse samples (on average, 32 normal and 29 tumour samples) with GenElute mammalian total RNA miniprep kit (Sigma, RTN70) and converted to cDNA with First Strand cDNA synthesis kit (Amersham biosciences) according to the manufacturer's protocol. cDNAs were purified using the QIAquick PCR purification kit (Qiagen, 28104) and 30 ng of cDNA were used for each reaction. TaqMan Gene Expression Assays (Applied Biosystems) were used for each gene of interest. Details are available from the authors on request. Absolute quantification qRT–PCR was performed on the ABI 7900HT (Applied Biosystems) according to the manufacturer's instructions and data were analysed with the comparative Ct method (with *GAPDH* serving as endogenous Control) as described in Applied Biosystems's User Bulletin No. 2.

For analysis of target mRNA expression in RKO, cells were seeded and left to grow to ~80% confluence. The cells were then treated with the protein synthesis inhibitor cycloheximide (60 ug/ml, Sigma) for 2 h, before addition of 20 mM LiCl (test) or KCl (Control). No-treatment and cycloheximide-only samples were also analysed. The cells were incubated for 24 h and RNAs were then extracted from pelleted cells using the GenElute total mammalian RNA kit (RTN-70, Sigma). RNAs were reverse transcribed into cDNA using the AffinityScript^{TM} QPCR cDNA Synthesis Kit (Stratagene), according to the manufacturer's protocol. cDNAs were purified using the QIAquick PCR purification kit (QIAGEN) and used (30 ng per reaction) to assess expression of the putative Wnt target genes *CXCL3*, *ETS2*, *GNA11*, *GREM1*, *GREM2*, *ITGA6*, *LGR5*, *PHLDA1*, *RARRES2*, *SEMA3F*, *SGK1* and *TACSTD2* with TaqMan Gene Expression Assays (details available on request), according to manufacturer's protocol. Each sample was run in duplicate and the data obtained were analysed using the comparative Ct method (as described in Applied Biosystems User Bulletin No. 2), using *GAPDH* as an endogenous Control and comparing LiCl- with KCl-treated

cells. The fold change in mRNA expression levels was calculated as 2^{-ddCt}.

In situ hybridization

Isotopic ISH was performed essentially as previously described (24) on the following genes: *MYC*, *EPHB2*, *MMP7*, *CD44*, *GPR49/LGR5*, *SGK1*, *GNA11*, *ETS2*, *GREM2*, *ITGA6* and *RARRES2*. Briefly, templates for riboprobe synthesis were generated by amplifying from cDNA PCR products of ~500 bp, spanning exon–exon boundaries, and cloning them into the vector pGEM3Z. All templates were sequenced prior to riboprobe synthesis. Details of primers used and correspondence between ISH probes and TaqMan Gene Expression Assays are shown in Supplementary Material, Table S3. Riboprobes were generated by *in vitro* transcription using the RNA polymerase Sp6 and hybridized to 4 μm sections, cut from eight formalin-fixed, paraffin-embedded specimens. A beta-actin probe was generated and used as a hybridization control.

Data analysis

Raw gene expression data were imported into the GeneSpring program and the Mas5 algorithm was used to average across the different probes per gene present in each array. Data for each chip were normalized by median-centering (i.e. to the median of all values on a given array). Data were included if a probeset was present in at least one sample probe set per data series. Differential expression was defined as exceeding a specified fold change in either direction with a specified degree of confidence; in general, higher fold-changes and lower *P*-values or FDRs were specified for the initial screens than for the verification experiments. For meta-analysis of our own data and those of others, signal intensity estimates for the extra data sets were Mas5 generated and the data median centred as for our own human data set. Since both data sets were from Affymetrix arrays, probesets with identical IDs were identified and analysed. Tumour and experiment effect variance was modelled using ANOVA. Differential probe sets between the adenoma and normal cases were identified using a 0.001 FDR correcting for any experiment effect across the samples. The analysis was carried out using the Limma and Affy packages from Bioconductor 2.2 within R 2.7.0.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* Online.

ACKNOWLEDGEMENTS

We are grateful to colleagues who supplied tissue, to those involved in animal husbandry and to several groups who undertook vital aspects of the expression microarray analysis.

Conflict of Interest statement. None declared.

FUNDING

The study was funded by Cancer Research UK. Funding to pay the Open Access publication charges for this article was provided by Cancer Research UK.

REFERENCES

1. Van der Flier, L.G., Sabates-Bellver, J., Oving, I., Haegebarth, A., De Palo, M., Anti, M., Van Gijn, M.E., Suijkerbuijk, S., Van de Wetering, M., Marra, G. *et al.* (2007) The intestinal Wnt/TCF signature. *Gastroenterology*, **132**, 628–632.
2. Reichling, T., Goss, K.H., Carson, D.J., Holdcraft, R.W., Ley-Ebert, C., Witte, D., Aronow, B.J. and Groden, J. (2005) Transcriptional profiles of intestinal tumors in Apc(Min) mice are unique from those of embryonic intestine and identify novel gene targets dysregulated in human colorectal tumors. *Cancer Res.*, **65**, 166–176.
3. Kaiser, S., Park, Y.K., Franklin, J.L., Halberg, R.B., Yu, M., Jessen, W.J., Freudenberg, J., Chen, X., Haigis, K., Jegga, A.G. *et al.* (2007) Transcriptional recapitulation and subversion of embryonic colon development by mouse colon tumor models and human colon cancer. *Genome Biol.*, **8**, R131.
4. Sabates-Bellver, J., Van der Flier, L.G., de Palo, M., Cattaneo, E., Maake, C., Rehrauer, H., Laczko, E., Kurowski, M.A., Bujnicki, J.M., Menigatti, M. *et al.* (2007) Transcriptome profile of human colorectal adenomas. *Mol. Cancer Res.*, **5**, 1263–1275.
5. Gaspar, C., Cardoso, J., Franken, P., Molenaar, L., Morreau, H., Moslein, G., Sampson, J., Boer, J.M., de Menezes, R.X. and Fodde, R. (2008) Cross-species comparison of human and mouse intestinal polyps reveals conserved mechanisms in adenomatous polyposis coli (APC)-driven tumorigenesis. *Am. J. Pathol.*, **172**, 1363–1380.
6. Yamamoto, Y., Sakamoto, M., Fujii, G., Tsuiji, H., Kenetaka, K., Asaka, M. and Hirohashi, S. (2003) Overexpression of orphan G-protein-coupled receptor, Gpr49, in human hepatocellular carcinomas with beta-catenin mutations. *Hepatology*, **37**, 528–533.
7. McClanahan, T., Koseoglu, S., Smith, K., Grein, J., Gustafson, E., Black, S., Kirschmeier, P. and Samatar, A.A. (2006) Identification of overexpression of orphan G protein-coupled receptor GPR49 in human colon and ovarian primary tumors. *Cancer Biol. Ther.*, **5**, 419–426.
8. Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J. *et al.* (2007) Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature*, **449**, 1003–1007.
9. Naishiro, Y., Yamada, T., Idogawa, M., Honda, K., Takada, M., Kondo, T., Imai, K. and Hirohashi, S. (2005) Morphological and transcriptional responses of untransformed intestinal epithelial cells to an oncogenic beta-catenin protein. *Oncogene*, **24**, 3141–3153.
10. Chung, E.J., Sung, Y.K., Farooq, M., Kim, Y., Im, S., Tak, W.Y., Hwang, Y.J., Kim, Y.I., Han, H.S., Kim, J.C. *et al.* (2002) Gene expression profile analysis in human hepatocellular carcinoma by cDNA microarray. *Mol. Cells*, **14**, 382–387.
11. Leong, M.L., Maiyar, A.C., Kim, B., O'Keeffe, B.A. and Firestone, G.L. (2003) Expression of the serum- and glucocorticoid-inducible protein kinase, Sgk, is a cell survival response to multiple types of environmental stress stimuli in mammary epithelial cells. *J. Biol. Chem.*, **278**, 5871–5882.
12. Dehner, M., Hadjihannas, M., Weiske, J., Huber, O. and Behrens, J. (2008) Wnt signaling inhibits forkhead box O3a-induced transcription and apoptosis through up-regulation of serum- and glucocorticoid-inducible kinase 1. *J. Biol. Chem.*, **283**, 19201–19210.
13. Asada, K., Miyamoto, K., Fukutomi, T., Tsuda, H., Yagi, Y., Wakazono, K., Oishi, S., Fukui, H., Sugimura, T. and Ushijima, T. (2003) Reduced expression of GNA11 and silencing of MCT1 in human breast cancers. *Oncology*, **64**, 380–388.
14. Sussan, T.E., Yang, A., Li, F., Ostrowski, M.C. and Reeves, R.H. (2008) Trisomy represses Apc(Min)-mediated tumours in mouse models of Down's syndrome. *Nature*, **451**, 73–75.
15. Wolvetang, E.J., Wilson, T.J., Sanij, E., Busciglio, J., Hatzistavrou, T., Seth, A., Hertzog, P.J. and Kola, I. (2003) ETS2 overexpression in transgenic models and in Down syndrome predisposes to apoptosis via the p53 pathway. *Hum. Mol. Genet.*, **12**, 247–255.

16. Jaeger, E., Webb, E., Howarth, K., Carvajal-Carmona, L., Rowan, A., Broderick, P., Walther, A., Spain, S., Pittman, A., Kemp, Z. *et al.* (2008) Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk. *Nat. Genet.*, **40**, 26–28.
17. Kosinski, C., Li, V.S., Chan, A.S., Zhang, J., Ho, C., Tsui, W.Y., Chan, T.L., Mifflin, R.C., Powell, D.W., Yuen, S.T. *et al.* (2007) Gene expression patterns of human colon tops and basal crypts and BMP antagonists as intestinal stem cell niche factors. *Proc. Natl Acad. Sci. USA*, **104**, 15418–15423.
18. Sneddon, J.B., Zhen, H.H., Montgomery, K., van de Rijn, M., Tward, A.D., West, R., Gladstone, H., Chang, H.Y., Morganroth, G.S., Oro, A.E. *et al.* (2006) Bone morphogenetic protein antagonist gremlin 1 is widely expressed by cancer-associated stromal cells and can promote tumor cell proliferation. *Proc. Natl Acad. Sci. USA*, **103**, 14842–14847.
19. Namkoong, H., Shin, S.M., Kim, H.K., Ha, S.A., Cho, G.W., Hur, S.Y., Kim, T.E. and Kim, J.W. (2006) The bone morphogenetic protein antagonist gremlin 1 is overexpressed in human cancers and interacts with YWHAH protein. *BMC Cancer*, **6**, 74.
20. Hourihan, R.N., O'Sullivan, G.C. and Morgan, J.G. (2003) Transcriptional gene expression profiles of oesophageal adenocarcinoma and normal oesophageal tissues. *Anticancer Res.*, **23**, 161–165.
21. Cariati, M., Naderi, A., Brown, J.P., Smalley, M.J., Pinder, S.E., Caldas, C. and Purushotham, A.D. (2008) Alpha-6 integrin is necessary for the tumourigenicity of a stem cell-like subpopulation within the MCF7 breast cancer cell line. *Int. J. Cancer*, **122**, 298–304.
22. Fodde, R., Smits, R. and Clevers, H. (2001) APC, signal transduction and genetic instability in colorectal cancer. *Nat. Rev. Cancer*, **1**, 55–67.
23. Hatzis, P., van der Flier, L.G., van Driel, M.A., Guryev, V., Nielsen, F., Denissov, S., Nijman, I.J., Koster, J., Santo, E.E., Welboren, W. *et al.* (2008) Genome-wide pattern of TCF7L2/TCF4 chromatin occupancy in colorectal cancer cells. *Mol. Cell. Biol.*, **28**, 2732–2744.
24. Poulsom, R., Longcroft, J.M., Jeffery, R.E., Rogers, L.A. and Steel, J.H. (1998) A robust method for isotopic riboprobe in situ hybridisation to localise mRNAs in routine pathology specimens. *Eur. J. Histochem.*, **42**, 121–132.