



---

Year: 2014

---

## A comprehensive look at transcription factor gene expression changes in colorectal adenomas

Vonlanthen, Janine; Okoniewski, Michal J; Menigatti, Mirco; Cattaneo, Elisa; Pellegrini-Ochsner, Daniela; Haider, Ritva; Jiricny, Josef; Staiano, Teresa; Buffoli, Federico; Marra, Giancarlo

**Abstract:** **BACKGROUND:** Biological processes are controlled by transcription networks. Expression changes of transcription factor (TF) genes in precancerous lesions are therefore crucial events in tumorigenesis. Our aim was to obtain a comprehensive picture of these changes in colorectal adenomas. **METHODS:** Using a 3-pronged selection procedure, we analyzed transcriptomic data on 34 human tissue samples (17 adenomas and paired samples of normal mucosa, all collected with ethics committee approval and written, informed patient consent) to identify TFs with highly significant tumor-associated gene expression changes whose potential roles in colorectal tumorigenesis have been under-researched. Microarray data were subjected to stringent statistical analysis of TF expression in tumor vs. normal tissues, MetaCore-mediated identification of TF networks displaying enrichment for genes that were differentially expressed in tumors, and a novel quantitative analysis of the publications examining the TF genes' roles in colorectal tumorigenesis. **RESULTS:** The 261 TF genes identified with this procedure included DACH1, which plays essential roles in the proper proliferation and differentiation of retinal and leg precursor cell populations in *Drosophila melanogaster*. Its possible roles in colorectal tumorigenesis are completely unknown, but it was found to be markedly overexpressed (mRNA and protein) in all colorectal adenomas and in most colorectal carcinomas. However, DACH1 expression was absent in some carcinomas, most of which were DNA mismatch-repair deficient. When networks were built using the set of TF genes identified by all three selection procedures, as well as the entire set of transcriptomic changes in adenomas, five hub genes (TGFB1, BIRC5, MYB, NR3C1, and TERT) were identified as putatively crucial components of the adenomatous transformation process. **CONCLUSION:** The transcription-regulating network of colorectal adenomas (compared with that of normal colorectal mucosa) is characterized by significantly altered expression of over 250 TF genes, many of which have never been investigated in relation to colorectal tumorigenesis.

DOI: 10.1186/1471-2407-14-46

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <http://doi.org/10.5167/uzh-92989>

Published Version



Originally published at:

Vonlanthen, Janine; Okoniewski, Michal J; Menigatti, Mirco; Cattaneo, Elisa; Pellegrini-Ochsner, Daniela; Haider, Ritva; Jiricny, Josef; Staiano, Teresa; Buffoli, Federico; Marra, Giancarlo (2014). A comprehen-

sive look at transcription factor gene expression changes in colorectal adenomas. BMC Cancer, 14:46.  
DOI: 10.1186/1471-2407-14-46

RESEARCH ARTICLE

Open Access

# A comprehensive look at transcription factor gene expression changes in colorectal adenomas

Janine Vonlanthen<sup>1</sup>, Michal J Okoniewski<sup>2</sup>, Mirco Menigatti<sup>1</sup>, Elisa Cattaneo<sup>1</sup>, Daniela Pellegrini-Ochsner<sup>3</sup>, Ritva Haider<sup>1</sup>, Josef Jiricny<sup>1</sup>, Teresa Staiano<sup>4</sup>, Federico Buffoli<sup>4</sup> and Giancarlo Marra<sup>1\*</sup>

## Abstract

**Background:** Biological processes are controlled by transcription networks. Expression changes of transcription factor (TF) genes in precancerous lesions are therefore crucial events in tumorigenesis. Our aim was to obtain a comprehensive picture of these changes in colorectal adenomas.

**Methods:** Using a 3-pronged selection procedure, we analyzed transcriptomic data on 34 human tissue samples (17 adenomas and paired samples of normal mucosa, all collected with ethics committee approval and written, informed patient consent) to identify TFs with highly significant tumor-associated gene expression changes whose potential roles in colorectal tumorigenesis have been under-researched. Microarray data were subjected to stringent statistical analysis of TF expression in tumor vs. normal tissues, MetaCore-mediated identification of TF networks displaying enrichment for genes that were differentially expressed in tumors, and a novel quantitative analysis of the publications examining the TF genes' roles in colorectal tumorigenesis.

**Results:** The 261 TF genes identified with this procedure included *DACH1*, which plays essential roles in the proper proliferation and differentiation of retinal and leg precursor cell populations in *Drosophila melanogaster*. Its possible roles in colorectal tumorigenesis are completely unknown, but it was found to be markedly overexpressed (mRNA and protein) in all colorectal adenomas and in most colorectal carcinomas. However, *DACH1* expression was absent in some carcinomas, most of which were DNA mismatch-repair deficient. When networks were built using the set of TF genes identified by all three selection procedures, as well as the entire set of transcriptomic changes in adenomas, five hub genes (*TGFB1*, *BIRC5*, *MYB*, *NR3C1*, and *TERT*) were identified as putatively crucial components of the adenomatous transformation process.

**Conclusion:** The transcription-regulating network of colorectal adenomas (compared with that of normal colorectal mucosa) is characterized by significantly altered expression of over 250 TF genes, many of which have never been investigated in relation to colorectal tumorigenesis.

**Keywords:** Transcription factors, Gene expression, Colorectal adenomas, *DACH1*

## Background

Colorectal adenomas are benign tumors of the large intestinal epithelium. They are found in roughly one third of asymptomatic adults who undergo colonoscopy before the age of 50. Endoscopic removal of these lesions is associated with high rates of recurrence (up to 60% at three years, depending on the size, number, histological features, and degree of dysplasia [1]). In addition, it has

been estimated that 15% of adenomas measuring 1 cm or more become carcinomas within 10 years of their detection [2].

Adenomatous transformation of normal colorectal mucosa is associated with profound changes in the tissue's gene expression profile [3]. These changes are caused by epigenetic and/or genetic events that "reprogram" the regulation of gene transcription [4]. An early—and probably fundamental—event in this reprogramming process involves qualitative, quantitative, and spatial subversion of the Wnt signaling pathway, the physiological regulator of epithelial homeostasis [5]. Almost invariably, it stems from

\* Correspondence: marra@imcr.uzh.ch

<sup>1</sup>Institute of Molecular Cancer Research, University of Zurich, Winterthurerstrasse 190, Zurich 8051, Switzerland

Full list of author information is available at the end of the article

mutations in genes encoding Wnt pathway components (*APC*, *adenomatous polyposis coli*, in most cases), which lead to the accumulation of  $\beta$ -catenin in both the cytoplasm and nucleus. In the latter compartment, it interacts with DNA-binding proteins of the T-cell factor/lymphoid-enhancer factor family, transforming them from transcriptional repressors into transcriptional activators.

The abnormal activation of Wnt signaling can affect the expression of numerous genes involved in epithelial homeostasis, including the oncogenic transcription factor (TF)-encoding gene *MYC*. It is one of the genes most frequently found to be overexpressed in intestinal adenomas and carcinomas (and many other tumors as well) [6,7]. Genes directly targeted by *MYC* have been identified in various tumors [8,9], but more recent studies suggest that this oncogene might be a “universal amplifier” with effects on most of the cell’s actively expressed genes. This phenomenon might account for the broad spectrum of effects ascribed to this oncogene in normal and tumor cells [10,11].

However, while *MYC* undoubtedly plays a central role in tumors that overexpress it, the adenomatous phenotype is likely to be underpinned by transcription networks in which the expression of numerous TFs is altered. These networks are characterized by cross-regulation and redundant regulation of component TFs and TF-gene binding that occurs over a wide range of DNA occupancy levels [12]. Understanding how the concentration of a given TF in a neoplastic tissue differs from that in its normal tissue counterpart is therefore of paramount importance to elucidate the tumorigenic process.

Gene expression studies can reveal potentially important factors in colorectal tumorigenesis by pinpointing genes with markedly up- or downregulated expression levels in early precancerous lesions [3,13,14]. For this reason, we attempted in the present study to comprehensively characterize the TF gene expression changes that occur in colorectal adenomas. Many of the numerous changes we identified involve TF genes that have not been previously linked to colorectal tumorigenesis. One of these, *DACH1*, consistently displayed marked upregulation in the colorectal adenomas we examined, and it was subjected to further investigation in a series of neoplasms representing different types and stages of colorectal tumor progression.

## Methods

### Microarray data

We analyzed previously collected [13] gene expression data on 17 pedunculated colorectal adenomas and 17 peritumoral samples of normal mucosa (> 2 cm from the adenoma). The pathologic features of the tumor series are summarized in Additional file 1: Table S1. Human colorectal tissues were prospectively collected

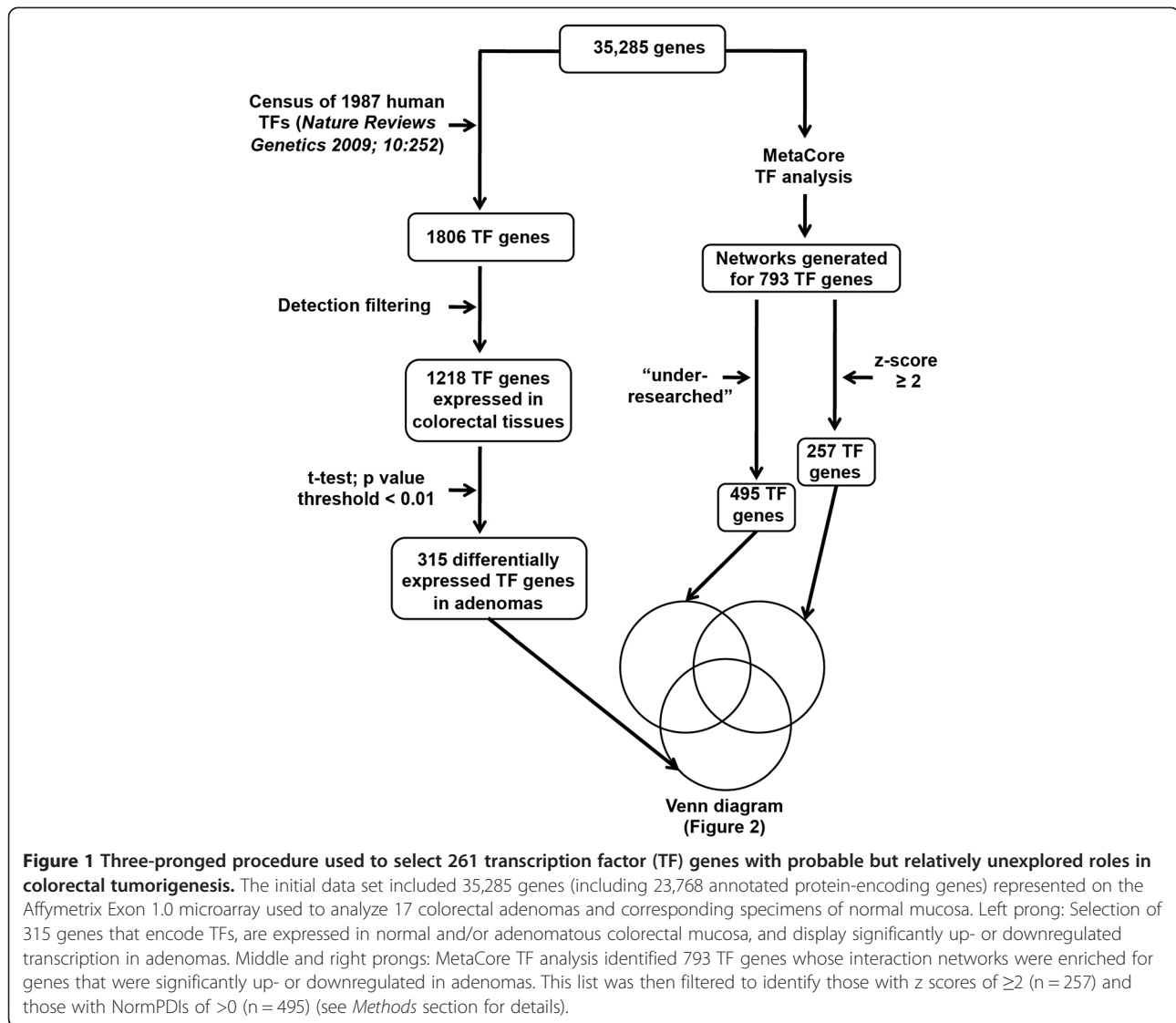
from patients undergoing colonoscopy in the *Istituti Ospitalieri* of Cremona, Italy. The approval of the ethics committee of this institution was obtained, and tissues were used in accordance with the Declaration of Helsinki. Each donor provided written informed consent to sample collection, data analysis, and publication of the findings. Detailed descriptions of RNA extraction method and the Affymetrix Exon 1.0 microarray analysis are available in the report of our original study [13]. Raw transcriptomic data have been deposited in GEO (accession number GSE21962).

### Selection of TF genes

A three-pronged selection procedure (Figure 1) was used to identify TFs likely to play important but unsuspected roles in colorectal tumorigenesis. The starting point was a list of 35,285 genes, i.e., the 23,768 protein-encoding genes examined in the original study [13] plus 11,517 non-protein-encoding genes.

First (Figure 1, left prong), these genes were screened against a census of human TFs published in 2009 by Vaquerizas *et al.* [15]. This manually curated compilation contains 1987 sequence-specific DNA-binding TF genes, each with information on its function, genomic organization, and evolutionary conservation. Most were identified with the Ensemble Genome Browser [16], but 27 are probable TF genes from other sources, such as Gene Ontology [GO] or TRANSCRIPTION FACTOR [TRANSFAC] database [17]. One thousand eight hundred six of the 1987 TF genes in the census were also found in our original data set. These genes were selected on the basis of gene-level Brainarray summaries [18] of the Exon 1.0 microarray data, so exon-level and splicing information were not taken into account. A detection filter was then applied to select TF genes likely to be expressed in either normal or adenomatous colorectal tissues. Candidates were thus excluded unless their expression values exceeded an arbitrarily defined cut-off of 5.8 (log<sub>2</sub> scale) in  $\geq 50\%$  of the samples in one or both of the tissue groups (adenomas, normal mucosa). The 1218 TF genes selected with this step are listed in Additional file 2: Table S2. This list was then further reduced to include only those TF genes that had exhibited significantly up- or downregulated expression in the adenomas vs. normal mucosa (TF genes in bold face in Additional file 2: Table S2). For this final selection, a p value threshold of < 0.01 in a paired two-tailed t test was chosen. Unadjusted p values were used for the ranking, which is not influenced by multiple testing correction [19].

The second and third prongs of the selection procedure (Figure 1, middle and right-hand columns) began with analysis of TF genes in the original data set with commercially available MetaCore™ software (version 6.14, build 61508) from GeneGo, Inc. In MetaCore, each gene is assigned to a network of related genes (e.g., a TF gene is included in a network of genes that it likely



regulates). Network size varies widely: some contain less than 10 genes, others (like that of the transcription factor SP1), well over 2000. The MetaCore TF analysis used the hypergeometric test to select TF genes regulating networks enriched in genes that had displayed significant differential expression in our adenomas, as compared with normal mucosa. The results are expressed in terms of a z-score, which reflects the deviation stretch from the mean of a normally distributed population, and a p value, which is inversely correlated with the significance of the TF network (Additional file 3: Table S3). We set a relaxed significance threshold (a t-test p value of 0.2 and an absolute logarithmic fold change of 0.2) to select TF networks with enough significant elements to allow efficient calculation of enrichment. The significance of a given TF gene network in the context of the selected genes, measured by hypergeometric test, is described by its p value and additionally by the z-score of

network enrichment. The 793 TF genes whose networks were enriched in genes displaying significant differential expression in adenomas (Figure 1) are listed in Additional file 4: Table S4, where those with z-scores  $> 2$  are reported in bold-face type.

MetaCore is based on a curated database of human protein-protein and protein-DNA interactions, transcription factors, signaling and metabolic pathways, diseases and toxicity, and the effects of bioactive molecules. It is constructed and edited manually by GeneGo scientists on the basis of data from full-text articles published in relevant journals (<https://portal.genego.com>). The size of a gene network therefore depends on the data (and therefore the number of publications) available on a given gene. In GeneGo, TF significance (measured by the parameters described above) is related to network size. Therefore, genes that have been researched more intensively and are therefore well-represented in published reports might be

reported as more significant than those that have been less thoroughly investigated. In other words, higher connectivity might be partly rooted in investigative biases.

The third prong of our selection procedure (Figure 1) was designed to correct for such biases by identifying TFs that are under-represented in scientific publications dealing with colorectal tumors. For each TF gene identified by the Metacore analysis, we manually reviewed the GeneCard ([www.genecards.org](http://www.genecards.org)) links to research articles dealing with the gene indexed in PubMed (as well as Novoseek, HGNC, Entrez Gene, UniProB, PharmaGKB and/or GAD) and recorded the number of articles that also dealt with colorectal tumors (*actual* publications). Correlation between the number of *actual* publications and the z-score of each TF gene was assessed with a scatter plot, and a trend line was drawn to identify the *expected* number of publications for each TF (Additional file 5: Figure S1). The trend line was obtained by multiplying the z-score for each TF by the slope value (142 in this case, with the fixed intercept = 0). The correlation was fairly strong ( $r=0.4$ ) for such heterogeneous data, so the linear approximation appeared to be justified. By subtracting the *actual* number from the *expected* number of publications calculated for each TF, the *difference in publications* (DP) was obtained. The normalized DP (NormDP) was then calculated [i.e.,  $\text{NormDP} = (\text{actual} - \text{expected publication number}) / \text{expected publication number}$ ], which correlates with the distance to the trend line. Higher NormDPs reflect larger discrepancies between the expected and actual numbers of publications and are therefore associated with TFs whose possible links to colorectal tumorigenesis have been relatively "under-researched." The TF genes with a NormDP > 0 were therefore termed "under-researched" (the 495 TF genes in red colour in Additional file 4: Table S4). Their importance and number of connections in the Metacore database may be underestimated owing to their limited presence in the literature.

The TF gene sets generated by the three selection procedures were compared and their intersections represented in a Venn diagram (see *Results* and *Discussion* sections). Hierarchical clustering analysis of the microarray data was carried out using heatmap.2 function from the gplots library (CRAN repository at <http://cran.rproject.org/web/packages/gplots/index.html>) with Pearson correlation as a distance function and Ward agglomeration method for clustering.

The TF gene expression perturbations found in our adenoma series were also compared with those reported in advanced colorectal tumors. For this purpose, we applied the same TF gene selection procedure to the Exon 1.0 microarray-based, gene expression data reported by Maglietta et al. [14] (raw data available in Array Express E-MTAB-829) relative to 13 colorectal *carcinomas* and paired samples of normal mucosa.

### Immunohistochemistry

Immunostaining was used to assess DACH1 protein expression patterns in 20 archival, formalin-fixed, paraffin-embedded colorectal adenomas, 80 sporadic colorectal cancers, and the normal mucosa adjacent to these latter lesions. The cancers represented different stages and histologic grades (Additional file 6: Table S5). Forty were classified as mismatch repair (MMR)-proficient and 40 as MMR-deficient based on immunostaining for the protein encoded by the MMR gene *MLH1*, whose lack of expression in sporadic cancer is caused by CpG island hypermethylation at its promoter [20]. *MLH1* protein expression in a cancer tissue is usually uniformly strong (indicating MMR proficiency) or completely absent (MMR deficiency) [20].

In brief, 4- $\mu\text{m}$  sections of each cancer were mounted on glass slides coated with organosilane (DakoCytomation), deparaffinized, and rehydrated. Antigen retrieval was accomplished by heating the sections in a pressure cooker at 120°C for 2 min in 10 mM citrate-buffered solution (pH 6.0). DAKO peroxidase-blocking reagent and goat serum were used sequentially to suppress nonspecific staining due to endogenous peroxidase activity and nonspecific antibody binding, respectively. Sections were then incubated overnight at 4°C with the primary antibody (mouse monoclonal anti-*MLH1* antibody [BD, no. 551091, 1:200 dilution] or rabbit polyclonal anti-DACH1 antibody [Sigma, no. HPA012672, 1:400 dilution]). The sections were washed, and appropriate secondary antibodies conjugated to peroxidase-labeled polymer (DAKO EnVision+ kit) were applied for 30 min at RT. Finally, the sections were incubated with 3,3'-diaminobenzidine chromogen solution (DAKO) to develop the peroxidase activity and then counterstained with hematoxylin.

DACH1 immunostaining patterns proved to be complex and were evaluated as follows. The extension of staining in each cancer specimen (i.e., the percentage of tumor cells displaying any degree of staining) was rated as absent (no stained cells); limited ( $\leq 35\%$  cells); moderate (36%–69%); or extensive (70%–100%). As for immunostaining intensity, scores were first assigned to various areas of the cancer (1 = weak; 2 = moderate; 3 = strong). The highest score assigned anywhere in the cancer specimen was then added to the score that was most representative of the specimen. The sum was an intensity score ranging from 2 to 6. The Fisher exact test was used to examine the significance of associations between extension or intensity DACH1 staining score and various characteristics of the cancers (MMR status, TNM stage, and histologic grade).

The specificity of the DACH1 antibody we used was verified in immunostaining experiments performed as described above on sections of formalin-fixed, paraffin-

embedded pellets made from colon cancer cell lines with different *DACH1* gene expression levels.

### Evaluation of *DACH1* promoter methylation status in colorectal cancers

Using the QIAamp DNA FFPE Tissue kit (Qiagen, no. 56404), we extracted DNA from 18 of the 80 cancers described above. *DACH1* expression in these cancers was marked and ubiquitous in 6, patchy in 6, and completely lost in 6 (see examples in the *Results* section), and each of these 3 groups included 3 tumors that were MMR-proficient and 3 that were MMR-deficient. Sodium bisulfite conversion of genomic DNA was performed as previously described [21], and the resulting DNA was subjected to combined bisulfite restriction analysis (COBRA) to determine the methylation status of two CpG islands located respectively upstream the transcription start site (CpG I) and in the first intron (CpG II) of the *DACH1* gene. Amplifications were carried out with FastStart Taq DNA Polymerase (Roche, Basel, Switzerland) with the following primers: CpG I: 5'-GTAGTAGTAGAAGAGAAGTAGATGA-3' (sense) and 5'-ACCCAAATTATCCAACCAAAA ACTC-3' (antisense); CpG II: 5'-GGGTGAGGGTTTIGT TGGGA-3' (sense) and 5'-CCCTCCCCTCIACACT TC-3' (antisense). The amplified products were digested with the *Taq<sup>61</sup>* restriction enzyme (New England Biolabs, Beverly, MA, USA) and subjected to 2% agarose gel electrophoresis and ethidium bromide staining.

### Results

To isolate *bona fide* TFs from our original set of 35,285 genes, we screened it against the census of 1987 human TFs compiled by Vaquerizas *et al.* [15]. As shown in Figure 1 (left-hand prong), 1806 of the 1987 TF genes were identified among those in our original set, but only 1218 of these were significantly expressed in either normal colorectal mucosa or in colorectal adenomas or in both (see *Methods*). The expression levels of these 1218 TF genes in the normal and neoplastic tissue groups are illustrated in a hierarchical clustering analysis of the 34 tissue samples (Additional file 7: Figure S2). As shown in Figure 1 (and detailed in Additional file 2: Table S2), 315 of the 1218 TF genes were found to be significantly over- or under-expressed in adenomas relative to normal mucosa (*t* test:  $p < 0.01$ ).

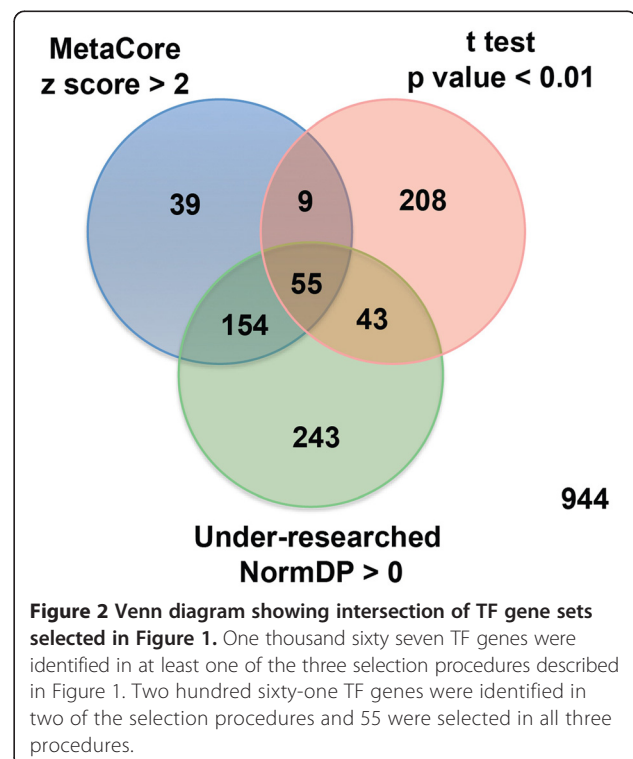
Parallel MetaCore analysis of the original gene set identified 793 TF genes whose interaction networks were enriched for genes displaying significant differential expression in adenomas, as compared with normal mucosa samples (Additional file 4: Table S4). This list, which was generated with the relatively relaxed criteria described in the *Methods* section, was then filtered (Figure 1, right-hand prong) to select the TF genes most likely to be involved in adenomatous transformation of the colorectal

epithelium. The result was a list of 257 TF genes with *z*-scores  $\geq 2$  in the hypergeometric enrichment test, reflecting gene expression changes in adenomas amounting to at least 2 standard deviations from the mean expression change.

In parallel, the MetaCore list of 793 TF genes was filtered to identify those whose possible role in colorectal tumorigenesis has been relatively under-researched (Figure 1, middle prong), as defined by the NormDP (see *Methods*). This analysis pinpointed 495 of the 793 TF genes with fewer than expected publications on their involvement in colorectal tumorigenesis (i.e., NormDPs of  $>0$ ; Additional file 4: Table S4).

Figure 2 shows the intersections of the three TF gene sets obtained with the procedures described above. Two hundred sixty one were identified with at least two selection procedures (Additional file 8: Table S8). Hierarchical clustering analysis of the 34 tissue samples based on the expression levels of these TF genes showed clear separation of the adenomas and normal mucosa samples (Figure 3). The sub-clusters of adenomas and normal samples seen in Figure 3 showed no correlation with the known clinical and pathologic features of the tissues (Additional file 1: Table S1), which is not particularly surprising given the relatively small number of samples analyzed.

We then used two different approaches to identify TF genes listed in Additional file 8: Table S8 that might be candidates for subsequent validation studies as drivers of colorectal transformation. First, using manual inspection

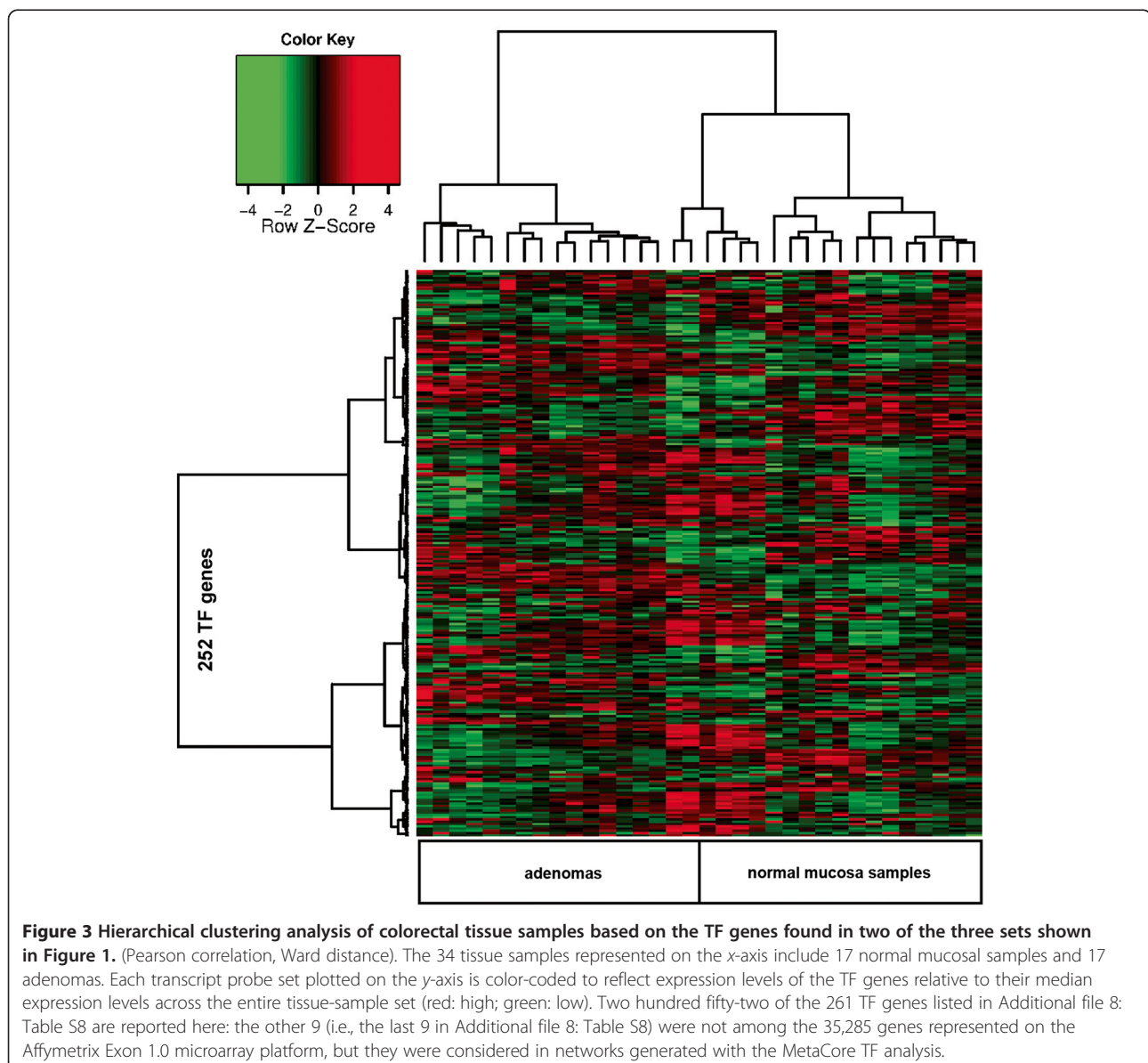


of the list, we selected the TF genes with the following characteristics: marked upregulation in adenomas (i.e., top upregulated genes in Additional file 8: Table S8) and *no* actual publications on the possible roles in colorectal tumorigenesis (regardless of whether research had been published on their involvement in other types of tumorigenesis). Upregulated TF genes were chosen since they were also more likely to represent potential biomarkers of adenomatous transformation.

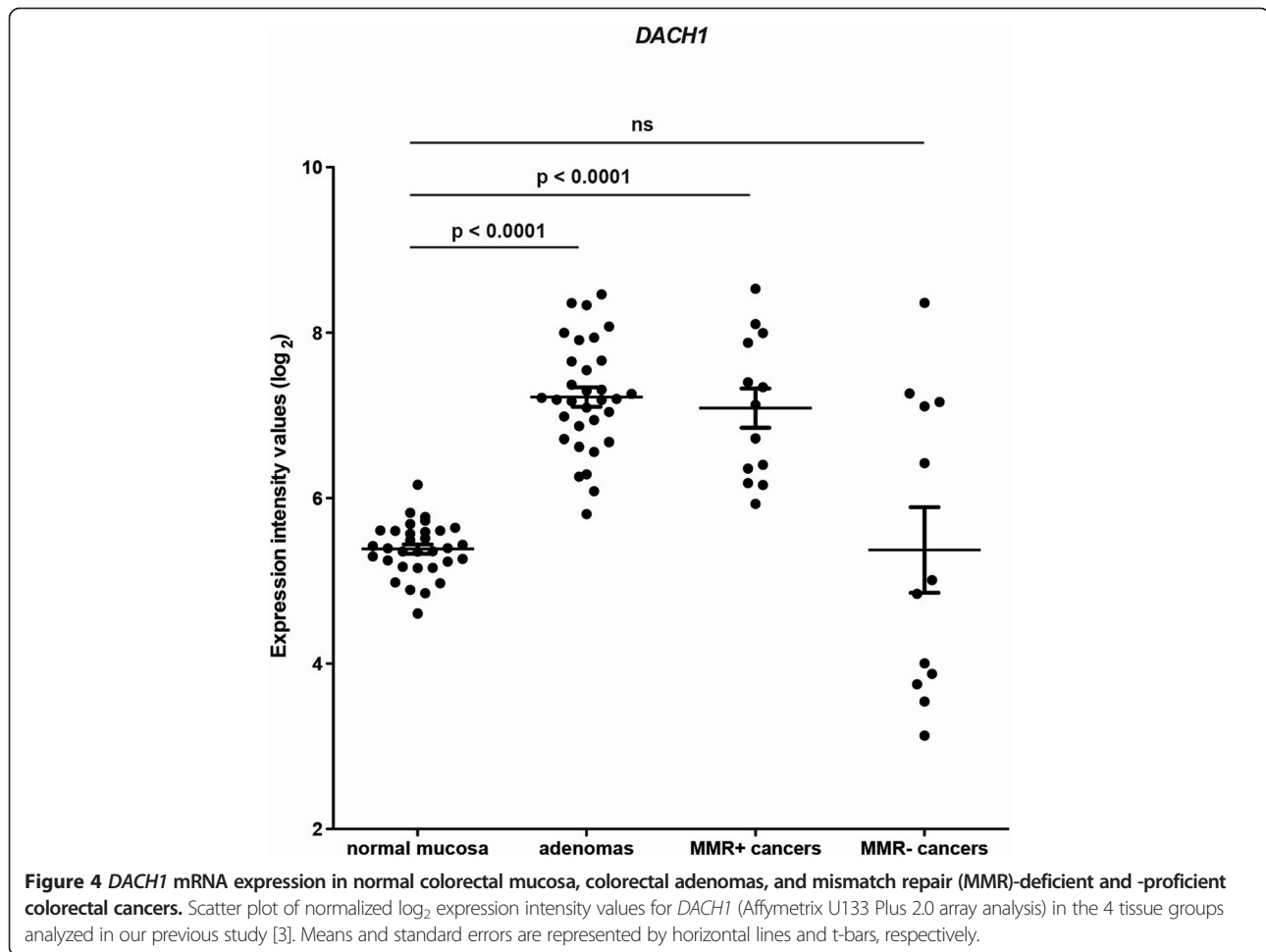
One of the genes that met these criteria was *DACH1*. Microarray data from a previous study by our group [3] had indicated that its expression is also upregulated in most colorectal cancers, although significantly reduced mRNA levels were observed in some of the cancers tested, all of which were MMR-deficient (Figure 4). This

observation prompted us to conduct immunohistochemistry experiments to investigate *DACH1* protein expression in colorectal adenomas and in colorectal cancers of different stages, histologic grades, and MMR status (40 MMR+ and 40 MMR-, Additional file 6: Table S5).

The *DACH1* antibody used for these studies displayed excellent specificity, as shown by Additional file 9: Figure S3. Immunostaining of normal mucosa revealed high nuclear expression of *DACH1*, which was confined mainly to the proliferating cells in the lower half of colorectal crypts (Figure 5A). Nuclear expression was also invariably strong in the adenomas we tested, but in this case it was almost ubiquitous (Figure 5B and C). As for the cancers, three different staining patterns emerged: marked and ubiquitous *DACH1* expression resembling





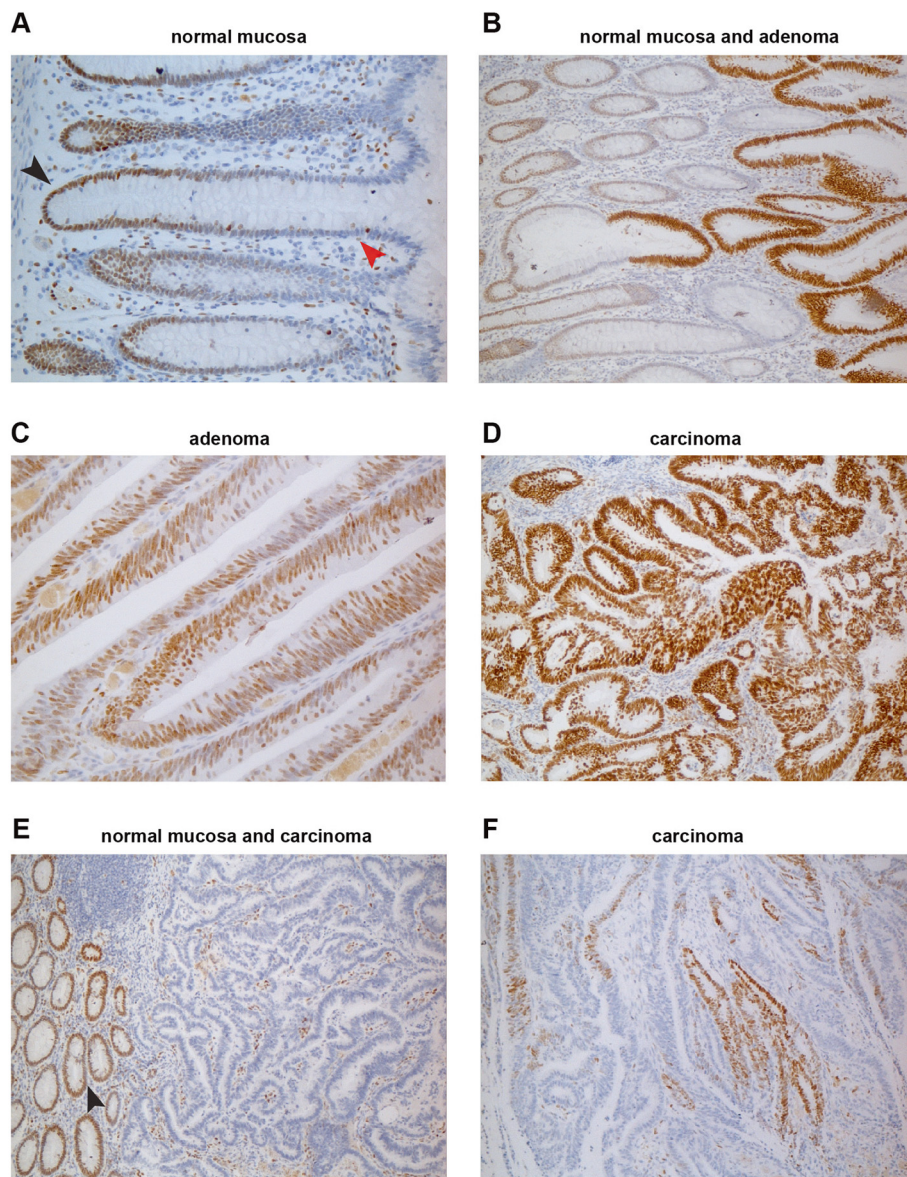


that seen in adenomas (Figure 5D); complete loss of expression throughout the lesion (Figure 5E); and patches of variable-intensity staining interspersed with areas of absent expression (Figure 5F). The latter two patterns were significantly more frequent in MMR- cancers (30/40 vs. 11/40 of those that were MMR+). Fisher's exact tests showed that *DACH1* expression in MMR- cancers was significantly more likely to be partially/completely lost (staining extension: <70% of cells;  $p = 0.00016$ ) or relatively weak (intensity scores of <5) ( $p = 0.054$ ) than that observed in MMR+ cancers. *DACH1* staining intensity scores were also significantly lower in poorly differentiated (G3) cancers ( $p = 0.019$  vs. G2 cancers), which were (as expected [20]) significantly associated with MMR deficiency ( $P = 0.0019$ ). *DACH1* staining patterns did not appear to be related to TNM stages, although this finding needs to be confirmed in larger groups of MMR+ and MMR- cancers.

Because our MMR- cancers showed loss of gene expression due to epigenetic silencing of the MMR gene *MLH1*, we wondered whether their diminished *DACH1* expression might be caused by methylation at the *DACH1*

promoter. The COBRA experiments we performed failed to confirm this hypothesis. The CpG island located in the *DACH1* promoter (CpG I in Figure 6A, primers in *Methods*) was not found to be methylated in any of the 18 cancers we tested (samples from each *DACH1* staining pattern group are shown in Figure 6B). Hypermethylation at this site may occur *in vitro*, however, as shown for the colon cancer cell lines HCT116 and CO115 (Figure 6B). Similar results were obtained with the COBRA analysis of a different CpG island located in the first intron of the *DACH1* gene (CpG II in Figure 6A).

The second approach we used involved the identification of genes that might represent important hubs in the transcriptional network of adenomas (as opposed to the one operating in the normal mucosa). To this end, we uploaded the 55 significant TF genes identified by all three selection procedures (Figure 2) into the MetaCore database and ran a comparative analysis of their networks. The most promising network included the following five target genes: *TGF-beta 1* (*TGFB1*), *TERT*, *Survivin* (*BIRC5*), *c-Myb* (*MYB*), and *GCR-alpha* (*NR3C1*) (see Figure 7, and Additional file 10: Figure S4 and Additional file 11: Figure

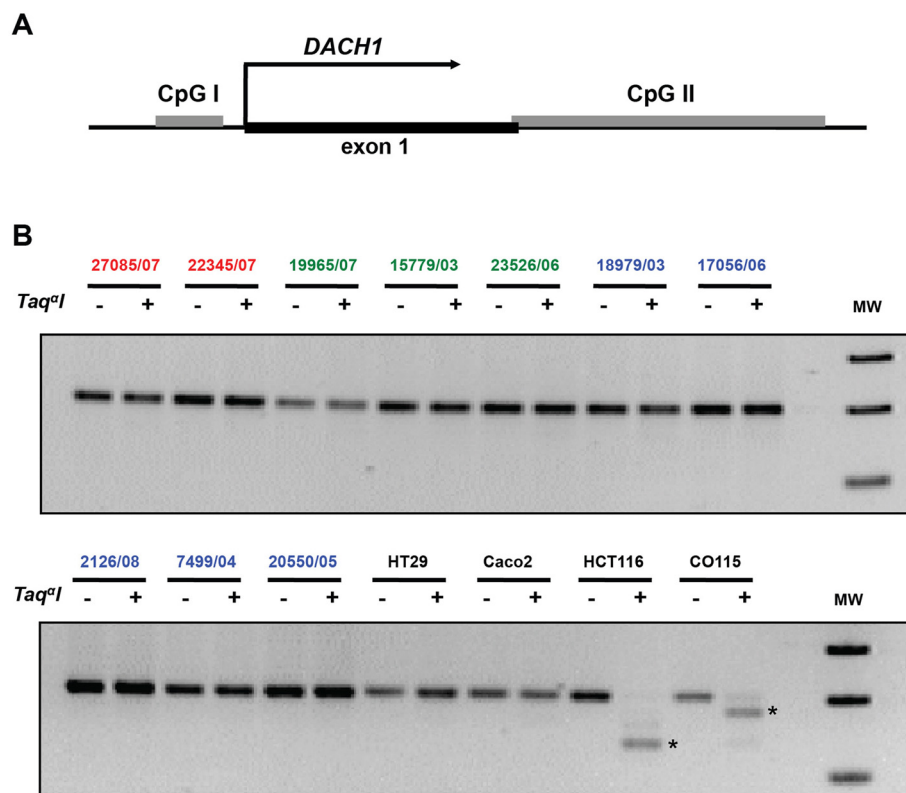


**Figure 5 Immunohistochemical staining for DACH1 protein in normal and neoplastic colon. (A)** In normal mucosa, DACH1 expression is present in the nuclei of proliferating cells in the lower portion of the epithelial crypts (black arrowhead) and completely absent in the differentiated cells in the upper crypts (red arrowhead). **(B)** High-level DACH1 expression is seen in rapidly proliferating cells of adenomatous glands taking over normal crypts. Abundant expression is also seen in most cells of a colorectal adenoma **(C)** and a colorectal carcinoma **(D)**. In another colorectal cancer **(E)**, DACH1 expression is absent in neoplastic glands, although proliferating cells in the normal mucosa and in the tumoral stroma are positive. **(F)** A third colorectal cancer with patchy staining for DACH1.

S5). This network was characterized by a p value of  $3.43e-64$  and 75 target genes, including 27 “seeds”, i.e., TF genes. These findings will be discussed in the next section.

Finally, we compared the perturbations of TF gene expression documented in our colorectal adenomas with those reported by Maglietta et al. [14] in 13 colorectal *carcinomas* and paired normal mucosa samples. These latter tissue pairs were a subset of the 17 analyzed by Maglietta et al. They were selected because they had all

been processed in the same laboratory [14]. As shown in Additional file 12: Figure S6), a substantial proportion of TF genes whose expression was dysregulated in the carcinomas were also dysregulated in our adenomas (46% using the *t* test based-approach of the left prong of our selection procedure, 57% using the MetaCore-based approach of the right prong [Figure 1]). The TF genes identified in colorectal carcinomas with these two approaches are reported in Additional file 13: Table S6 and Additional file 14: Table S7).



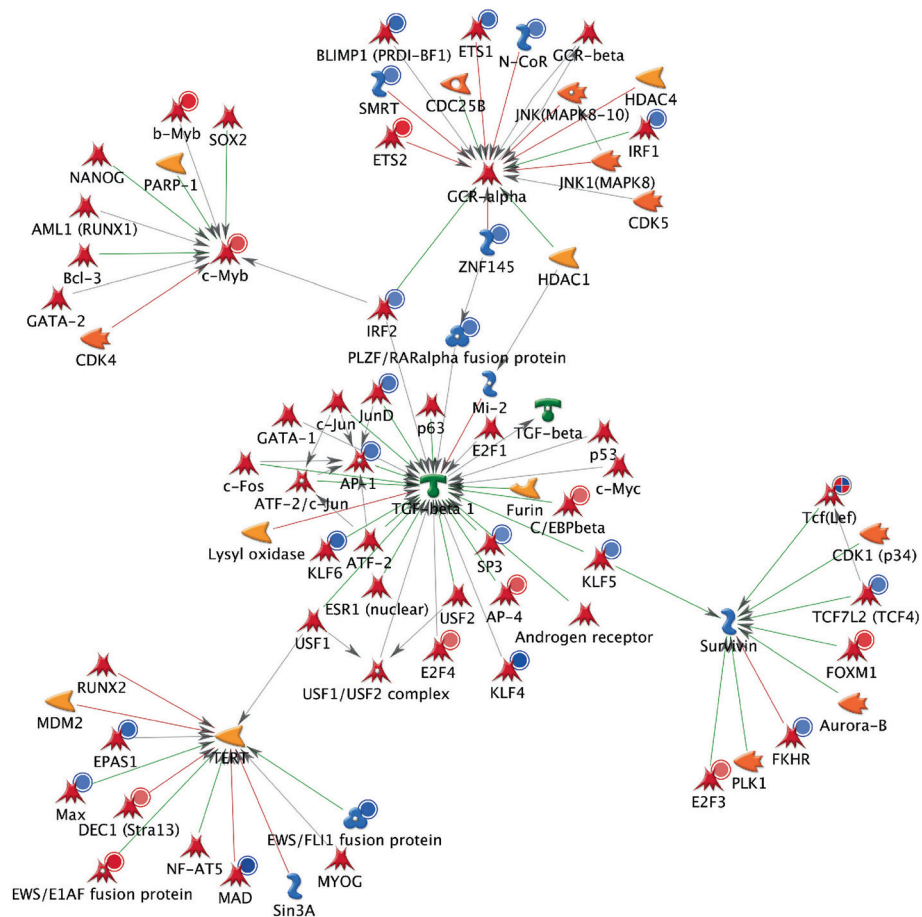
**Figure 6** Methylation analysis of the CpG island in the *DACH1* promoter. **(A)**: Schematic depiction of the CpG islands located respectively 5' upstream from the *DACH1* transcription start site (CpG I) and in the first intron of the *DACH1* gene (CpG II). **(B)**: Examples of CpG I COBRA analysis in colorectal cancers with intense (red), patchy (green), or no (blue) *DACH1* protein immunostaining and in 4 colon cancer cell lines characterized by low (HT29 and Caco2) or very low (HCT116 and CO115) *DACH1* expression (based on microarray-documented *DACH1* mRNA expression levels - see also Additional file 9). Asterisks indicate *Taq*<sup>I</sup>-digested DNA fragments representing methylated alleles; slower-migrating fragments correspond to undigested, unmethylated DNA. MW, molecular weight; bp, base pair.

## Discussion

The aim of this study was to identify TF genes with probable roles in the early stages of colorectal tumorigenesis, especially those whose roles in this setting have been less extensively investigated. The list we compiled contained 261 TF genes, including one, *DACH1*, which appeared particularly interesting. It was invariably overexpressed in the preinvasive stage of colorectal tumorigenesis (i.e., adenomas) and frequently upregulated in colorectal cancers as well. However, it was found to be silenced in certain colorectal cancers, especially those that were MMR-.

To our knowledge, this is the first attempt to comprehensively characterize the TF gene transcriptome of human colorectal adenomatous polyps, although several studies have been published on the overall transcriptional profile of colorectal tumors (GEO database [22] and our previous reports [3,13,23]). Vaquerizas *et al.* studied TF gene expression changes in 32 healthy human tissue types, but, surprisingly, the colon was not included.

The focus of our study was the adenomatous colorectal polyp (as compared with corresponding samples of normal colorectal mucosa), and genes were considered to be expressed in these tissues only if they had expression levels of  $\geq 5.8$  (log scale). This cutoff, which was chosen on the basis of our previous observations and the recommendations of the microarray manufacturer (Affymetrix), is less stringent than the one used by Vaquerizas *et al.* [15]. We deliberately chose a more relaxed cutoff to maximize our chances of identifying all TFs involved in colorectal carcinogenesis, even those with low-level expression. This is important because TF mRNA and TF proteins are less stable than those of other classes of genes [24], and TF protein levels span over four orders of magnitude [12]. The number of TF genes that met our criterion for expression in normal or adenomatous colorectal tissue (or both)—1218—was thus markedly higher than those reported by Vaquerizas *et al.* in normal tissues of other organ systems, which ranged from 150 to 300 [15]. The U133 Plus 2.0 array used by these investigators is also less sensitive than the



**Figure 7 Organic (hub-centric) layout of the most significant network identified by MetaCore.** The network includes 27 of the 55 TF genes found in all three sets depicted in Figure 2.

Affymetrix exon array platform we used [13]. In spite of these differences, however, in both studies, TF genes represented ~7% of all genes classified as “expressed” genes in most of the tissues examined.

The three-pronged selection procedure we used to identify TF genes involved in colorectal tumorigenesis generated a final list of 261 candidates (Additional file 8: Table S8). At the time of our analysis, only 15% of these genes had been implicated (putatively or otherwise) in colorectal tumorigenesis in more than 10 publications, including a few like *MYC* and *TP53*, whose links to this process are well-established. And for 102 (39%) of the candidate genes, our literature search revealed no data at all on their possible roles in colorectal tumors.

To extract meaningful biological information from this list, we initially focused on the TF genes displaying the most markedly *upregulated* expression in colorectal adenomas together with the lowest publication scores. One of the top genes in this subgroup was *DACH1*, a human homolog of the *Drosophila melanogaster* TF gene *dachs-hund*, which is essential for proper proliferation and

differentiation of retinal and leg precursor cell populations in these flies [25-27]. *DACH1* appears to regulate the transcription of several human genes involved in proliferation (e.g., *CDKN1B*, *CCND1*, *JUN*, and *TGFb*) [28-32]. Furthermore, our immunohistochemistry studies revealed abundant *DACH1* expression in the nuclei of epithelial cells in the lower half of normal colorectal crypts (Figure 5A), where proliferation predominates over differentiation. Therefore, the staining pattern strongly associates *DACH1* expression with cell proliferation and/or commitment to cell differentiation. It has also recently found to be highly expressed in cycling intestinal stem cells from mice [33].

In line with these findings, the expression of *DACH1* mRNA and protein was significantly increased in tumor lesions (Figures 4 and 5C/D), which are extensively populated by proliferating cells. However, it does not appear to be indispensable for cancer-cell proliferation and cancer progression since some of the colorectal cancers we examined were characterized by complete or partial loss of *DACH1* protein expression (Figure 5E and F). These

losses showed no correlation with the TNM stages of the cancers, but they were significantly more frequent in tumors that were poorly differentiated and/or MMR deficient.

The mismatch repair defect was the result of epigenetic silencing of the *MLH1* gene, and a similar phenomenon might have been responsible for the loss of *DACH1* expression in some cancers. However, COBRA revealed no evidence of cytosine hypermethylation at the CpG islands investigated (Figure 6A) in any of the colorectal cancers we examined (although analysis of colon cancer cell lines indicated that this phenomenon can occur *in vitro*) (Figure 6B). The loss of *DACH1* in certain cancers might stem from cytosine hypermethylation at other possible regulatory regions of the *DACH1* locus or from other types of epigenetic changes at this site. Additional work is needed to explore these possibilities. In any case, histone modifications are likely to play some role in the silencing of this gene. *DACH1* is one of the developmental TF genes whose chromatin in mouse embryonic stem cells is bivalent, i.e., it harbors permissive as well as repressive histone marks (Figure 2 in [34]). This epigenetic conformation, which facilitates the gene's ability to switch rapidly between transcriptionally active and inactive states, might account for the staining patterns shown in Figure 5.

Altered *DACH1* expression has already been reported in other human tumors. In a study of lung cancers based on whole-genome sequencing, *DACH1* emerged as a biologically significant target of mutation (loss-of-function alterations in particular) [35], and its homozygous deletion has been reported in some glioblastoma multiformes [36]. More recently, *DACH1* has been reported to inhibit the growth of lung adenocarcinoma cells through its binding to TP53 [37]. *DACH1*'s putative tumor suppressive function has also been documented in studies of breast, prostate, and uterine cancers (reviewed in [38]), where its expression was found to be frequently downregulated. In contrast, upregulated expression has been reported in advanced ovarian cancers [39] and in CD15+ myeloid progenitor cells harboring the t(9;11) translocation [40], and there is some evidence that it exerts oncogenic effects in t(9;11) acute myeloid leukemia [41].

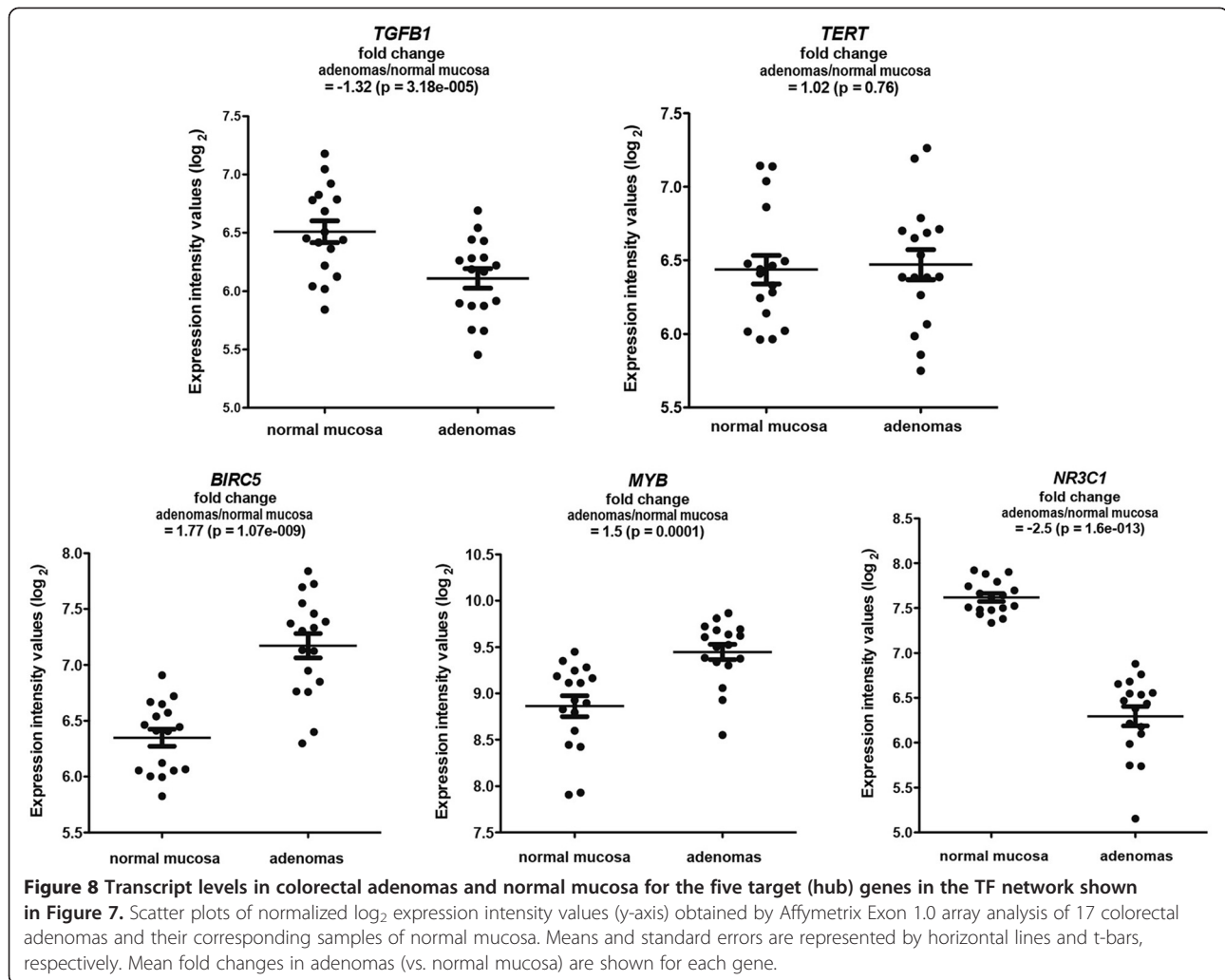
In light of these findings, the transcription-regulating roles of *DACH1* in somatic tissues—and their implications for tumorigenesis in a given tissue—might be expected to vary widely. The anti-*DACH1* antibody used in our study is highly reliable for exploring this question, but it has been commercially available only recently. Immunostaining patterns in extracolonic tissues and tumors obtained with older antibodies might therefore need to be re-examined.

A second approach used in the analysis of the TF genes listed in Additional file 8: Table S8 involved the identification of hub genes in networks that could be

built with the selected TF genes. Like many other proteins, TFs interact with the products of other genes. Interaction networks are very useful to better understand the functional significance of gene expression changes. Each TF influences the expression of several genes, producing changes in the levels of mRNA and, in many cases, also in the levels of the corresponding proteins. Thus, it is important to know the expression level of the TF gene itself, but also that of the other genes in its network(s).

For this reason, we used GeneGo (portal.genego.com) to build networks using the 55 TF genes identified by all three of our selection procedures. The most significant network included 27 of the 55 TF genes. Each of the five hub genes identified within this network (*TGFBI*, *BIRC5*, *NR3C1*, *MYB*, and *TERT*) (Figure 7) is known to play roles in at least one fundamental cellular process involved in tumorigenesis [42-47]. Figure 8 shows how the expression of these hub genes changes as normal colorectal mucosa undergoes adenomatous transformation. The downregulated *TGFBI* transcription we observed in colorectal adenomas (Figure 8) is consistent with previous reports, which described upregulation of this gene only in advanced colorectal tumors (severely dysplastic adenomas and cancers) [48-50]. These findings suggest that the proapoptotic function of *TGFBI* [51], which is important for maintenance of homeostasis in the normal colorectal epithelium, might decline in the early phases of colorectal tumor growth. Indeed, sulindac treatment has been shown to upregulate apoptosis in certain areas of colorectal adenomas, and these same areas also displayed increased *TGFBI* expression [52]. *TGFBI*'s growth inhibition is believed to be replaced by tumor-promoting functions, i.e., immunosuppression and angiogenesis, in more advanced tumors, where its expression is in fact increased [51]. Impaired apoptosis, an essential feature of early adenomatous growth, might also be related to the increased expression of *BIRC5* we documented in our adenomas (Figure 8). *BIRC5* is a well-known member of the inhibitor of apoptosis gene family [45], and its overexpression in precancerous colorectal lesions has been well-documented [53-56].

It is more difficult to predict the functional impact on colorectal tumorigenesis of the striking downregulated expression of the glucocorticoid receptor gene *NR3C1* in all the adenomas we examined (Figure 8). The mechanisms underlying this nuclear receptor's control of transcription in the intestinal epithelium (and other tissues or cells) are still unknown [57]. Its decreased expression in our adenomas might be related to epigenetic modifications involving its promoter region, which could eventually lead to cytosine hypermethylation as these lesions progress [58]. Upregulated *MYB* expression (Figure 8) has already been reported in human



and mouse colorectal tumors, including adenomas [59]. In *APC* (*Min/+*) mice that are also haploinsufficient for *Myb*, adenoma formation is delayed, and co-operation between *Myb* and Wnt signaling appears to play a crucial role in this process [60].

As for *TERT*, the fifth hub in this network, its expression in our adenomas was not significantly different from that in normal mucosa (Figure 8). *TERT* is normally expressed in progenitor cells, and its overexpression has been implicated in the transformation of colorectal epithelia and many other types of tumorigenesis as well. Its expression in colorectal adenomas has not been investigated in large studies, but it appears to undergo a gradual increase during progression from adenomas to carcinomas [61,62]. Our adenomas were probably not advanced enough to display significantly upregulated *TERT* expression (all were larger than 10 mm, but most were characterized by low-level dysplasia). Nonetheless, *TERT*'s putative role as a major player in colorectal cellular transformation emerged from our MetaCore TF analysis, owing in all probability to

significant expression changes involving other molecules that interact with *TERT* in the same network.

In a previous report, we provided a thorough description of the sequential dysregulation of biological pathways that occurs along the adenoma-to-carcinoma sequence, based on analysis of our transcriptomic data [14]. In the present study, we focused on precancerous colorectal lesions and compared our findings with those obtained in colorectal carcinomas using the same approach depicted in Figure 1. Roughly half the TF gene expression perturbations found in carcinomas were already evident in adenomas (Additional file 12: Figure S6; genes listed in Additional file 13: Table S6 and Additional file 14: Table S7), suggesting that the tumorigenic transcriptional program is already well underway during the preinvasive stage. However, a similarly large proportion of TF genes were dysregulated only in carcinomas, which indicates that this program undergoes an important change across the adenoma-to-cancer transition. Extensive validation studies will be necessary to shed light on the biological and clinical implications of the similarities

and differences of the transcriptional program between these two stages of transformation.

## Conclusions

This study provides novel information on the TF gene transcript levels associated with adenomatous transformation of the colorectal epithelium and identifies 261 TF genes that appear to play roles in colorectal tumorigenesis. We pinpointed the TF genes whose expression is significantly altered in colorectal adenomas and characterized the extent and direction of these changes. Integrating these findings with those observed in the entire transcriptome allowed us to identify a few hub genes, which may play crucial roles in the formation and progression of adenomas. Finally, we provide useful information on numerous TF genes whose roles in colorectal tumorigenesis have been relatively unexplored, such as *DACH1*, a development gene whose protein expression patterns in colorectal tissues raises interesting questions about its involvement in tumor growth. This study represents a very early step toward a better understanding the highly complex transcription network of a given tissue and tumor. The function of any TF does not depend solely on its expression level but on many other aspects, such as DNA occupancy levels and the tissue-specific availability of factors it interacts with [12,57]. Our findings must be complemented with studies designed to address these aspects of the transcriptional network in colorectal tissues.

## Additional files

**Additional file 1: Table S1.** Characteristics of the 17 patients with pedunculated adenomas included in the study.

**Additional file 2: Table S2.** Differential expression of 1218 TF genes in colorectal adenomas and corresponding normal mucosa samples. Genes are listed in order of ascending p value; bold-face type: 315 genes with a p value <0.01.

**Additional file 3: Table S3.** Legend of numeric attributes in the MetaCore TF analysis.

**Additional file 4: Table S4.** The 793 TF genes whose networks were enriched in genes displaying significant differential expression in adenomas (Figure 1, MetaCore analysis). Genes are listed in order of descending z-score. Bold-face type: genes with a z-scores > 2; red: "under-researched" genes (NormDP > 0).

**Additional file 5: Figure S1.** Actual and expected numbers of colorectal tumorigenesis-related publications dealing with each TF gene. Relationship between z-score (x-axis) and actual number of publications (y-axis) for each TF gene. The trend line has an intercept = 0 and a slope  $\alpha = 142$ . The  $\alpha$  value can be used to predict the expected number of publications. This allowed us to distinguish between TFs that have been "under-researched" (below the trend line) and "over-researched" (above the trend line) in the field of colorectal tumorigenesis.

**Additional file 6: Table S5.** Characteristics of the 80 colon cancer patients tested for DACH1 expression.

**Additional file 7: Figure S2.** Hierarchical clustering analysis of colorectal tissue samples based on expression levels of 1218 TF genes. (Pearson correlation, Ward distance). The 34 tissue samples represented

on the x-axis include 17 normal mucosal samples (normal) and 17 adenomas (adenoma). Each transcript probe set plotted on the y-axis is color-coded to reflect expression levels of the 1218 TF genes relative to their median expression levels across the entire tissue-sample set (red: high; green: low).

**Additional file 8: Table S8.** Scores of 261 TF genes that passed at least 2 of the 3 selection thresholds shown in Figure 1. Genes are listed in the order of ascending p values.

**Additional file 9: Figure S3.** DACH1 immunohistochemical staining of sections from formalin-fixed, paraffin-embedded pellets of 6 colon cancer cell lines. Affymetrix U133Plus2.0 raw mRNA *DACH1* expression values are reported for each cell line. These gene expression levels were consistent with the complete absence of DACH1 protein in CO115 cells (A) and its extremely weak expression in HCT116 cells (B). Strongly stained nuclei were rare in the population of Caco2 cells (C) but much more common in LS180 cells (D). Most of the SW620 cell nuclei were positive for DACH1 expression, which is consistent with the high *DACH1* mRNA expression value for these cells (E). Interestingly, DACH1 expression was absent in SW480 cells (F), which were established from a primary colon cancer whose lymph node metastasis was used to establish the SW620 cell line.

**Additional file 10: Figure S4.** Legend of the symbolic attributes of the gene networks shown in Figure 7 and Additional file 11: Figure S5.

**Additional file 11: Figure S5.** The most significant network included 27 of the 55 TF genes found in all three sets depicted in Figure 2. In this panel, the *subcellular localization* layout of MetaCore is shown.

**Additional file 12: Figure S6.** Venn diagrams showing intersection of TF gene sets identified in our analyses of colorectal adenomas and carcinomas (each compared with matched samples of normal mucosa). Top diagram: The number of TF genes was identified with the *t* test-based procedure shown in Figure 1 (left prong). Bottom diagram: The number of TF genes was identified with the MetaCore-based procedure shown in Figure 1 (right prong). Note that the number of TF genes identified with the latter approach in adenomas and carcinomas refer to version 6.16, build 63671 of MetaCore™ software.

**Additional file 13: Table S6.** Differential expression of 1038 TF genes in colorectal carcinomas vs. corresponding samples of normal mucosa. Genes are listed in order of ascending p values. Bold-face type: 232 genes with a p value <0.01. The 107 TF genes listed in red were also dysregulated in adenomas (see Venn diagram in Additional file 11: Figure S5).

**Additional file 14: Table S7.** The 633 TF genes whose networks were enriched in genes displaying significant differential expression in carcinomas (MetaCore analysis). Genes are listed in order of descending z-scores. The 255 genes listed in bold-face type had z-scores of > 2.

## Abbreviations

COBRA: Combined bisulfite restriction analysis; DP: Difference in publications; GO: Gene ontology; MMR: Mismatch repair; NormDP: Normalized DP; TF: Transcription factor; TRANSFAC: TRANSCRIPTION FACTOR.

## Competing interests

All authors deny conflicts of interest relevant to the manuscript.

## Authors' contributions

JV and MJO performed microarray data analyses and most of the experiments; MM performed COBRA experiments; EC processed microarrays; DP-O histologically classified tissue samples; RH performed immunohistochemistry; JJ conceived important experiments during the study; TS and FB performed endoscopies and tissue sampling; and GM conceived the project, obtained funding, prepared the manuscript, and served as supervising mentor for JV during her Master studies. All authors read and approved the final manuscript.

## Acknowledgements

The study was supported by a grant from the Swiss National Science Foundation (no. 31003A\_141225) to Giancarlo Marra. Marian Everett Kent provided writing assistance and editing services.

#### Author details

<sup>1</sup>Institute of Molecular Cancer Research, University of Zurich, Winterthurerstrasse 190, Zurich 8051, Switzerland. <sup>2</sup>Functional Genomics Center, University of Zurich and ETH Zurich, Zurich, Switzerland. <sup>3</sup>Institute of Pathology, Triemli Hospital Zurich, Zurich, Switzerland. <sup>4</sup>Endoscopy and Gastroenterology Unit, Hospital of Cremona, Cremona, Italy.

Received: 14 August 2013 Accepted: 23 January 2014

Published: 29 January 2014

#### References

- Winawer SJ, Zauber AG, Fletcher RH, Stillman JS, O'Brien MJ, Levin B, Smith RA, Lieberman DA, Burt RW, Levin TR, Bond JH, Brooks D, Byers T, Hyman N, Kirk L, Thorson A, Simmam C, Johnson D, Rex DK: **Guidelines for colonoscopy surveillance after polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer and the American Cancer Society.** *Gastroenterology* 2006, **130**(6):1872–1885.
- Stryker SJ, Wolff BG, Culp CE, Libbe SD, Ilstrup DM, MacCarty RL: **Natural history of untreated colonic polyps.** *Gastroenterology* 1987, **93**(5):1009–1013.
- Sabates-Bellver J, Van der Flier LG, de Palo M, Cattaneo E, Maake C, Rehrauer H, Laczko E, Kurovski MA, Bujnicki JM, Menigatti M, Luz J, Ranalli TV, Gomes V, Pastorelli A, Faggiani R, Anti M, Jiricny J, Clevers H, Marra G: **Transcriptome profile of human colorectal adenomas.** *Mol Cancer Res* 2007, **5**(12):1263–1275.
- Cattaneo E, Baudis M, Buffoli F, Bianco MA, Zorzi F, Marra G: **Pathways and crossroads to colorectal cancer.** In *Pre-Invasive Disease: Pathogenesis and Clinical Management*. Edited by Fitzgerald RC. Springer; 2011:369–394.
- Pinto D, Clevers H: **Wnt, stem cells and cancer in the intestine.** *Biol Cell* 2005, **97**(3):185–196.
- Dang CV: **MYC on the path to cancer.** *Cell* 2012, **149**(1):22–35.
- Sansom OJ, Meniel VS, Muncan V, Phesse TJ, Wilkins JA, Reed KR, Vass JK, Athineos D, Clevers H, Clarke AR: **Myc deletion rescues Apc deficiency in the small intestine.** *Nature* 2007, **446**(7136):676–679.
- Dang CV, O'Donnell KA, Zeller KI, Nguyen T, Osthus RC, Li F: **The c-Myc target gene network.** *Semin Cancer Biol* 2006, **16**(4):253–264.
- Ji H, Wu G, Zhan X, Nolan A, Koh C, De Marzo A, Doan HM, Fan J, Cheadle C, Fallahi M, Cleveland JL, Dang CV, Zeller KI: **Cell-type independent MYC target genes reveal a primordial signature involved in biomass accumulation.** *PLoS One* 2011, **6**(10):e26057.
- Lin CY, Loven J, Rahl PB, Paranal RM, Burge CB, Bradner JE, Lee TI, Young RA: **Transcriptional amplification in tumor cells with elevated c-Myc.** *Cell* 2012, **151**(1):56–67.
- Nie Z, Hu G, Wei G, Cui K, Yamane A, Resch W, Wang R, Green DR, Tessarollo L, Casellas R, Zhao K, Levens D: **c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells.** *Cell* 2012, **151**(1):68–79.
- Biggin MD: **Animal transcription networks as highly connected, quantitative continua.** *Dev Cell* 2011, **21**(4):611–626.
- Cattaneo E, Laczko E, Buffoli F, Zorzi F, Bianco MA, Menigatti M, Bartosova Z, Haider R, Helmchen B, Sabates-Bellver J, Tiwari A, Jiricny J, Marra G: **Preinvasive colorectal lesion transcriptomes correlate with endoscopic morphology (polypoid vs. nonpolypoid).** *EMBO Mol Med* 2011, **3**(6):334–347.
- Maglietta R, Liuzzi VC, Cattaneo E, Laczko E, Piepoli A, Panza A, Carella M, Palumbo O, Staiano T, Buffoli F, Andriulli A, Marra G, Ancona N: **Molecular pathways undergoing dramatic transcriptomic changes during tumor development in the human colon.** *BMC Cancer* 2012, **12**:608.
- Vaquerez JM, Kummerfeld SK, Teichmann SA, Luscombe NM: **A census of human transcription factors: function, expression and evolution.** *Nat Rev Genet* 2009, **10**(4):252–263.
- Hubbard TJ, Aken BL, Beal K, Ballester B, Caccamo M, Chen Y, Clarke L, Coates G, Cunningham F, Cutts T, Down T, Dyer SC, Fitzgerald S, Fernandez-Banet J, Graf S, Haider S, Hammond M, Herrero J, Holland R, Howe K, Howe K, Johnson N, Kahari A, Keefe D, Kokocinski F, Kulesha E, Lawson D, Longden I, Melsopp C, Megy K, et al: **Ensembl 2007.** *Nucleic Acids Res* 2007, **35**(Database issue):D610–617.
- Wingender E, Chen X, Hehl R, Karas H, Liebich I, Matys V, Meinhardt T, Pruss M, Reuter I, Schacherer F: **TRANSFAC: an integrated system for gene expression regulation.** *Nucleic Acids Res* 2000, **28**(1):316–319.
- Nikolsky Y, Ekins S, Nikolskaya T, Bugrim A: **A novel method for generation of signature networks as biomarkers from complex high throughput data.** *Toxicol Lett* 2005, **158**(1):20–29.
- Ferreira JA, Zwiderman A: **Approximate sample size calculations with microarray data: an illustration.** *Stat Appl Genet Mol Biol* 2006, **5**:25.
- Truninger K, Menigatti M, Luz J, Russell A, Haider R, Gebbers JO, Bannwart F, Yurtsever H, Neuweiler J, Riehle HM, Cattaruzza MS, Heinemann K, Schär P, Jiricny J, Marra G: **Immunohistochemical analysis reveals high frequency of PMS2 defects in colorectal cancer.** *Gastroenterology* 2005, **128**(5):1160–1171.
- Menigatti M, Cattaneo E, Sabates-Bellver J, Ilinsky VV, Went P, Buffoli F, Marquez VE, Jiricny J, Marra G: **The protein tyrosine phosphatase receptor type R gene is an early and frequent target of silencing in human colorectal tumorigenesis.** *Mol Cancer* 2009, **8**:124.
- Barrett T, Troup DB, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Muetterter RN, Holko M, Ayhanbule O, Yefanov A, Soboleva A: **NCBI GEO: archive for functional genomics data sets—10 years on.** *Nucleic Acids Res* 2011, **39**(Database issue):D1005–1010.
- di Pietro M, Sabates Bellver J, Menigatti M, Bannwart F, Schnider A, Russell A, Truninger K, Jiricny J, Marra G: **Defective DNA mismatch repair determines a characteristic transcriptional profile in proximal colon cancers.** *Gastroenterology* 2005, **129**(3):1047–1059.
- Schwanhaussner B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M: **Global quantification of mammalian gene expression control.** *Nature* 2011, **473**(7347):337–342.
- Mardon G, Solomon NM, Rubin GM: **Dachshund encodes a nuclear protein required for normal eye and leg development in Drosophila.** *Development* 1994, **120**(12):3473–3486.
- Chen R, Amoui M, Zhang Z, Mardon G: **Dachshund and eyes absent proteins form a complex and function synergistically to induce ectopic eye development in Drosophila.** *Cell* 1997, **91**(7):893–903.
- Li X, Oghii KA, Zhang J, Krones A, Bush KT, Glass CK, Nigam SK, Aggarwal AK, Maas R, Rose DW, Rosenfeld MG: **Eya protein phosphatase activity regulates Six1-Dach-Eya transcriptional effects in mammalian organogenesis.** *Nature* 2003, **426**(6964):247–254.
- Li X, Perissi V, Liu F, Rose DW, Rosenfeld MG: **Tissue-specific regulation of retinal and pituitary precursor cell proliferation.** *Science* 2002, **297**(5584):1180–1183.
- Lee JW, Kim HS, Kim S, Hwang J, Kim YH, Lim GY, Sohn WJ, Yoon SR, Kim JY, Park TS, Park KM, Ryoo ZY, Lee S: **DACH1 regulates cell cycle progression of myeloid cells through the control of cyclin D, Cdk 4/6 and p21Cip1.** *Biochem Biophys Res Commun* 2012, **420**(1):91–95.
- Wu K, Yang Y, Wang C, Davoli MA, D'Amico M, Li A, Cveklava K, Kozmik Z, Lisanti MP, Russell RG, Cvekl A, Pestell RG: **DACH1 inhibits transforming growth factor-beta signaling through binding Smad4.** *J Biol Chem* 2003, **278**(51):51673–51684.
- Wu K, Liu M, Li A, Donniger H, Rao M, Jiao X, Lisanti MP, Cvekl A, Birrer M, Pestell RG: **Cell fate determination factor DACH1 inhibits c-Jun-induced contact-independent growth.** *Mol Biol Cell* 2007, **18**(3):755–767.
- Wu K, Li A, Rao M, Liu M, Dailey V, Yang Y, Di Vizio D, Wang C, Lisanti MP, Sauter G, Russell RG, Cvekl A, Pestell RG: **DACH1 is a cell fate determination factor that inhibits cyclin D1 and breast tumor growth.** *Mol Cell Biol* 2006, **26**(19):7116–7129.
- Munoz J, Stange DE, Schepers AG, van de Wetering M, Koo BK, Itzkovitz S, Volckmann R, Kung KS, Koster J, Radulescu S, Myant K, Versteeg R, Sansom OJ, van Es JH, Barker N, van Oudenaarden A, Mohammed S, Heck AJ, Clevers H: **The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent +4 cell markers.** *Embo J* 2012, **31**(14):3079–3091.
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES: **A bivalent chromatin structure marks key developmental genes in embryonic stem cells.** *Cell* 2006, **125**(2):315–326.
- Govindan R, Ding L, Griffith M, Subramanian J, Dees ND, Kanchi KL, Maher CA, Fulton R, Fulton L, Wallis J, Chen K, Walker J, McDonald S, Bose R, Orntz D, Xiong D, You M, Dooling DJ, Watson M, Mardis ER, Wilson RK: **Genomic landscape of non-small cell lung cancer in smokers and never-smokers.** *Cell* 2012, **150**(6):1121–1134.
- Watanabe A, Ogiwara H, Ehata S, Mukasa A, Ishikawa S, Maeda D, Ueki K, Ino Y, Todo T, Yamada Y, Fukuyama M, Saito N, Miyazono K, Aburatani H: **Homozygously deleted gene DACH1 regulates tumor-initiating activity of glioma cells.** *Proc Natl Acad Sci U S A* 2011, **108**(30):12384–12389.
- Chen K, Wu K, Cai S, Zhang W, Zhou J, Wang J, Ertel A, Li Z, Rui H, Quong A, Lisanti MP, Tozeren A, Tanes C, Addya S, Gormley M, Wang C, McMahon SB, Pestell RG: **Dachshund binds p53 to block the growth of lung adenocarcinoma cells.** *Cancer Res* 2013, **73**(11):3262–3274.



38. Popov VM, Wu K, Zhou J, Powell MJ, Mardon G, Wang C, Pestell RG: **The Dachshund gene in development and hormone-responsive tumorigenesis.** *Trends Endocrinol Metab* 2010, **21**(1):41–49.
39. Sunde JS, Donninger H, Wu K, Johnson ME, Pestell RG, Rose GS, Mok SC, Brady J, Bonome T, Birrer MJ: **Expression profiling identifies altered expression of genes that contribute to the inhibition of transforming growth factor-beta signaling in ovarian cancer.** *Cancer Res* 2006, **66**(17):8404–8412.
40. Lee S, Chen J, Zhou G, Shi RZ, Bouffard GG, Kocherginsky M, Ge X, Sun M, Jayathilaka N, Kim YC, Emmanuel N, Bohlander SK, Minden M, Kline J, Ozer O, Larson RA, LeBeau MM, Green ED, Trent J, Karrison T, Liu PP, Wang SM, Rowley JD: **Gene expression profiles in acute myeloid leukemia with common translocations using SAGE.** *Proc Natl Acad Sci U S A* 2006, **103**(4):1030–1035.
41. Lee JW, Kim HS, Hwang J, Kim YH, Lim GY, Sohn WJ, Yoon SR, Kim JY, Park TS, Oh SH, Park KM, Choi SU, Ryou ZY, Lee S: **Regulation of HOXA9 activity by predominant expression of DACH1 against C/EBPalpha and GATA-1 in myeloid leukemia with MLL-AF9.** *Biochem Biophys Res Commun* 2012, **426**(3):299–305.
42. Hanahan D, Weinberg RA: **Hallmarks of cancer: the next generation.** *Cell* 2011, **144**(5):646–674.
43. Massague J: **TGFbeta in Cancer.** *Cell* 2008, **134**(2):215–230.
44. Donate LE, Blasco MA: **Telomeres in cancer and ageing.** *Philos Trans R Soc Lond B Biol Sci* 2011, **366**(1561):76–84.
45. Hunter AM, LaCasse EC, Korneluk RG: **The inhibitors of apoptosis (IAPs) as cancer targets.** *Apoptosis* 2007, **12**(9):1543–1568.
46. Ramsay RG, Gonda TJ: **MYB function in normal and cancer cells.** *Nat Rev Cancer* 2008, **8**(7):523–534.
47. Lind GE, Kleivi K, Meling GI, Teixeira MR, Thiis-Evensen E, Rognum TO, Lothe RA: **ADAMTS1, CRABP1, and NR3C1 identified as epigenetically deregulated genes in colorectal tumorigenesis.** *Cell Oncol* 2006, **28**(5–6):259–272.
48. Daniel P, Wagrowska-Danilewicz M, Danilewicz M, Stasikowska O, Malecka-Panas E: **Transforming growth factor beta 1 and metalloproteinase-9 overexpression in colorectal cancer (CC) and adenoma.** *Int J Colorectal Dis* 2007, **22**(10):1165–1172.
49. Hawinkels LJ, Verspaget HW, van der Reijden JJ, van der Zon JM, Verheijen JH, Hommes DW, Lamers CB, Sier CF: **Active TGF-beta1 correlates with myofibroblasts and malignancy in the colorectal adenoma-carcinoma sequence.** *Cancer Sci* 2009, **100**(4):663–670.
50. Bellone G, Gramigni C, Vizio B, Mauri FA, Prati A, Solerio D, Dughera L, Ruffini E, Gasparri G, Camandona M: **Abnormal expression of Endoglin and its receptor complex (TGF-beta1 and TGF-beta receptor II) as early angiogenic switch indicator in premalignant lesions of the colon mucosa.** *Int J Oncol* 2010, **37**(5):1153–1165.
51. Siegel PM, Massague J: **Cytostatic and apoptotic actions of TGF-beta in homeostasis and cancer.** *Nat Rev Cancer* 2003, **3**(11):807–821.
52. Masunaga R, Kohno H, Dhar DK, Kotoh T, Tabara H, Tachibana M, Kubota H, Nagasue N: **Enhanced apoptosis and transforming growth factor-beta1 expression in colorectal adenomas and carcinomas after Sulindac therapy.** *Dis Colon Rectum* 2001, **44**(7):1008–1015.
53. Kawasaki H, Toyoda M, Shinohara H, Okuda J, Watanabe I, Yamamoto T, Tanaka K, Tenjo T, Tanigawa N: **Expression of survivin correlates with apoptosis, proliferation, and angiogenesis during human colorectal tumorigenesis.** *Cancer* 2001, **91**(11):2026–2032.
54. Lin LJ, Zheng CQ, Jin Y, Ma Y, Jiang WG, Ma T: **Expression of survivin protein in human colorectal carcinogenesis.** *World J Gastroenterol* 2003, **9**(5):974–977.
55. Parfitt JR, Driman DK: **Survivin and hedgehog protein expression in serrated colorectal polyps: an immunohistochemical study.** *Hum Pathol* 2007, **38**(5):710–717.
56. Hernandez JM, Farma JM, Coppola D, Hakam A, Fulp WJ, Chen DT, Siegel EM, Yeatman TJ, Shibata D: **Expression of the antiapoptotic protein survivin in colon cancer.** *Clin Colorectal Cancer* 2011, **10**(3):188–193.
57. Clark AR, Belvisi MG: **Maps and legends: the quest for dissociated ligands of the glucocorticoid receptor.** *Pharmacol Ther* 2012, **134**(1):54–67.
58. Ahlquist T, Lind GE, Costa VL, Meling GI, Vatn M, Hoff GS, Rognum TO, Skotheim RI, Thiis-Evensen E, Lothe RA: **Gene methylation profiles of normal mucosa, and benign and malignant colorectal tumors identify early onset markers.** *Mol Cancer* 2008, **7**:94.
59. Ernst M, Ramsay RG: **Colorectal cancer mouse models: integrating inflammation and the stroma.** *J Gastroenterol Hepatol* 2012, **27**(1):39–50.
60. Ciznadija D, Tothill R, Waterman ML, Zhao L, Huynh D, Yu RM, Ernst M, Ishii S, Mantamadiotis T, Gonda TJ, Ramsay RG, Malaterre J: **Intestinal adenoma formation and MYC activation are regulated by cooperation between MYB and Wnt signaling.** *Cell Death Differ* 2009, **16**(11):1530–1538.
61. Kolquist KA, Ellisen LW, Counter CM, Meyerson M, Tan LK, Weinberg RA, Haber DA, Gerald WL: **Expression of TERT in early premalignant lesions and a subset of cells in normal tissues.** *Nat Genet* 1998, **19**(2):182–186.
62. Terrin L, Rampazzo E, Pucciarelli S, Agostini M, Bertorelle R, Esposito G, DelBianco P, Nitti D, De Rossi A: **Relationship between tumor and plasma levels of hTERT mRNA in patients with colorectal cancer: implications for monitoring of neoplastic disease.** *Clin Cancer Res* 2008, **14**(22):7444–7451.

doi:10.1186/1471-2407-14-46

**Cite this article as:** Vonlanthen et al.: A comprehensive look at transcription factor gene expression changes in colorectal adenomas. *BMC Cancer* 2014 **14**:46.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

