







## *BRAF*<sup>V600E</sup>-mutated serrated colorectal neoplasia drives transcriptional activation of cholesterol metabolism

Paulina Rzasa <sup>1,8</sup>, Sarah Whelan<sup>1,8</sup>, Pooyeh Farahmand <sup>1,8</sup>, Hong Cai<sup>1</sup>, Inna Guterman<sup>1</sup>, Raquel Palacios-Gallego<sup>1</sup>, Shanthi S. Undru<sup>1</sup>, Lauren Sandford <sup>1,7</sup>, Caleb Green <sup>1</sup>, Catherine Andreadi<sup>1</sup>, Maria Mintseva<sup>1,2</sup>, Emma Parrott<sup>1</sup>, Hong Jin<sup>1</sup>, Fiona Hey<sup>1</sup>, Susan Giblett<sup>1</sup>, Nicolas B. Sylvius<sup>3</sup>, Natalie S. Allcock<sup>4</sup>, Anna Straatman-Iwanowska<sup>4</sup>, Roberto Feuda<sup>5</sup>, Cristina Tufarelli <sup>1</sup>, Karen Brown<sup>1</sup>, Catrin Pritchard<sup>1</sup> & Alessandro Rufini <sup>1,6</sup>✉

*BRAF* mutations occur early in serrated colorectal cancers, but their long-term influence on tissue homeostasis is poorly characterized. We investigated the impact of short-term (3 days) and long-term (6 months) expression of *Braf*<sup>V600E</sup> in the intestinal tissue of an inducible mouse model. We show that *Braf*<sup>V600E</sup> perturbs the homeostasis of intestinal epithelial cells, with impaired differentiation of enterocytes emerging after prolonged expression of the oncogene. Moreover, *Braf*<sup>V600E</sup> leads to a persistent transcriptional reprogramming with enrichment of numerous gene signatures indicative of proliferation and tumorigenesis, and signatures suggestive of metabolic rewiring. We focused on the top-ranking cholesterol biosynthesis signature and confirmed its increased expression in human serrated lesions. Functionally, the cholesterol lowering drug atorvastatin prevents the establishment of intestinal crypt hyperplasia in *Braf*<sup>V600E</sup>-mutant mice. Overall, our work unveils the long-term impact of *Braf*<sup>V600E</sup> expression in intestinal tissue and suggests that colorectal cancers with mutations in *BRAF* might be prevented by statins.

<sup>1</sup>Leicester Cancer Research Centre, University of Leicester, Leicester, UK. <sup>2</sup>Area of Neuroscience, International School for Advanced Studies (SISSA), Trieste, Italy. <sup>3</sup>NUCLEUS Genomics, Core Biotechnology Services, University of Leicester, Leicester, UK. <sup>4</sup>University of Leicester Core Biotechnology Services Electron Microscopy Facility, Leicester, UK. <sup>5</sup>Department of Genetics and Genome Biology, University of Leicester, Leicester, UK. <sup>6</sup>Dipartimento di Bioscienze, University of Milan, Milan, Italy. <sup>7</sup>Present address: Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, UK. <sup>8</sup>These authors contributed equally: Paulina Rzasa, Sarah Whelan, Pooyeh Farahmand. ✉email: [Alessandro.Rufini@unimi.it](mailto:Alessandro.Rufini@unimi.it)

The single-layered intestinal epithelium contains functionally distinct cellular populations. Proliferating intestinal stem cells (ISCs), whose self-renewal depends on active Wnt/ $\beta$ -catenin signaling<sup>1</sup>, are interspersed among Paneth cells at the bottom of epithelial invaginations, known as crypts of Lieberkühn, and can be identified by selective markers such as LGR5 and OLFM4<sup>2,3</sup>. ISCs generate transient amplifying cells, which have limited proliferative capability and migrate upwards towards intestinal villi where they differentiate into the absorptive enterocyte lineage or the secretory lineage, which includes Paneth cells, enteroendocrine cells (EECs), and goblet cells.

Colorectal cancer (CRC) is a heterogeneous disease arising from the intestinal epithelium through two main routes<sup>4,5</sup>. First, tumors originating from tubular, villous or tubulovillous precursor adenomas localize to the left colon, are characterized by sustained Wnt/ $\beta$ -catenin signaling, and are thought to develop from ISCs through a bottom-up fashion<sup>6–9</sup>. Second, tumors located in the right colon arise from precursor sessile serrated lesions, characterized by a saw-tooth-shaped folding of the dysplastic epithelium, a flat sessile morphology, and a mucinous histology<sup>10</sup>. This subtype of CRC develops in a top-bottom fashion through transdifferentiation of committed epithelial cells to gastric metaplasia and enrichment in fetal markers<sup>6,7,11</sup>. The most common driver mutation of serrated lesions is an activating mutation of the *BRAF* oncogene<sup>12–14</sup>. A valine to glutamate substitution at codon 600 (V600E) generates a constitutive active mutant (*BRAF*<sup>V600E</sup>) serine-threonine kinase that is responsible for the activation of the downstream MEK/ERK arm of the MAPK pathway<sup>15</sup>. Additional molecular features associated with right-sided colonic neoplasia include CpG island methylator phenotype (CIMP), microsatellite instability (MSI), and loss of TGF signaling<sup>10–12,14,16</sup>. Recently, heterogeneity of *BRAF*-driven CRCs has been uncovered through bulk RNA sequencing, and two main *BRAF*<sup>V600E</sup> CRC subtypes have been identified based on gene expression profile<sup>17</sup>. The BM1 subtype displays a gene expression profile enriched in EMT-related processes, *KRAS* signaling, and immune response, whereas the BM2 subtype is enriched in cell-cycle and cycle checkpoints-related processes, such as target genes of the E2F transcription factors and genes involved in the G2 to M transition of the cell cycle.

*Braf*<sup>V600E</sup> has been reported to be a poor oncogene when expressed in the mouse intestinal epithelium, whereby few tumors develop after long latency<sup>18–20</sup>. This prolonged latency agrees with the knowledge that human CRC develops covertly over 10–15 years<sup>21</sup> and indicates that intestinal tissue harbors the oncogenic *BRAF* mutation for a considerable time despite the absence of any clinical manifestation of the disease. Notwithstanding, little is known about the long-term impact of *BRAF*<sup>V600E</sup> expression on intestinal homeostasis, and the mechanisms that enable *BRAF*-driven CRC development remain incompletely understood. Recently, the limited intestinal tumorigenesis of *Braf*<sup>V600E</sup> mutant mice has been attributed to an imbalance between stemness and differentiation, whereby expression of *Braf*<sup>V600E</sup> promotes differentiation of intestinal epithelial cells at the expense of ISCs<sup>18,22</sup>.

The overall aim of the work described here was to attain a deeper understanding of serrated tumorigenesis and ascertain the impact of long-term expression of *Braf*<sup>V600E</sup>. To this end, we used a mouse model carrying an inducible *Braf*<sup>V600E</sup> allele<sup>20</sup> targeted to the intestinal epithelium and performed transcriptional profiling over a 6-month time course. Together with dynamic changes in tissue homeostasis, we show that *Braf*<sup>V600E</sup> orchestrates a rapid and persistent transcriptional reprogramming characterized by the enrichment of numerous gene signatures associated with *Braf*-driven CRC and additional signatures suggestive of metabolic rewiring. In particular, we observed a robust increase in the

expression of cholesterol biosynthesis genes. The functional relevance of those adaptations was confirmed by the ability of atorvastatin, a commonly prescribed pharmacological inhibitor of the mevalonate pathway and cholesterol biosynthesis, to prevent the establishment of crypt hyperplasia in *Braf*<sup>V600E</sup>-mutant mice. Finally, through the analysis of human bulk and single-cell transcriptomic datasets, we confirmed that an increased cholesterol gene signature is a hallmark of serrated CRC.

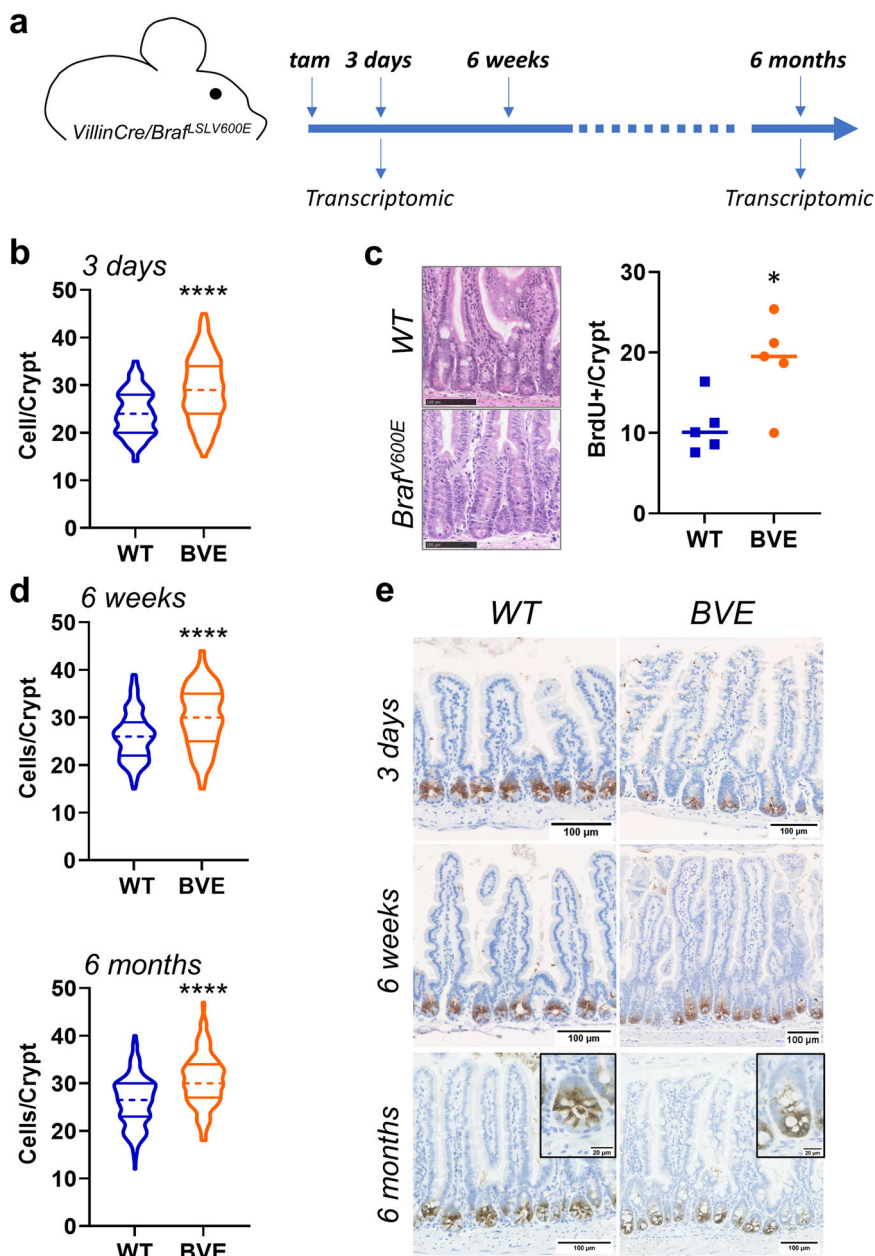
## Results

**Activation of *Braf*<sup>V600E</sup> in the intestine leads to crypt hyperplasia and persistence of crypt-resident ISCs.** Mice carrying an inducible *Braf*<sup>V600E</sup> knock-in allele were crossed with *VillinCre*<sup>ER</sup> mice to generate *VillinCre*<sup>ER/0</sup>/*Braf*<sup>V600E-LSL/+</sup> (BVE) and control *VillinCre*<sup>ER</sup>/*Braf*<sup>+/+</sup> (WT) mice. Intraperitoneal injection of tamoxifen in BVE mice enabled expression of the mutant *Braf*<sup>V600E</sup> allele. Mutant BVE mice displayed reduced survival (median survival 414 days versus 622 days in WT mice) (Supplementary Fig. 1a) and, as previously reported, neoplasia developed after long latency and with limited penetrance (Supplementary Fig. 1)<sup>18–20</sup>. To gain insight into short- and long-term outcomes of oncogene expression, we performed tissue histology (3 days, 6 weeks and 6 months after induction of *Braf*<sup>V600E</sup>) and transcriptomic analysis (3 days and 6 months after induction) of intestinal tissue of BVE and control mice (Fig. 1a and Supplementary Fig. 2). Analysis of transcriptomic data using GSEA confirmed a robust expression of a MAPK gene signature in mutant mice (Supplementary Fig. 3).

First, we assessed whether expression of mutant *BRAF* alters homeostasis of the intestinal crypt. Cell count indicated that as early as 3 days post induction, mutant crypts were hyperplastic (Fig. 1b and Supplementary Data 1). Hyperplasia was linked to increased cell proliferation (Fig. 1c, Supplementary Data 1) and was persistent, with a significantly increased number of cells detected in intestinal crypts of aged mutant mice (Fig. 1d and Supplementary Data 1). Previous reports have suggested that the expression of *Braf*<sup>V600E</sup> causes transient amplifying cell generation and cellular differentiation at the expense of the ISC population<sup>18,22</sup>. Using the stem cells marker *Olfm4*<sup>2</sup>, we were able to identify a persistent *Olfm4*<sup>+</sup> population localized at the bottom of the intestinal crypt in tissue specimens collected 3 days, 6 weeks, and 6 months after induction (Fig. 1e). Moreover, stem cell marker genes *Lgr5* and *Olfm4* in BVE mice were expressed at levels comparable to those in control mice 6 months after *Braf*<sup>V600E</sup> expression (Supplementary Fig. 4a). We also interrogated transcriptomic data, performing GSEA of an ISC signature. This analysis showed a mild downregulation early after *Braf* activation, but this was reverted in 6-month tissues (Supplementary Fig. 4b).

However, whereas in WT mice ISCs were interspersed with Paneth cells, in long-term induced mice ISCs clustered at the bottom of the intestinal crypt in long-term induced mice (Fig. 1e, insets), following a generalized loss and delocalization of Paneth cells (Supplementary Fig. 5). Finally, in keeping with the role of *Braf*<sup>V600E</sup> in serrated neoplasia, villi from BVE mice developed a saw-toothed appearance early after oncogene expression (Supplementary Fig. 6a) and, consistent with the establishment of tissue hyperplasia, villi length was significantly increased in 6-week and 6-month mutant mice (Supplementary Fig. 6b).

**Expression of *Braf*<sup>V600E</sup> modifies the abundance and distribution of differentiated intestinal populations.** We then assessed the impact of mutant *Braf* on differentiated intestinal cell populations. As early as 6 weeks post induction of *Braf*<sup>V600E</sup>, the intestinal crypts were deprived of Paneth cells, which migrated



**Fig. 1 Persistence of ISCs and crypt hyperplasia in *Braf<sup>V600E</sup>* mouse intestinal tissue.** **a** Schematic of the experimental work: *VillinCre<sup>ER/0</sup>/Braf<sup>V600E</sup>-LSL/+* were culled 3 days and 6 months post induction of *Braf<sup>V600E</sup>* for tissue histology and transcriptomic analysis. Mice were also culled at an intermediate time point 6 weeks post induction to corroborate changes in tissue histology. **b** Violin plots showing the distribution of number of cells per crypt in control mice and mice induced with *Braf<sup>V600E</sup>* for 3 days (BVE). Dotted and solid lines indicate median and quartiles, respectively. Data were analyzed by Kolmogorov-Smirnov test.  $n = 6$  mice per group in 3 days, 4 WT and 3 mutant mice in 6 weeks,  $n = 4$  mice per group in 6 months. 50 crypts were counted per each animal. \*\*\*\* $P \leq 0.0001$ . **c** H&E representative images of hyperplastic crypts and quantification of proliferative cells (BrdU+) 3 days post tamoxifen injection ( $n = 5$  animals per group, 1 female and 4 males, 31 crypts were counted for each animal, see also Supplementary Data 1). Data were analyzed using unpaired two-tailed *t* test. \* $P \leq 0.05$ . Bar size = 100  $\mu$ m. **d** Violin plots showing the distribution of number of cells per crypt in mice induced for 6 weeks ( $n = 4$  control males and 3 mutant males) and 6 months ( $n = 4$  mice per group). Dotted and solid lines indicate median and quartiles, respectively. Data were analyzed as in (a). \*\*\*\* $P \leq 0.0001$ . **e** Histological images of small intestinal tissue stained with the stem cell marker *Olm4* from WT control mice and mice with the *Braf<sup>V600E</sup>* mutation (BVE) at 3 days, 6 weeks and 6 months following tamoxifen induction. Bar size = 100  $\mu$ m.

upwards toward the villi (Supplementary Fig. 5) and were progressively lost, a phenotype previously observed in mouse models with altered MAPK activity<sup>23</sup>.

In addition, GSEA revealed an initial increase in the expression of an intestinal differentiation gene signature 3 days post *Braf<sup>V600E</sup>* induction. However, unexpectedly, the same signature was downregulated following long-term expression of mutant *Braf* (Fig. 2a). At this stage, we also identified a significant

downregulation of an intestinal enterocyte signature, which was reflected by reduced *Alpl* histological staining (Fig. 2b, c). A late impairment in the differentiation of enterocytes was confirmed by TEM, which showed shortened and less abundant microvilli on the surface of mutant enterocytes (Fig. 2d).

Next, we visualized and quantified mucus-producing goblet cells and EECs using Alcian Blue staining and immunofluorescence staining for Chromogranin A (ChgA), respectively (Supplementary

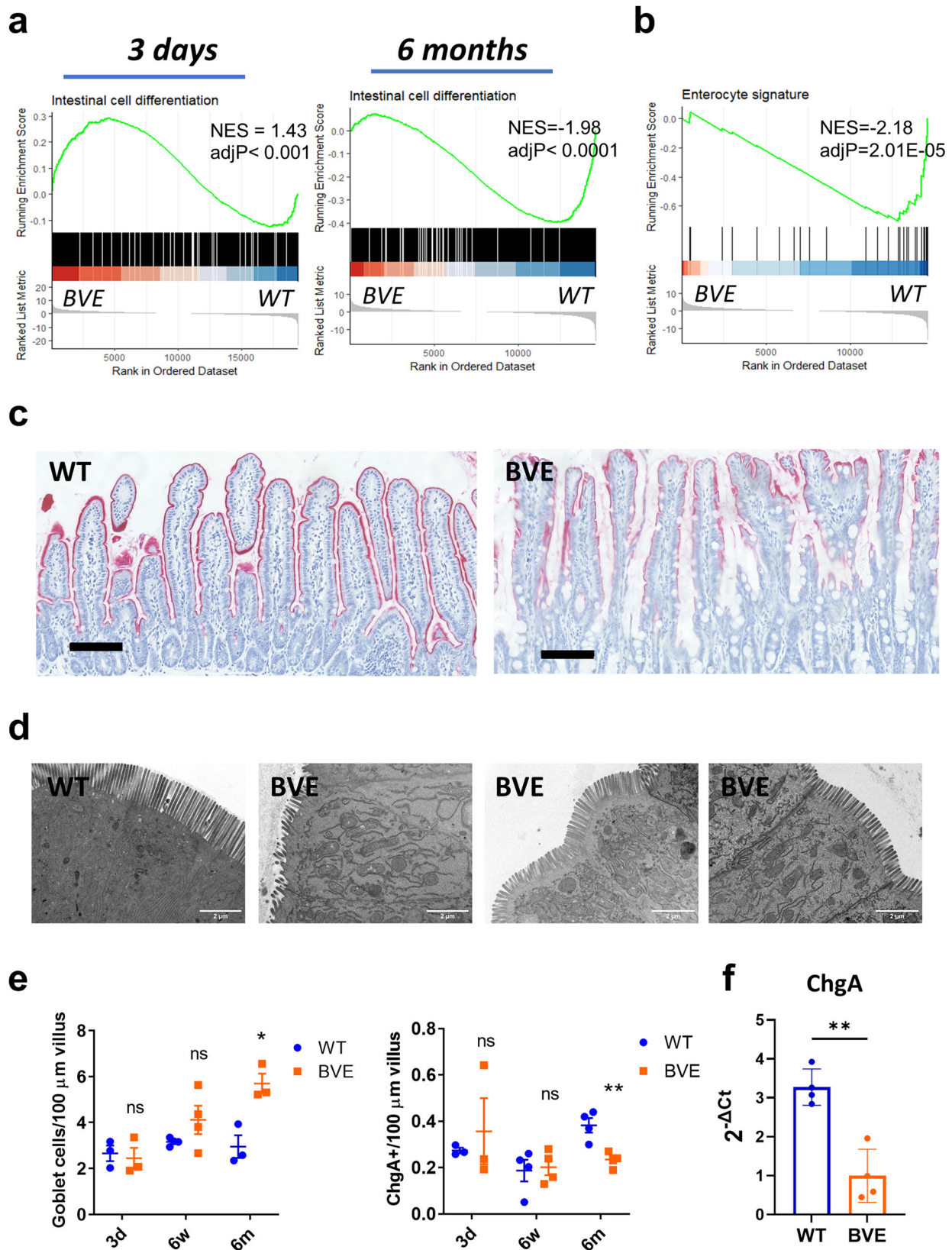


Fig. 7). Quantification of Alcian Blue- and ChgA-positive cells revealed an increased number of goblet cells, whereas EECs were reduced after long-term (6 months) expression of *Braf<sup>V600E</sup>* (Fig. 2e and Supplementary Data 1). rt-qPCR analysis confirmed the reduced expression of ChgA in aged mice (Fig. 2f and Supplementary Data 1).

These data suggest that *Braf<sup>V600E</sup>* alters the balance of intestinal secretory and absorptive cell lineages. The analysis of the long-term impact of oncogene expression particularly unveiled a significant impact on the numbers of goblet cells and EECs, and a broader downregulation of gene signatures

**Fig. 2 *Braf*<sup>V600E</sup> expression in the mouse intestinal epithelium alters intestinal homeostasis and affects absorptive and secretory cell lines.** **a** GSEA showing changes in differentiation gene signatures 3 days and 6 months following tamoxifen induction. **b** GSEA showing changes in enterocyte gene signatures 6 months following tamoxifen induction. NES normalized enrichment score. **c** Histological images of small intestinal tissue from WT control mice and mice with the *Braf*<sup>V600E</sup> mutation (BVE) 6 months following tamoxifen induction stained with alkaline phosphatase (Apl) to assess enterocyte cells. Bar size = 100  $\mu$ m. **d** TEM images of WT and *Braf*-mutant intestinal tissues showing altered microvilli in mutant mice. scale bar = 2  $\mu$ m. **e** Quantification of Goblet and ChgA+ enteroendocrine cells from the indicated mice. Each dot represents a single mouse, and lines represent mean  $\pm$  SEM. Data were analyzed by unpaired two-tailed *t* test ( $n = 3$  per group). \* $P \leq 0.05$ , \*\* $P \leq 0.001$ . **f** rt-qPCR analysis of the expression of the intestinal biomarkers Chromogranin A, (*ChgA*, EECs) 6 months after induction of *Braf*<sup>V600E</sup>. Data are plotted as mean  $\pm$  SD. Dots represent single replicates. Data were analyzed by unpaired two-tailed *t* test ( $n = 4$  per group). \* $P \leq 0.05$ .

associated with intestinal differentiation and, more specifically, with enterocyte differentiation.

Overall, our data confirm previous evidence of reduced stemness and increased differentiation early after induction of *Braf*<sup>V600E</sup>. They show, in addition, that those changes are transient and that stemness is restored in the intestinal tissue at the expense of differentiation. These findings emphasize the importance of long-term analysis of oncogene expression in the intestinal tissue and they are congruent with recent data suggesting a top-down origin of human serrated neoplasia triggered by dedifferentiation of intestinal cells<sup>6</sup>.

***Braf*<sup>V600E</sup> expression regulates a transcriptional program linked to CRC.** Next, we identified the molecular pathways regulated by *Braf*<sup>V600E</sup> by interrogating the Molecular Signatures Database (MSigDB)<sup>24</sup> (Supplementary Fig. 8). Signatures that scored high included pathways associated with cellular proliferation and the BM2 subtype of human CRC, such as *Myc*, E2F targets and G2/M checkpoint signatures (Fig. 3a), which were significantly enriched as early as 3 days post *Braf*<sup>V600E</sup> induction. The same signatures were confirmed to be enriched in *BRAF*<sup>V600E</sup> mutant CRCs compared with normal colon tissue from the TCGA dataset<sup>8</sup> (Supplementary Fig. 9). We also observed an enrichment of a canonical Wnt pathway signature in the mouse mutant intestinal tissue (Fig. 3b), in agreement with previous findings indicating that activation of *Braf*<sup>V600E</sup> results primarily in intestinal tumors displaying Wnt pathway activation<sup>19,20,22</sup>. More tellingly, long-term activation of *Braf*<sup>V600E</sup> was associated with the expression of *BRAF*-mutant CRC signatures that predict poor prognosis in CRC patients<sup>25</sup> (Fig. 3c).

To ascertain whether MAPK activity was necessary for the establishment of the observed transcriptional changes, we treated mice with the MEK inhibitor PD184352 (MEKi). Mice received daily treatments with vehicle or MEKi for 3 consecutive days and were then culled for transcriptomic analysis (Supplementary Fig. 10a). MEKi administration achieved a marked decrease in phospho-Erk in intestinal tissue, as assessed by Western blot analysis (Supplementary Fig. 10b), and stark repression of the MAPK signature<sup>19</sup> (Supplementary Fig. 10f). Notably, pharmacological inhibition of the MAPK pathway reverted the enrichment of some, but not all, signatures. In particular, the *Myc* signature, the canonical Wnt signature, and the Popovici poor-prognosis signature all depended on a functional MAPK pathway (Fig. 3d).

The origin of *BRAF*<sup>V600E</sup> CRC remains elusive. Recent evidence suggests that serrated neoplasia arises through intermediate gastric metaplasia and a reversion to an embryonic cellular stage, sustained by a fetal gene signature<sup>6,7,11</sup>. Indeed, human serrated lesions show enrichment for gene expression associated with metaplasia and fetal genes (Supplementary Fig. 11a). In agreement with previous observations<sup>11</sup>, expression of *Braf*<sup>V600E</sup> triggers significant and persistent enrichment of the fetal signature in intestinal tissue, which we show to be dependent on MAPK signaling (Supplementary Fig. 11b). However, we did not detect significant changes in the metaplasia gene signature<sup>6</sup>, although two genes within the signature, namely Aquaporin 5

(*Aqp5*) and Annexin A10 (*Anxa10*), were consistently upregulated in mutant tissue in a MAPK-dependent fashion (Supplementary Fig. 12). Interestingly, *Anxa10* has been previously reported to be a specific marker for human serrated lesions and CRCs of the serrated pathway<sup>22,26,27</sup>.

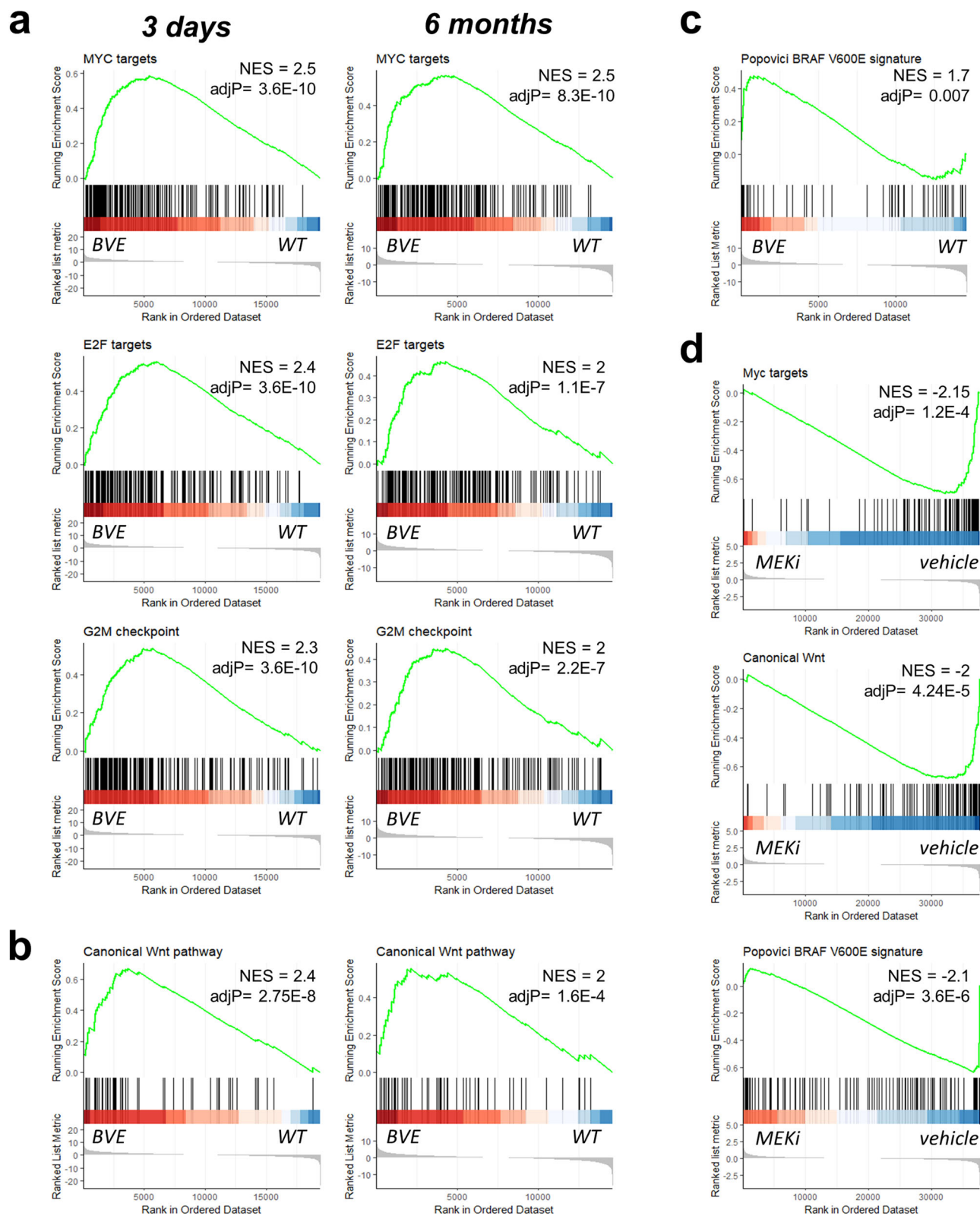
Overall, these data indicate that activation of *Braf*<sup>V600E</sup> orchestrates changes in the intestinal transcriptome that reflect the transcriptional reprogramming underlying colorectal carcinogenesis. These changes are durable, persisting for up to 6 months following oncogene expression.

***Braf*<sup>V600E</sup> expression rewires gene signatures of cholesterol metabolism to drive crypt hyperplasia.** We noticed that the top-scoring gene signature enriched in *Braf*-mutant mice was a cholesterol signature at both 3 days and 6 months following *Braf*<sup>V600E</sup> induction (Fig. 4a). The robust increase in this signature was also MAPK-dependent (Fig. 4b). The leading-edge genes responsible for the signature enrichment included several key metabolic enzymes of the mevalonate pathway, such as *Idi1*, *Mvk*, *Fdft1*, *Fdps*, *Sqle*, *Hmgcs1*, *Mvd*, *Hmgcs2*. Increased expression of cholesterol biosynthesis genes was also confirmed by rt-qPCR on RNA extracted from intestinal tissue (Fig. 4c and Supplementary Data 1). Next, we confirmed enrichment of the cholesterol biosynthesis signatures in RNA-seq datasets from two independent mouse models of *BRAF*-driven CRC (Fig. 4d)<sup>11,18</sup>.

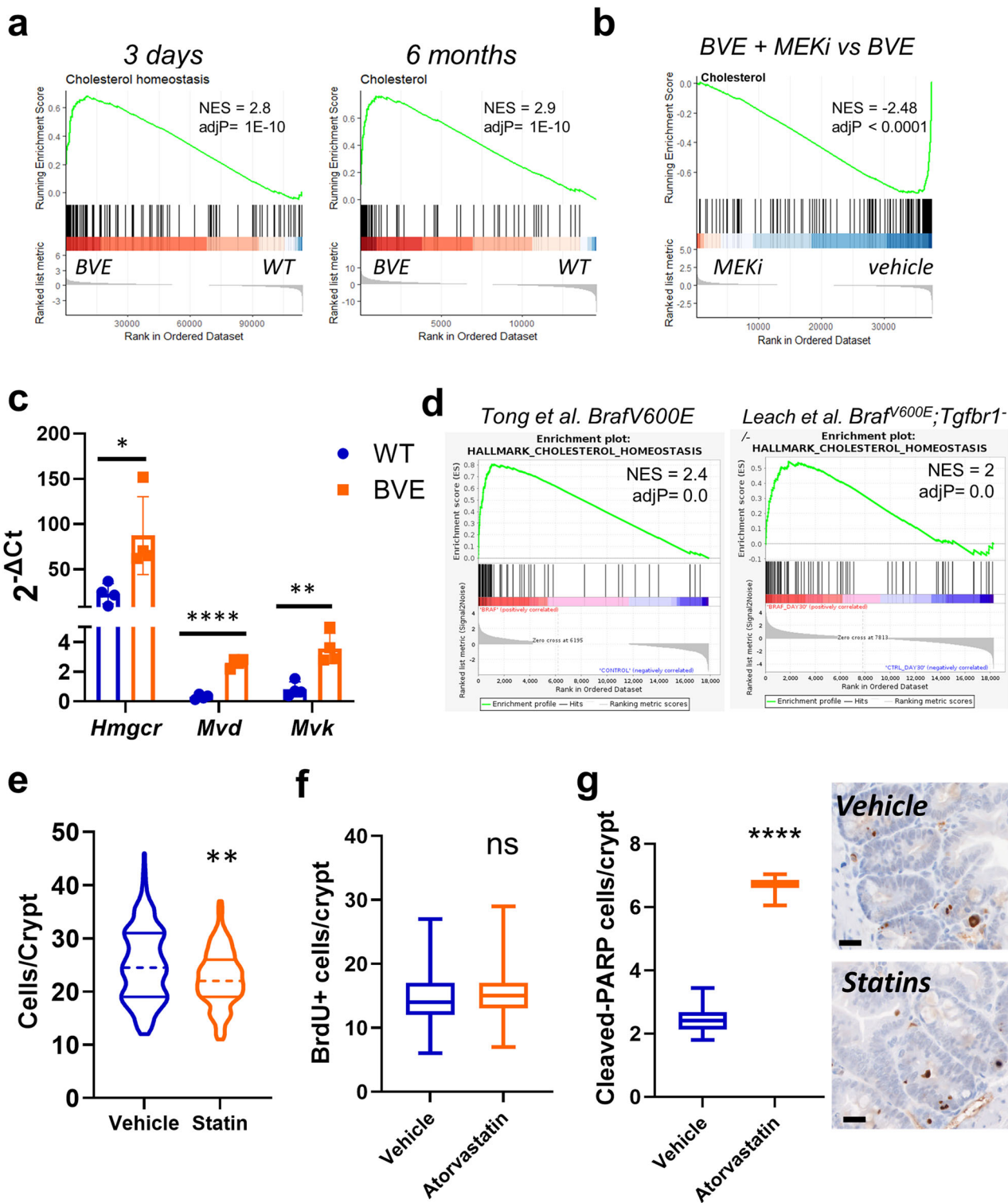
Cholesterol biosynthesis contributes to tumorigenesis and can be inhibited using statins, a widely prescribed category of drugs that target the mevalonate pathway rate-limiting enzyme HMG-CoA reductase. Several studies indicate that regular statin use reduces the incidence of several malignancies, including CRC<sup>28,29</sup>. Hence, we reasoned that cholesterol biosynthesis contributes to the establishment of the durable crypt hyperplasia observed following induction of *Braf*<sup>V600E</sup> and its inhibition by statins would prevent the formation of hyperplastic crypts. To test this hypothesis, we treated mice with daily doses of atorvastatin starting 1 week before induction of *Braf*<sup>V600E</sup> and collected tissue 3 days post tamoxifen administration. Counting of crypt epithelial cells confirmed that statin treatment prevented the establishment of crypt hyperplasia in mutant intestinal tissue (Fig. 4e and Supplementary Data 1). Mechanistically, atorvastatin did not alter crypt proliferation (Fig. 4f and Supplementary Data 1), but significantly increased apoptosis assessed through cleaved PARP immunohistochemistry (Fig. 4g and Supplementary Data 2).

Overall, these data indicate that *Braf*<sup>V600E</sup> elicits a transcriptional reprogramming of cellular metabolism indicative of increased cholesterol biosynthesis, which contributes to the establishment of tissue hyperplasia by increasing the survival of crypt cells. The results suggest that statins might prevent CRC harboring mutant *BRAF*<sup>V600E</sup>.

**Expression of *BRAF*<sup>V600E</sup> establishes a network of transcription factors that mediates the enrichment in the cholesterol biosynthesis gene signature in human serrated lesions.** Next, we investigated whether the enrichment of the cholesterol



**Fig. 3 Expression of *Braf<sup>V600E</sup>* in the intestinal epithelium induces persistent gene signatures related to CRC. **a** GSEA showing changes in signatures associated with *BRAF<sup>V600E</sup>*-driven CRC 3 days and 6 months following tamoxifen induction. **b** GSEA showing changes in the Wnt signature 3 days and 6 months following tamoxifen induction. **c** GSEA showing enrichment in a poor-prognosis gene signature for *BRAF<sup>V600E</sup>*-drive CRC 6 months following tamoxifen induction. **d** GSEA showing dependence on MAPK pathway for the establishment of key gene signatures. Mice expressing *Braf<sup>V600E</sup>* for 3 days were treated with vehicle control or the MEK inhibitor PD184352 (MEKi). NES normalized enrichment score.**



metabolism gene signature induced by *Braf<sup>V600E</sup>* in the mouse intestine was preserved in human serrated lesions. To this end, we analyzed three independent datasets of serrated lesions, which enabled comparison between serrated lesions, traditional adenomas and normal right colon tissue<sup>30–32</sup>. These analyses showed a robust enrichment in the cholesterol biosynthesis signature when serrated lesions were compared to either normal tissue or adenomas (Fig. 5a, b). Moreover, we confirmed in two independent datasets that CRCs harboring mutant *BRAF<sup>V600E</sup>* display an enrichment of the cholesterol gene signature compared to WT

cancers (Supplementary Fig. 13a)<sup>7</sup>. Notably, no significant enrichment was observed when adenomas were compared to normal tissue control, suggesting that the transcriptional rewiring of cholesterol metabolism is specific to serrated lesions (Supplementary Fig. 13b). Since bulk transcriptomic data cannot distinguish between the contributions by tumor or non-tumor cells to the enrichment of gene signatures, we ascertained expression of the cholesterol metabolism signature in transformed epithelial cells, by querying a scRNA-seq dataset of serrated lesions and normal tissue<sup>6</sup>. Clustering of the serrated lesions and normal

**Fig. 4** *Braf*<sup>V600E</sup> in the intestinal epithelium shows the enriched expression of a gene signature related to cholesterol metabolism. **a** GSEA showing changes in gene expression of the cholesterol biosynthesis signature 3 days and 6 months following tamoxifen induction. NES normalized enrichment score. **b** Treatment of mice with the MEK inhibitor PD184352 (MEKi) reverts the enrichment of the cholesterol biosynthesis gene signature in intestinal tissue. **c** rt-qPCR analysis of cholesterol metabolism genes in intestinal tissue collected 6 months after tamoxifen injection. Bars represent mean  $\pm$  SD. Dots represent single replicates. Data were analyzed by unpaired two-tailed *t* test ( $n = 4$  per group). \*\* $P \leq 0.01$ ; \*\*\* $P \leq 0.001$ . **d** GSEA showing increased transcriptional expression of cholesterol biosynthesis gene signature in the intestinal tissue of two independent *Braf*<sup>V600E</sup> mouse datasets. NES normalized enrichment score. **e** Violin plot showing the distribution of the number of cells per crypt in *Braf*-mutant mice induced for 3 days and treated with vehicle control or atorvastatin 10 mg/mL. Dotted and plain lines indicate median and quartiles, respectively. Data were analyzed by Kolmogorov-Smirnov test.  $n = 6$  female mice in the vehicle group and 3 female mice in the atorvastatin-treated group. In total, 50 crypts were counted per each animal. \*\* $P \leq 0.01$ . **f, g** Cell proliferation and apoptosis were evaluated in *Braf*-mutant intestinal crypts of mice treated with vehicle control or atorvastatin using IHC for BrdU and cleaved-PARP1, respectively. Examples of cleaved PARP positive crypt cells are provided (bar size = 20  $\mu$ m). Data are plotted as box and whisker plots. Boxes represent median and quartile values, whiskers represent min and max values. Data were analyzed using unpaired, two-tailed *t* test ( $n = 6$  vehicle and 3 atorvastatin-treated female animals). ns not significant, \*\*\*\* $P \leq 0.0001$ .

single cells revealed seven canonical cell types and one serrated specific cell (SSC) subtype<sup>6</sup> (Fig. 5c and Supplementary Fig. 14a). Notably when the expression of the cholesterol biosynthesis signatures was scored in the cell clusters using the UCell tool<sup>33</sup>, SSCs exhibited significantly higher expression of the signature, compared with other cell types (Fig. 5d, e and Supplementary Data 1), including cancer cells from conventional adenomas (Supplementary Fig. 14b). Key metabolic enzymes of the mevalonate pathway were expressed in SSCs, some of them showing a significant enrichment compared to non-cancerous cells (Supplementary Fig. 14c).

These data indicate that the transcriptional metabolic rewiring of cholesterol metabolism observed in *Braf*<sup>V600E</sup> mouse intestinal tissue recapitulates comparable adaptations of human colorectal lesions and is a feature of serrated lesions.

To investigate how *BRAF*<sup>V600E</sup> increases transcription of the cholesterol gene signature in human serrated lesions, we first applied the SCENIC method<sup>34</sup> to infer transcription factors and gene regulatory networks from scRNA-seq data of serrated lesions and normal cells. By doing so, we identified regulons that are specific to annotated cell types based on RSS<sup>35</sup> (Supplementary Fig. 15). Next, we employed iRegulon<sup>36</sup> to pinpoint which transcription factors active in SSCs were predicted to regulate the differentially expressed genes within the cholesterol gene signature (cut off  $\geq 5$  target genes). We singled out six transcription factors, namely *CEBPB*, *CREB3*, *NR2F1*, *KLF16*, *SP6*, and *FOSL1*, whose regulons are strongly enriched in SSCs (Fig. 6a) and which contribute towards the establishment of the cholesterol metabolism signature (Fig. 6b).

## Discussion

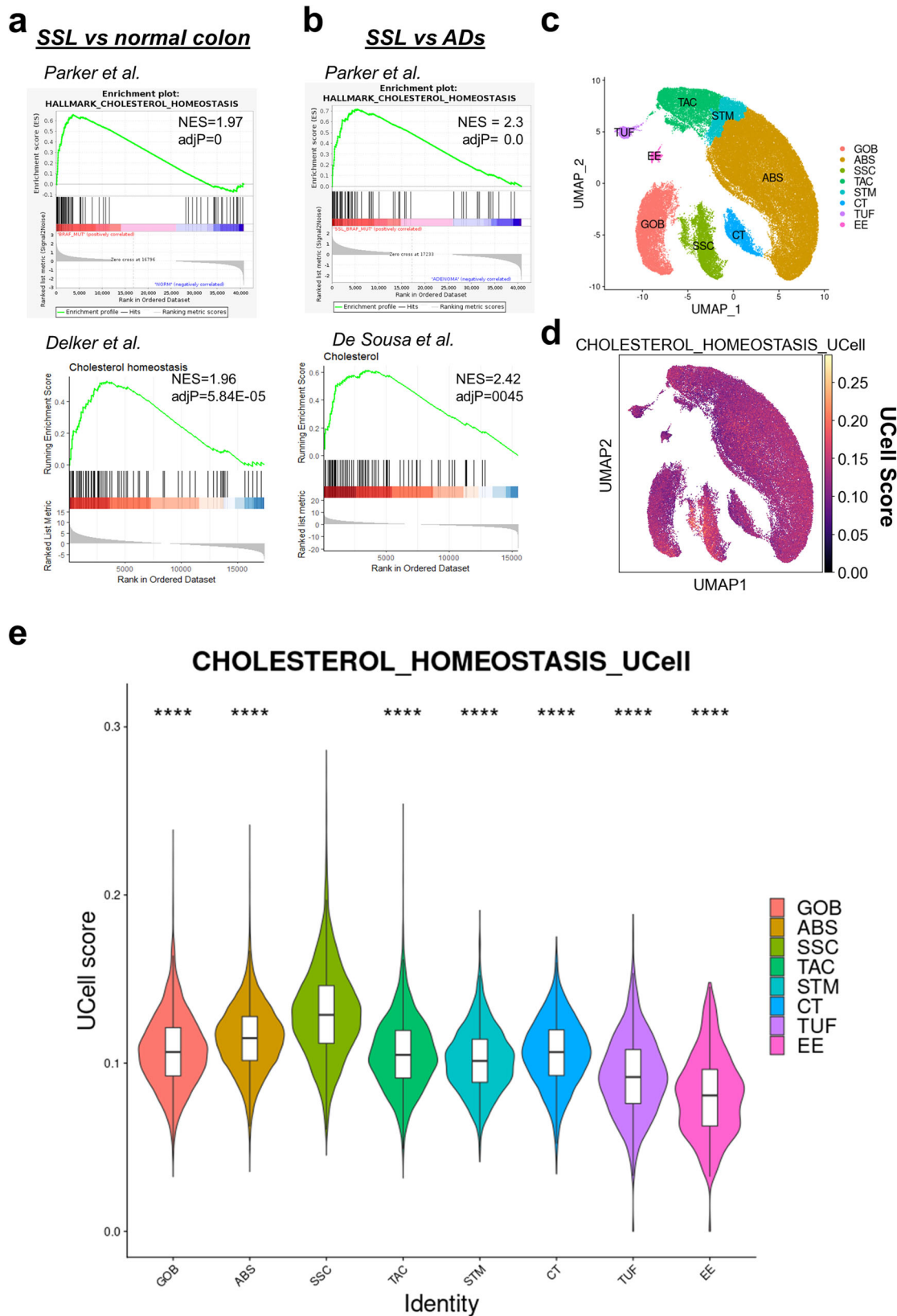
Ten percent of human CRCs, which originate from serrated lesions, harbor the V600E mutation in the *BRAF* oncogene<sup>13</sup>. This mutation is thought to contribute to tumorigenesis through the activation of the MEK/ERK arm of the MAPK pathway. Yet, despite the development of CRC being a decade-long process, the long-term impact of oncogene expression on the intestinal epithelium has not been investigated. Here, we assessed this issue using a genetically modified mouse model, the *VillinCre*<sup>ER/0</sup>/*Braf*<sup>V600E-LSL/+</sup> (BVE), which enables the tamoxifen-inducible expression of mutant *Braf*<sup>V600E</sup> in the intestinal epithelium. We used a combination of transcriptomic and histological analyses to compare short-term and long-term impacts of *Braf*<sup>V600E</sup> expression. By doing so, we were able to identify changes in intestinal epithelial populations that were evident only after prolonged oncogene expression. Interestingly, despite failing to observe any loss of ISCs as reported by others<sup>18,22</sup>, we did find downregulation of the ISC signature early after *Braf*<sup>V600E</sup> induction. However, ISC signature was recovered and downregulation reverted after 6 months. Although it is currently unclear what enables this adaptation in the BVE mice,

these changes could be explained by the well-established plasticity of the intestinal epithelial cells<sup>37</sup>. A striking alteration of Paneth cells was evident in 6-week and 6-month-induced animals. Paneth cells lost their localization at the base of the intestinal crypt and migrated towards the villi, where they displayed a rounded morphology and were similar to the intermediary cells expressing both goblet and Paneth cell markers described in ref. 22. A similar phenotype was reported in other mouse models with intestinal expression of *Braf*<sup>V600E</sup>, and a MAPK-dependent loss of Paneth cells was also described in models of mutant *Kras*<sup>18,23,38</sup>. Consistent with these observations, analysis of intestinal tissue after long-term *Braf*<sup>V600E</sup> activation revealed an impact on other secretory lineages. We observed changes in the numbers of both goblet cells and EECs. However, it should be noted that this reduction does not reflect the total number of cells. Indeed, since villi length was increased in mutant mice, cell numbers were normalized to villi length. Notwithstanding, the reduction in RNA levels of *ChgA* confirmed a general reduction in EECs. Whereas the increase in goblet cells probably reflects the mucinous nature of serrated neoplasia, the changes in EECs are particularly intriguing. Indeed, human *BRAF*<sup>V600E</sup> CRCs accumulate EEC progenitor cells that fail to complete their differentiation. These progenitor EECs secrete the TFF3 protein, a member of the Trefoil family, which supports CRC progression through activation of the PI3K/AKT pathway<sup>39</sup>. These data warrant additional research into the impact of *BRAF* mutations on the differentiation and activity of EECs in the context of serrated neoplasia.

Transcriptomic analysis through GSEA confirmed a negative enrichment of an intestinal differentiation signature after prolonged *Braf*<sup>V600E</sup> expression. This was likely driven by a significant downregulation of an enterocyte differentiation signature, which, to the best of our knowledge, has not been reported previously. The possible reduction in enterocytes was confirmed by reduced Alpl staining in intestinal tissue. Overall, our data suggest that the long-term impact of *BRAF*<sup>V600E</sup> is associated with a reduction in tissue differentiation, perhaps necessary to rebalance the initial decrease in stemness<sup>18,22</sup>. These data are in agreement with published evidence indicating that *BRAF*<sup>V600E</sup> suppresses hallmarks of intestinal differentiation, which can be restored upon pharmacological inhibition of the MAPK pathway<sup>40</sup>.

The activation of *Braf*<sup>V600E</sup> led to a rapid and persistent enrichment in MAPK-dependent gene signatures associated with proliferation and CRC progression, including a fetal-like gene signature, a Myc signature, a G2/M signature and an E2F target gene signature. The latter two have been described as significant features of the BM2 molecular subtype of CRC<sup>17</sup>. We also observed the induction of a Wnt signaling-related signature, which was promptly downregulated by inhibition of the MAPK pathway. The association between the Wnt pathway and *BRAF*-driven CRC is still unclear. In mice, Rad and colleagues observed

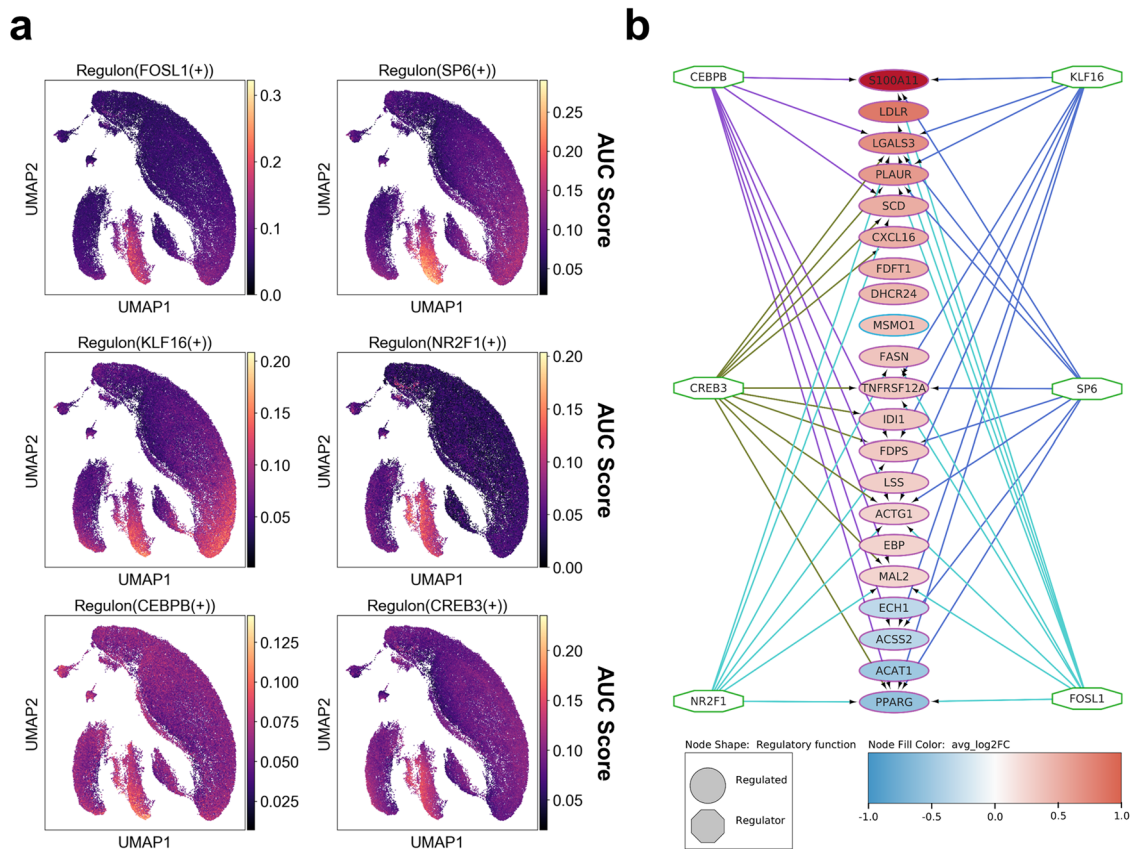




a progressive increase in Wnt activation during neoplasia progression<sup>19</sup>, which was associated with nonsense or frameshift mutations in the *Apc* gene and activating mutations in the *Ctnnb1* gene, which encodes  $\beta$ -catenin. Mutations affecting Wnt-related genes have also been reported in human serrated CRCs. Truncating mutations of the Wnt negative regulator gene *RNF43* are

common in MSI CRCs<sup>41</sup>, and non-truncating mutations in the *APC* gene have also been detected<sup>42</sup>. However, the contribution of Wnt signaling to MSI CRCs remains controversial with some studies reporting lack of aberrant nuclear  $\beta$ -catenin localization<sup>43</sup>, a surrogate marker for Wnt pathway activation, but others showing significant nuclear accumulation in more advanced

**Fig. 5 Gene signatures of cholesterol biosynthesis are enriched in transcriptomic datasets of human colorectal neoplasia.** **a, b** GSEA showing increased expression of the cholesterol gene signature in human serrated lesions (SSL) from three independent datasets comparing serrated neoplasia to normal matched control (**a**) or to adenomatous lesions (AD). NES normalized enrichment score. (**b**). **c** UMAP visualization of cells from serrated and normal datasets colored by annotated cell type. ABS absorptive cells, CT crypt top colonocytes, EE enteroendocrine cells, GOB goblet cells, SSC serrated specific cells, STM stem cells, TAC transient amplifying cells, TUF tuft cells. **d** UCell score distribution for the cholesterol gene signature shown in UMAP space. **e** Violin plot showing a comparison of UCell score distribution of a cholesterol gene signature in different cell populations. The boxes represent median and interquartile ranges. The whiskers show 95% confidence interval. The distribution of UCell score in SSC population was compared with the remaining cell lineages using the Wilcoxon statistical test. \*\*\*\* $P \leq 0.0001$ .



**Fig. 6 Analysis of regulons associated with the transcription of the cholesterol metabolism genes in a human single-cell transcriptomic dataset.** **a** Regulon activity projected in UMAP space for the regulons associated with transcription factors (FOSL1, SP6, KLF16, NR2F1, CEBPB, CREB3) involved in the establishment of the cholesterol gene signature in the SSCs. **b** A network diagram generated with iRegulon in Cytoscape using the differentially expressed gene of the cholesterol biosynthesis signature in SSC identified by Seurat as an input. Circles indicate 21 significant genes, colored by avg\_log2FC. Octagons symbolize selected TFs (cut off >5 target genes) involved in the regulation of DEGs in the cholesterol biosynthesis signature. The edges represent connections between each of six transcription factors and their target genes and are colored based on the source of integration.

serrated lesions<sup>42,44</sup>. Notwithstanding, GSEA in BVE mice indicates that, despite the lack of neoplastic lesions, expression of *Braf*<sup>V600E</sup> triggers a transcriptional reprogramming congruent with tumorigenesis.

Metabolic rewiring is a common hallmark of cancer, including CRC<sup>45–47</sup> and we have recently, reported on the role of amino acid metabolism in CRC progression<sup>48,49</sup>. Here, we provide data that strongly support the establishment of a transcriptionally-driven metabolic reprogramming of cholesterol biosynthesis following activation of *Braf*<sup>V600E</sup>. We have also confirmed similar changes in human datasets of serrated lesions, suggesting that this metabolic adaptation is a feature of serrated neoplasia. Using bioinformatics tools SCENIC and iRegulon, we have finally identified the chief transcription factors that are responsible for the transcriptional rewiring: *CEBPB*, *CREB3*, *NR2F1*, *KLF16*, *SP6*, and *FOSL1*. The transcription factor *CREB3* has been linked to regulation of cholesterol in response to the unfolded protein

response<sup>50</sup>. Similarly, the orphan nuclear receptors NR2F1 and 2 (also known as chicken ovalbumin upstream-promoter transcription factor, COUP-TFI and II) have been reported to regulate genes associated with cholesterol metabolism<sup>51–53</sup>. A role for *FOSL1* in promoting cholangiosarcoma growth through regulation of HMG-CS1 along the mevalonate pathway has also been described<sup>54</sup>. Our data warrant further investigation to dissect the contribution of these transcription factors to the regulation of metabolic rewiring in *BRAF*-mutant CRCs and details of their regulation by MAPK. Importantly, statins have been suggested to possess cancer-preventive properties, but the association between the occurrence of adenoma or CRC and exposure to statins is controversial<sup>28,55–57</sup>. Very little is known about whether protection by statins depends on the specific molecular makeup of the tumor<sup>58</sup>. Notwithstanding, here, we show that treatment with atorvastatin prevents the establishment of crypt hyperplasia through induction of apoptosis in intestinal epithelial cells, but

without any evident impact on cell proliferation. However, it remains to be established how inhibition of cholesterol metabolism promotes apoptosis and how this relates to the presence of mutations in the *BRAF* oncogene. Moreover, our treatment study did not include non-induced mice and, therefore, we cannot formally rule out an effect of atorvastatin on normal intestinal epithelial cells. However, despite the inhibition of the mevalonate pathway displays some degree of unselective growth-suppressive function, tumor cells are more responsive than normal cells to pathway inhibition<sup>59,60</sup>. It is also noteworthy that atorvastatin treatment of *BRAF*-mutant mice reduces crypt cell numbers to levels comparable to WT animals, suggesting that statins likely revert the increased proliferation caused by the expression of mutant *BRAF*. The sensitivity of *BRAF*-mutant CRC cell lines to statin treatment has also been suggested to depend on BMP signaling through a functional *SMAD4* gene<sup>61</sup>. BMP signaling belongs to the TGF- $\beta$  superfamily and mutations in the TGF- $\beta$  pathway are common in *BRAF*-mutant CRCs. Whether this affects sensitivity to statins remains to be ascertained. Nonetheless, we observed a similar behavior in the response of *BRAF*-mutant CRC cells to atorvastatin in vitro: RKO cells with intact *SMAD4* showed sensitivity to statin treatment, whereas HT29 cell harboring homozygous mutations in *SMAD4* were resistant (Supplementary Fig. 16). These observations are in agreement with the evidence provided here that atorvastatin prevents crypt hyperplasia in the *Braf*-mutant intestine of BVE mice, which harbor a functional BMP pathway.

Overall, our data show that one of the most significant and consistent effects of *BRAF*<sup>V600E</sup> expression in mouse tissue and human serrated lesions is the enrichment of a gene expression signature of cholesterol biosynthesis that contributes to the establishment of epithelial hyperplasia. These findings suggest that the incidence of *BRAF*-mutant CRC could be reduced using statins and warrant further targeted investigations into the specific statin sensitivity of this subtype of CRC.

Despite mouse models are suitable tools for recapitulating the development of CRC in vivo<sup>62</sup>, they harbor defined genetic alterations within an inbred genetic background and carefully controlled, pathogen-free housing conditions, and thus do not fully recapitulate the complex genetic heterogeneity and micro-environmental complexity encountered in human tumors. In addition, a proper evaluation of the preventive activity of statins towards *BRAF*-mutant CRC would necessitate the assessment of tumor development in the *VillinCre*<sup>ER/0</sup>/*Braf*<sup>V600E-LSL/+</sup> or other models of *BRAF*-driven colorectal carcinogenesis<sup>11</sup>. It should also be acknowledged that the influence of additional genetic alterations (such as mutation in *SMAD4*)<sup>61,63</sup> or/and environmental factors (such as inflammation or the microbiome)<sup>64</sup> could influence and perhaps overcome any protective effect of statins or alter tissue response to expression of the *BRAF* oncogene. A thorough analysis of statin anti-cancer activity in diverse experimental conditions is therefore warranted.

## Methods

**Cell lines and MTT assay.** HT29 CRC cells were grown in DMEM supplemented with 10% fetal calf serum (FCS) and Glutamax. RKO CRC cells were grown in MEM supplemented with 10% FCS and Glutamax. Cells were incubated at 37 °C and 5% CO<sub>2</sub>. For MTT assay, 500–2000 cells were plated in 96-well plates. A 50 mM atorvastatin in DMSO stock solution was diluted in media (supplemented with dialyzed serum) to concentrations of 1, 5, and 10  $\mu$ M. Three 96-well plates were run for each assay, each treated for either 24, 48, or 72 h. Cells were then incubated with MTT (2.5 mg/ml) at 37 °C 2 h. To dissolve the formazan produced by the metabolism of MTT, 50  $\mu$ l of DMSO was added

to each well. The plates were then incubated for 30 min on a shaker before the optical density was read using a FluoStar OPTIMA plate reader (BMG-Labtech Ltd, UK).

**Mouse colony and genotyping.** All animal experiments were performed according to Home Office guidelines under project licenses (PPL) PC4E1710A and P7B8067BB. Experiments received ethical approval from the local ethics committee at the University of Leicester. All mice were C57BL6/j background aged between 8 weeks and 6 months. Unless otherwise stated in this section, figure legends and Supplementary Data 1, an equal number of males and females were used for experiments. Mice were housed in the pathogen-free Preclinical Research Facility (PRF) at the University of Leicester and were fed ad libitum with AIN93 diet (TestDiet, USA, Cat#5801-G), under a climate-controlled environment with 12 h day/night cycle. For transgene induction, 8–12-week-old double heterozygous *VillinCre*<sup>ER/0</sup>/*Braf*<sup>V600E-LSL/+</sup> mice received intraperitoneal injections of 1 mg tamoxifen on 5 consecutive days at 24 h intervals, unless otherwise stated. Tamoxifen solution was prepared as a 10 mg/ml stock solution solubilized in corn oil. Control mice were *VillinCre*<sup>ER/0</sup>/*Braf*<sup>+/+</sup> which received tamoxifen, or *VillinCre*<sup>ER/0</sup>/*Braf*<sup>V600E-LSL/+</sup> treated with vehicle. After euthanasia, the small intestine was flushed with PBS and cut into six sections rolled to create a Swiss roll, and fixed in 4% [w/v] paraformaldehyde on a shaker overnight at room temperature. Fixed tissues were placed in embedding cassettes and submerged in 70% ethanol and processed by the Core Biotechnology Services (CBS) histology facility, the University of Leicester. Small pieces of tissue (~0.5 cm) were snap-frozen in liquid nitrogen for protein extraction or snap-frozen in RNAlater solution (Sigma) for RNA extraction.

DNA from mouse ear snips was extracted using the DNeasy Blood & Tissue Kit (QIAGEN) according to the manufacturer's instructions. PCR reaction mix for the *Braf* locus was: 10.5  $\mu$ l of template DNA, OCP 125 (FWD primer) 1  $\mu$ l (10 pmol), OCP 137 (REV-HET primer), 0.5  $\mu$ l (5 pmol) OCP 143 (REV-WT primer) 0.5  $\mu$ l (5 pmol), MyTaq Red (2 $\times$ ) (Meridian Bioscience BIO-25043) 12.5  $\mu$ l. The PCR program was: 95 °C for 2 min, 30 cycles of 94 °C for 15 s, 60 °C for 15 s, 72 °C for 15 s; 5 min at 72 °C. For *Cre*<sup>ER</sup> genotyping the PCR mix was: 10.5  $\mu$ l of template DNA, OCP 361 (FWD primer) 1  $\mu$ l (10 pmol), OCP 362 (REV primer) 1  $\mu$ l (10 pmol), MyTaq Red (2 $\times$ ) (Meridian Bioscience BIO-25043) 12.5  $\mu$ l. The PCR program was: 95 °C for 4 min, 30 cycles of 94 °C for 20 s, 57 °C for 20 s, 72 °C for 25 s; 5 min at 72 °C. Primer (Sigma) sequences are listed in Supplementary Table 1. The PCR products were visualized using 1.8% agarose gel with ethidium bromide (0.5  $\mu$ g/mL) and imaged using the Syngene G:BOX Chemi XR5. Gels were examined for the presence of PCR bands at 140 bp (*Braf*<sup>LSL-V600E</sup> allele), 466 bp (wild-type *Braf* allele), and 300 bp (*Villin-CreERT* allele).

**Statin treatment in vivo.** Thirteen-week-old female *VillinCre*<sup>ER/0</sup>/*Braf*<sup>V600E-LSL/+</sup> mice received either 10 mg/kg atorvastatin by oral gavage or 5% DMSO/PBS vehicle control daily for 12 days. A 20 mg/mL atorvastatin (Generon) stock in DMSO was freshly diluted to a 1 mg/mL working solution in PBS. Days 1–7 of the study were a daily pre-treatment period of atorvastatin/vehicle. Days 8–10 consisted of both atorvastatin/vehicle gavage and intraperitoneal tamoxifen injection to induce *Braf*<sup>V600E</sup> expression. On the final day of the study at 3 days post *Braf*<sup>V600E</sup> induction, mice received intraperitoneal injections of 200  $\mu$ l of 20 mM BrdU solution (Sigma) 3 h before tissue harvesting.

**Histology.** Tissue sections (4  $\mu$ m), and hematoxylin and eosin (H&E) staining were performed by the Histology Facility, Core

Biotechnological Services, University of Leicester. Immunohistochemistry was performed using the Novolink Polymer Detection Systems kit (Leica Biosystems), according to the manufacturer's instructions. SuperFrost Plus™ Adhesion slides (Fisher Scientific) containing FFPE tissue sections were incubated at 65 °C for 30 min. Tissues were dewaxed and rehydrated through 3-min serial immersions in xylene, 99% industrial methylated spirit (IMS), and 95% IMS twice each, followed by 5 min in running tap water. Following antigen retrieval (Supplementary Table 2), slides were incubated for 5 min in a peroxidase block, washed twice in PBS, incubated for 5 min in the protein block solution and washed twice in PBS. Tissues were incubated with primary antibodies (Supplementary Table 2) (3% BSA/PBS) overnight at 4 °C. After incubation, sections were incubated in Polymer solution for 30 min, washed in PBS, and incubated in DAB solution for 5 min. Tissues were counterstained in Mayer's hematoxylin for 5 min and subsequently washed in tap water. Tissues were dehydrated through 3-min serial immersions in 95% IMS twice, 99% IMS twice, xylene twice. Slides were then mounted onto glass coverslips using DPX (Sigma Aldrich). Whole slide images were acquired using the Nanozoomer XR digital slide scanner (Hamamatsu Photonics) and visualized using NDP.view2 software. For immunofluorescence (IF), antigen-retrieved tissues were blocked in 5% BSA/0.3% Triton X-100 in PBS for 30 min, washed with 1% BSA/0.3% Triton X-100 in PBS twice for 10 min each and incubated with SP-1 Chromogranin A (ChgA) primary antibody (ImmunoStar) overnight at 4 °C. Slides were washed the next day in wash buffer solution twice for 10 min each. Alexa Fluor 568 secondary antibody (Invitrogen) was incubated for 1 h at RT in darkness. Slides were washed and sealed with coverslips using ProLong™ Glass Antifade Mountant containing DAPI (Invitrogen™). IF images were captured using the Vectra® Polaris™ imaging system (Akoya Biosciences), and images were viewed using Phenochart software v1.1.0 (Akoya Biosciences).

Alcian blue staining for visualization of goblet cells was performed using the Alcian Blue (pH 2.5) Stain Kit (Vector Laboratories). Rehydrated slides were heated at 65 °C and incubated in 3% acetic acid solution for 3 min, followed by incubation with Alcian Blue solution (pH 2.5) for 30 min at RT. Sections were rinsed in 3% acetic acid solution, running tap water, and distilled water. Sections were then counterstained with Enhanced Nuclear Fast Red (Vector Laboratories) for 5 min, rinsed in running tap water and distilled water, and dehydrated for mounting.

Intestinal alkaline phosphatase staining for the detection of enterocytes was performed using Vector Red alkaline phosphatase staining kit (Vector Laboratories). Rehydrated slides were incubated in substrate working solution for 30 min and washed in running water for 5 min. Tissues were counterstained in Mayer's hematoxylin (Leica Biosystems), dehydrated, and mounted in DPX.

**Electron microscopy (EM).** All EM reagents were from Agar Scientific unless otherwise stated. For Transmission, Electron Microscope material was fixed overnight at 4 °C in 2.5% glutaraldehyde and 4% paraformaldehyde in 0.1 M cacodylate buffer, pH 7.2, and washed in 0.1 M cacodylate buffer. After the secondary fixation in 1% osmium tetroxide and 1.5% potassium ferricyanide (Merck Life Science) at RT, samples were treated with 1% tannic acid (VWR) dehydrated in ethanol series followed by propylene oxide (Merck Life Science) and embedded in Epon 812 (TAAB Laboratory Equipment Ltd). Samples were sectioned to 70-nm thick using a Reichert Ultracut E ultramicrotome, collected onto copper mesh grids, and stained for 5 min in lead

citrate. Samples were viewed on a JEOL JEM-1400 TEM with an accelerating voltage of 120 kV. Digital images were collected with an EMSIS Xarosa digital camera with Radius software.

**Western blotting.** Tissues were homogenized in 400 µl of ice-cold RIPA buffer, containing protease inhibitor cocktail, and phosphatase inhibitor cocktail (Roche). The homogenized solution was kept on ice for 10 min with vortexing every minute. Samples were then centrifuged for 10 min at 12,000 rpm, at 4 °C. Supernatants were collected and protein concentration estimated with Lowry assay. Protein lysates were diluted to 1 mg/ml with SDS loading buffer (0.0625 M Tris-HCl pH 6.8, 2% SDS, 0.4 ml of 0.025% Bromophenol blue, 20% glycerol, 5% β-mercaptoethanol) and incubated at 95 °C for 10 min. In total, 15 µl of each sample was then run on acrylamide gels (Miniprotean III cell, BioRad) with pre-stained SDS-PAGE protein markers (All Blue Precision Plus Protein Standards, BioRad). Samples were then transferred to nitrocellulose membranes (Syngene) using a semi-dry transfer blot (BioRad) as per the manufacturer's instructions. For p-ERK1/2 (Rabbit, Cell Signaling, 9101 S), membranes were blocked in 5% BSA (Sigma Aldrich, A3059) in Tris-Buffered Saline, 0.1% Tween (TBST) for 90 min followed by incubation with primary antibody diluted 1:500 in 5% BSA in TBST overnight at 4 °C. For ERK2 (Mouse, Santa Cruz, sc-1647) membranes were blocked in 5% milk in TBST and incubated with the primary antibody diluted 1:1000 in 5% milk in TBST overnight at 4 °C. Membranes were washed in TBST, three times for 10 min each, and then incubated with secondary antibody (1:5000) at room temperature for 1 h (anti-Rabbit HRP Sigma A6154, anti-Mouse HRP Sigma A4416) followed by TBST washes three times. Bands were visualized using SuperSignal West Pico Chemiluminescent substrate kit (Thermo Fisher Scientific, 34580) and photographic film (Fuji).

**RNA extraction.** RNA was extracted from intestinal tissue on a Promega Maxwell 16 using the Maxwell 16 LEV simplyRNA Cells Kit (Reference AS1270). RNA quality and quantity was assessed using the 2100 Bioanalyzer Instrument (Agilent). The same RNA samples were used for RNA sequencing and Real-time PCR at 6 months.

#### Transcriptomic analysis

**RNA sequencing (RNA-seq).** For RNA-seq, four *VillinCre<sup>ER/0</sup>/Braf<sup>V600E-LSL/+</sup>* 8–12-week-old mice were induced with tamoxifen and four were treated with vehicle to be left uninduced and used as controls to identify differentially expressed (DE) genes. Each experimental group contained two male and two female animals. Mice received intraperitoneal injections of 2 mg tamoxifen on 2 consecutive days. Intestinal tissue was harvested 6 months post induction. Indexed RNA libraries were prepared using the NEBNext® Single Cell/Low Input RNA Library Prep Kit for Illumina (NEB ref # E6420S) according to manufacturer's standard protocol and sequenced by 75 bp-paired-end sequencing on Illumina Novaseq, with the aim to obtain approximately 25 million reads per sample.

**Microarrays.** For microarrays, 12 mice 8–12 weeks old were split into two groups of six with an equal number of males and females. Six mice were *VillinCre<sup>ER/0</sup>* controls, and six were *VillinCre<sup>ER/0</sup>/Braf<sup>V600E-LSL/+</sup>*. Both groups received 3 consecutive daily doses of 2 mg tamoxifen, and tissue was harvested 3 days after the last injection. Animals subjected to inhibition of MEK kinase with PD184352 treatment received two injections of 100 µl tamoxifen (10 mg/ml) on 2 consecutive days, followed by three injections of inhibitor/vehicle on 3 consecutive days. PD184352

(Selleckchem) was prepared in 10% DMSO (Sigma), 10% Crephosphor EL (Calbiochem) in water. RNA was extracted as described above, and transcriptome profiling was conducted using the SurePrint G3 Mouse Gene Expression v2 8x60K Microarray Kit (Agilent, #G4852B), according to the manufacturer's instructions.

**Real-time PCR (rt-qPCR).** Reverse transcription of RNA was performed using the High-Capacity RNA-to-cDNA™ Kit (Applied Biosystems™) with a random primer approach. The 2X master mix was prepared according to the manufacturer's instructions, and 2 µg of purified RNA was added per 20 µL reaction. Reactions with and without the reverse transcriptase enzyme were performed in tandem, and samples without the enzyme were used as negative controls in subsequent rt-qPCR assays. The reverse transcription reaction was performed using the program: 25 °C for 10 min, 37 °C for 120 min, 85 °C for 5 min.

TaqMan® Gene Expression Assays (Fisher Scientific) were used for rt-qPCR reactions (Supplementary Table 3), using the StepOnePlus™ rt-qPCR System (Fisher Scientific) equipped with StepOnePlus™ Software v2.3. The thermocycler program was: 50 °C for 2 min, 95 °C for 20 s, then 40 cycles of: 95 °C for 1 s, 60 °C for 20 s. Each PCR reaction contained 1 µL of template DNA/control/water, 5 µL of 2× TaqMan Fast Advanced Master Mix (Applied Biosystems) and a TaqMan Gene Expression Assay (Thermo Scientific) containing 900 nM forward and reverse primers and 250 nM probe and 10 µL with nuclease-free water. Each sample was run in technical triplicates. rt-qPCR data were then analyzed using the  $\Delta\Delta C_t$  method. Ct values were normalized to endogenous control (EC) genes. We employed the NormFinder Excel Add-In, v0.953<sup>65</sup> to assess four candidate genes for use as endogenous controls: glucuronidase beta (*Gusb*), beta-2-microglobulin (*B2m*), beta-actin (*Actb*), and TATA-box binding protein (*Tbp*). The top three genes (*Gusb*, *B2m*, *Tbp*) with the most stable expression across experimental groups were selected as EC genes for the normalization of RT-qPCR assays in accordance with the MIQE guidelines<sup>66</sup>.

**Bioinformatics analysis.** Bioinformatics analysis was performed using the statistical computing software R version 4.0.5 and the integrated development environment program RStudio version 1.3.1093. For RNA sequencing, FASTQ files sequencing quality was assessed using FastQC (v0.11.9). Sequencing adapter content was trimmed using FastP (v0.20.0)<sup>67</sup>. The pre-processed reads were then aligned to the mouse GRCm38/mm10 reference genome using Spliced Transcripts Alignment to a Reference (STAR) (v2.7.3a)<sup>68</sup>. Gene mapping and quantification were performed using DESeq2 (v1.34.0)<sup>69</sup>. For microarray data, the R Bioconductor package Linear Models for Microarray Data (Limma) (v3.50.1) was used to read raw text files from single-channel Agilent RNA microarrays and perform differential expression analysis.

**Gene set enrichment analysis (GSEA).** Agilent gene identifiers (IDs) from microarray analysis and RefSeq gene IDs from RNA sequencing were converted into Entrez IDs using BioMart (v2.50.3)<sup>70</sup> to ensure gene IDs were compatible with GSEA functions and The DE genes were ranked using Wald statistic. The R Bioconductor package clusterProfiler<sup>71</sup> was used to perform GSEA and statistical testing using Hallmark gene signatures from the Molecular Signatures Database (MSigDB)<sup>72</sup>. The *P* values were adjusted for multiple-hypothesis testing using the Benjamini–Hochberg method. Enrichment plots were generated using R Bioconductor package enrich plot (v1.14.2).

**Analysis of publicly available datasets.** Publicly available datasets (GSE106330<sup>18</sup>, GSE168478<sup>11</sup>, E-MTAB-6951<sup>32</sup>, GSE45270<sup>30</sup>, syn26720761 and TCGA-COAD<sup>8</sup>), were obtained from Gene Expression Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo/>), ArrayExpress (<https://www.ebi.ac.uk/arrayexpress/>), synapse (<https://www.synapse.org/>) and The Cancer Genome Atlas (TCGA) (<https://www.cancer.gov/tcga>). For GSE106330, normalized gene expression matrix was retrieved from GEO. For GSE168478, matrix of raw counts was obtained and further normalized using DESeq2 (v1.36.0) R package. Fastq files were downloaded from ArrayExpress for E-MTAB-6951. Reads, after quality control with fastQC (v0.11.5), were aligned to the human reference genome GRCh38 and quantified using STAR (v2.7.9a) with the following parameters: --quantMode GeneCounts, --outFilterScoreMinOverLread 0.33, --outFilterMatchNmin 40, --outFilterMatchNminOverLread 0.33. Count files were loaded into R and normalized using Deseq2 package. Expression CEL files were downloaded from GSE45270 and loaded into R using read.celfiles() function from affy R package (v1.74.0)<sup>73</sup>. Background correction normalization and expression estimation was done with rma() function. For syn26720761, processed TPM matrix was retrieved from synapse. For TCGA-COAD, the samples were downloaded in fastq format using GDC data transfer tool gdc client (v1.6.1)<sup>74</sup>. The quality of fastq files were analyzed with fastQC and samples, which failed quality control were removed. Universal Unique Identifier (UUID) of remaining samples was used to access TCGA-COAD STAR count matrices with TCGAbiolinks R package (v2.24.3)<sup>75</sup> and normalized with DESeq2 package. In total, 41 samples from normal tissue and 271 CRC samples having BRAF mutation status information (35 BRAF<sup>V600E</sup>; 236 WT) were analyzed.

GSEA of analyzed public datasets was performed with GSEA software (v4.2.2)<sup>76</sup>. All matrixes, generated as described above, were saved in GCT format. The categorical class file format (cls) was created for each dataset to define phenotype labels. GSEA was performed using default settings except for the Permutation type set to gene\_set.

### Single-cell RNA sequencing (scRNA-seq) data analysis

*scRNA-seq, data subset, and integration.* Publicly available, scRNA-seq QC-filtered (Level 4) discovery (DIS) and validation (VAL) epithelial datasets were downloaded from HTAN data portal: <https://data.humantumoratlas.org/> (HTAN Vanderbilt)<sup>6</sup>. To subset cells of interest, datasets were loaded into Python (v3.9.1) as AnnData objects (v0.7.6)<sup>77</sup>. Both datasets contained metadata attributes, such as polyp type and assigned cell type. This information was used to subset cells annotated as sessile serrated lesion and normal colonic biopsy with the *isin* function of the pandas (v1.3.1) package (Reback et al.<sup>78</sup>). In total, 64,382 normal cells (VAL: 34,008, DIS: 30,374) and 12,986 serrated cells (VAL: 6892, DIS: 6094) were subsetted for downstream analysis. Subsets of cells were loaded into R (v3.6.1) to perform integration with Seurat (v3.2.3) R package<sup>79</sup>. Prior to integration, datasets were independently log-normalized with the *NormalisationData* function. 2000 features being highly variable for each dataset and across datasets were identified using *FindVariableFeatures* and *SelectingIntegrationFeatures* functions, respectively. A *FindIntegrationAnchors* function with default settings was implemented to perform the unsupervised identification of cells that were used to integrate DIS and VAL datasets together by applying *IntegrateData* function. Further bioinformatics analysis was performed in Seurat (v4.1.1)<sup>80</sup>.

*scRNA-seq, dimensional reduction, and clustering.* Principal component analysis (PCA) was performed on normalized and scaled data by implementing *RunPCA* function with a feature

argument set to use variable features. The first 10 dimensions of the PCA were used to calculate the neighbors with *FindNeighbors* function. Finally, the graph was segmented into clusters using *FindCluster* function with a resolution adjusted to 0.5. Uniform Manifold Approximation and Projection (UMAP), a non-linear dimensional reduction method, was selected to explore an integrated dataset by implementing *RunUMAP* function with a number of dimensions specified as 10. *DimPlot* function was further used to visualize cells in a low-dimensional UMAP space.

*scRNA-seq, cell-type annotation.* The original cell-type annotation, which was included in the metadata, was projected in the UMAP space with *DimPlot* function and used to rename 39 identified clusters by implementing *RenameIdents* function. In total, eight cell types were annotated within an analyzed subset of serrated and normal cells.

*scRNA-seq, gene signature scores, UCell.* UCell (v1.3.1) R package<sup>33</sup> was applied to evaluate cholesterol homeostasis signature distribution in serrated lesion and normal dataset, as well as in normalized and scaled original DIS and VAL datasets. List of genes involved in cholesterol homeostasis were retrieved from MsigDB<sup>72</sup>. UCell score for cholesterol signature was calculated by implementing *AddModuleScore\_UCell* function on Seurat object. For serrated lesion and normal dataset, obtained cholesterol signature distribution in different cell types was visualized on UMAP space and a violin plot by implementing *FeaturePlot* and *VlnPlot* functions, respectively. For original DIS and VAL datasets, violin plots were used to compare cholesterol signature distribution in SSC and adenoma-specific (ASC) cell populations. For statistical analysis, mean comparison p values were added to violin plots with *stat\_compare\_means* function from *ggpubr* (v0.4.0) R package<sup>81</sup>. Means were compared using the nonparametric Wilcoxon rank sum test with SSC cluster as a reference group.

*scRNA-seq, differential expression testing.* The differential testing of expression of cholesterol biosynthesis genes between SSC and remaining cell populations was performed in Seurat. *FindMarkers* function was run with the “ident.1” argument defining SSC and “feature” argument specifying a list of cholesterol genes (74 genes) to test. As a result, 21 genes of cholesterol signature were identified as up- (17 genes) and downregulated (4 genes) in SSC population. In addition, the average expression of genes upregulated in SSC was visualized across all cell types with *DotPlot* function.

*scRNA-seq, gene regulatory network.* Single-Cell rEgulatory Network Inference and Clustering (SCENIC) method was applied to infer gene regulatory network in different cell types of serrated lesion and normal dataset. The required input files included scRNA-seq expression matrix, which was generated from Seurat object using *build\_loom* function of *ScopeLoomR* R package. List of human transcription factors (1839 genes) and three hg19 gene-motifs ranking databases comparing ten species in .feather format were downloaded from <https://resources.aertslab.org/cistarget/>. First, the candidate regulatory modules of transcription factors and target genes were inferred by using the GRNBoost2 method (arboreto, v0.1.6). Next, *RcisTarget* was implemented to remove indirect targets from these modules based on motif discovery. Finally, the activity of predicted regulons was quantified at the cellular resolution by implementing the AUCell algorithm. Further, resulting regulon activity enrichment matrix was jointly analyzed with count-based matrix by using *Scanpy* (v1.9.1) python package<sup>77</sup>. To identify regulons that are specific to the SSC cluster we computed the regulon specificity score (RSS)<sup>35</sup>. The RSS was calculated for each cell type with

*regulon\_specificity\_scores* function of *pySCENIC* and the top 30 regulons for each cell type were shown on plot. Regulon activity of selected regulons was visualized on UMAP transcriptomic space with *sc.pl.umap* function of *scanpy* package.

*scRNA-seq, regulation detection of deregulated cholesterol genes.* A Seurat’s table of 21 deregulated cholesterol genes in SSC was imported to Cytoscape (v3.9.1)<sup>82</sup>, where iRegulon (v1.3)<sup>36</sup> was run with default settings to predict transcription factors potentially regulating the expression of these genes. Transcription factors that were shown previously to be specific for SSC were selected from the combination of motifs and tracks output (cut off >5 target genes). Generated regulatory network was visualized in Cytoscape with target genes node colored by *avg\_log2FC*.

**Statistics and reproducibility.** Statistical significance testing was performed using *GraphPad Prism 7* software. Comparisons between the two groups were performed using the two-tailed, unpaired Student’s *t* test. Statistical analysis was performed experiments with more than three biological replicates. The number of replicates are reported in the figure legend for each experiment. Comparisons between more than two groups were carried on using ANOVA with appropriate post hoc analysis for multiple comparison (Tukey, Dunnett). Distribution data for intestinal crypts were analyzed using the Kolmogorov–Smirnov test. Attempts to reproduce the data were successful.

**Reporting summary.** Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

Uncropped gel images for Supplementary Fig. 10b are reported in Supplementary Fig. 17. The transcriptomic data are available on GEO through accession numbers GSE236515 (RNA-seq data) and GSE234372 (microarray data). Any other data are available from the corresponding author upon reasonable request. Source data for generating graphs in Fig. 1b–d; Fig. 2e, f; Fig. 4c, e, f; Fig. 5e are reported in Supplementary Data 1. Source data for generating the graph in Fig. 4g are reported in Supplementary Data 2.

Received: 23 March 2023; Accepted: 7 September 2023;

Published online: 21 September 2023

## References

- Santos, A. J. M., Lo, Y. H., Mah, A. T. & Kuo, C. J. The intestinal stem cell niche: homeostasis and adaptations. *Trends Cell Biol.* **28**, 1062–1078 (2018).
- van der Flier, L. G., Haegerbarth, A., Stange, D. E., van de Wetering, M. & Clevers, H. OLFM4 is a robust marker for stem cells in human intestine and marks a subset of colorectal cancer cells. *Gastroenterology* **137**, 15–17 (2009).
- Barker, N. et al. Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* **449**, 1003–1007 (2007).
- Minoo, P., Zlobec, I., Peterson, M., Terracciano, L. & Lugli, A. Characterization of rectal, proximal and distal colon cancers based on clinicopathological, molecular and protein profiles. *Int. J. Oncol.* **37**, 707–718 (2010).
- Stintzing, S., Tejpar, S., Gibbs, P., Thiebach, L. & Lenz, H. J. Understanding the role of primary tumour localisation in colorectal cancer treatment and outcomes. *Eur. J. Cancer* **84**, 69–80 (2017).
- Chen, B. et al. Differential pre-malignant programs and microenvironment chart distinct paths to malignancy in human colorectal polyps. *Cell* **184**, 6262–6280.e6226 (2021).
- Joanito, I. et al. Single-cell and bulk transcriptome sequencing identifies two epithelial tumor cell states and refines the consensus molecular classification of colorectal cancer. *Nat. Genet.* **54**, 963–975 (2022).
- Muzny, D. M. et al. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).
- Fearon, E. R. & Vogelstein, B. A genetic model for colorectal tumorigenesis. *Cell* **61**, 759–767 (1990).

10. Missiaglia, E. et al. Distal and proximal colon cancers differ in terms of molecular, pathological, and clinical features. *Ann. Oncol.* **25**, 1995–2001 (2014).
11. Leach, J. D. G. et al. Oncogenic BRAF, unrestrained by TGF $\beta$ -receptor signalling, drives right-sided colonic tumorigenesis. *Nat. Commun.* **12**, 3464 (2021).
12. Weisenberger, D. J. et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat. Genet.* **38**, 787–793 (2006).
13. Davies, H. et al. Mutations of the BRAF gene in human cancer. *Nature* **417**, 949–954 (2002).
14. Rajagopalan, H. et al. RAF/RAS oncogenes and mismatch-repair status. *Nature* **418**, 934–934 (2002).
15. Dhillon, A. S., Hagan, S., Rath, O. & Kolch, W. MAP kinase signalling pathways in cancer. *Oncogene* **26**, 3279–3290 (2007).
16. Parsons, R. et al. Microsatellite instability and mutations of the transforming growth factor beta type II receptor gene in colorectal cancer. *Cancer Res.* **55**, 5548–5550 (1995).
17. Barras, D. et al. BRAF V600E mutant colorectal cancer subtypes based on gene expression. *Clin. Cancer Res.* **23**, 104–115 (2017).
18. Tong, K. et al. Degree of tissue differentiation dictates susceptibility to BRAF-driven colorectal cancer. *Cell Rep.* **21**, 3833–3845 (2017).
19. Rad, R. et al. A genetic progression model of Braf(V600E)-induced intestinal tumorigenesis reveals targets for therapeutic intervention. *Cancer Cell* **24**, 15–29 (2013).
20. Carragher, L. A. et al. V600EBraf induces gastrointestinal crypt senescence and promotes tumour progression through enhanced CpG methylation of p16INK4a. *EMBO Mol. Med.* **2**, 458–471 (2010).
21. Jänne, P. A. & Mayer, R. J. Chemoprevention of colorectal cancer. *N. Engl. J. Med.* **342**, 1960–1968 (2000).
22. Riemer, P. et al. Transgenic expression of oncogenic BRAF induces loss of stem cells in the mouse intestine, which is antagonized by  $\beta$ -catenin activity. *Oncogene* **34**, 3164–3175 (2015).
23. Feng, Y. et al. Mutant KRAS promotes hyperplasia and alters differentiation in the colon epithelium but does not expand the presumptive stem cell pool. *Gastroenterology* **141**, 1003–1013.e1010 (2011).
24. Liberzon, A. et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
25. Popovici, V. et al. Identification of a poor-prognosis BRAF-mutant-like population of patients with colon cancer. *J. Clin. Oncol.* **30**, 1288–1295 (2012).
26. Tsai, J.-H. et al. Aberrant expression of annexin A10 is closely related to gastric phenotype in serrated pathway to colorectal carcinoma. *Mod. Pathol.* **28**, 268–278 (2015).
27. Gonzalo, D. H. et al. Gene expression profiling of serrated polyps identifies annexin A10 as a marker of a sessile serrated adenoma/polyp. *J. Pathol.* **230**, 420–429 (2013).
28. Lytras, T., Nikolopoulos, G. & Bonovas, S. Statins and the risk of colorectal cancer: an updated systematic review and meta-analysis of 40 studies. *World J. Gastroenterol.* **20**, 1858–1870 (2014).
29. Mullen, P. J., Yu, R., Longo, J., Archer, M. C. & Penn, L. Z. The interplay between cell signalling and the mevalonate pathway in cancer. *Nat. Rev. Cancer* **16**, 718–731 (2016).
30. De Sousa, E. M. F. et al. Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions. *Nat. Med.* **19**, 614–618 (2013).
31. Delker, D. A. et al. RNA sequencing of sessile serrated colon polyps identifies differentially expressed genes and immunohistochemical markers. *PLoS ONE* **9**, e88367 (2014).
32. Parker, H. R. et al. The proto CpG island methylator phenotype of sessile serrated adenomas/polyps. *Epigenetics* **13**, 1088–1105 (2018).
33. Andreatta, M. & Carmona, S. J. UCell: Robust and scalable single-cell gene signature scoring. *Comput. Struct. Biotechnol. J.* **19**, 3796–3798 (2021).
34. Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
35. Suo, S. et al. Revealing the critical regulators of cell identity in the mouse cell atlas. *Cell Rep.* **25**, 1436–1445.e1433 (2018).
36. Janky, R. et al. iRegulon: from a gene list to a gene regulatory network using large motif and track collections. *PLoS Comput. Biol.* **10**, e1003731 (2014).
37. Beumer, J. & Clevers, H. Regulation and plasticity of intestinal stem cells during homeostasis and regeneration. *Development* **143**, 3639–3649 (2016).
38. de Jong, P. R. et al. ERK5 signalling rescues intestinal epithelial turnover and tumour cell proliferation upon ERK1/2 abrogation. *Nat. Commun.* **7**, 11551 (2016).
39. Miller, S. A. et al. LSD1 and aberrant DNA methylation mediate persistence of enteroendocrine progenitors that support BRAF-mutant colorectal cancer. *Cancer Res.* **81**, 3791–3805 (2021).
40. Herr, R. et al. B-Raf inhibitors induce epithelial differentiation in BRAF-mutant colorectal cancer cells. *Cancer Res.* **75**, 216–229 (2015).
41. Giannakis, M. et al. RNF43 is frequently mutated in colorectal and endometrial cancers. *Nat. Genet.* **46**, 1264–1266 (2014).
42. Borowsky, J. et al. The role of APC in WNT pathway activation in serrated neoplasia. *Mod. Pathol.* **31**, 495–504 (2018).
43. Panarelli, N. C., Vaughn, C. P., Samowitz, W. S. & Yantiss, R. K. Sporadic microsatellite instability-high colon cancers rarely display immunohistochemical evidence of Wnt signaling activation. *Am. J. Surg. Pathol.* **39**, 313–317 (2015).
44. Yachida, S., Mudali, S., Martin, S. A., Montgomery, E. A. & Iacobuzio-Donahue, C. A. Beta-catenin nuclear labeling is a common feature of sessile serrated adenomas and correlates with early neoplastic progression after BRAF activation. *Am. J. Surg. Pathol.* **33**, 1823–1832 (2009).
45. Cairns, R. A., Harris, I. S. & Mak, T. W. Regulation of cancer cell metabolism. *Nat. Rev. Cancer* **11**, 85–95 (2011).
46. Vander Heiden, M. G. et al. Metabolic pathway alterations that support cell proliferation. *Cold Spring Harb. Symp. Quant. Biol.* **76**, 325–334 (2011).
47. Satoh, K. et al. Global metabolic reprogramming of colorectal cancer occurs at adenoma stage and is induced by MYC. *Proc. Natl. Acad. Sci. USA* **114**, E7697–e7706 (2017).
48. Alaqbi, S. S. et al. Increased mitochondrial proline metabolism sustains proliferation and survival of colorectal cancer cells. *PLoS ONE* **17**, e0262364 (2022).
49. Alexandrou, C. et al. Sensitivity of colorectal cancer to arginine deprivation therapy is shaped by differential expression of urea cycle enzymes. *Sci. Rep.* **8**, 12096 (2018).
50. Ying, Z. et al. The unfolded protein response and cholesterol biosynthesis link luman/CREB3 to regenerative axon growth in sensory neurons. *J. Neurosci.* **35**, 14557–14570 (2015).
51. Myers, S. A., Wang, S. C. & Muscat, G. E. The chicken ovalbumin upstream promoter-transcription factors modulate genes and pathways involved in skeletal muscle cell metabolism. *J. Biol. Chem.* **281**, 24149–24160 (2006).
52. Rodríguez, J. C., Ortiz, J. A., Hegardt, F. G. & Haro, D. Chicken ovalbumin upstream-promoter transcription factor (COUP-TF) could act as a transcriptional activator or repressor of the mitochondrial 3-hydroxy-3-methylglutaryl-CoA synthase gene. *Biochem. J.* **326**, 587–592 (1997).
53. Yamazaki, T. et al. The COUP-TFII variant lacking a DNA-binding domain inhibits the activation of the Cyp7a1 promoter through physical interaction with COUP-TFII. *Biochem. J.* **452**, 345–357 (2013).
54. Vallejo, A. et al. FOSL1 promotes cholangiocarcinoma via transcriptional effectors that could be therapeutically targeted. *J. Hepatol.* **75**, 363–376 (2021).
55. Bardou, M., Barkun, A. & Martel, M. Effect of statin therapy on colorectal cancer. *Gut* **59**, 1572 (2010).
56. Jung, Y. S., Park, C. H., Eun, C. S., Park, D. I. & Han, D. S. Statin use and the risk of colorectal adenoma: A meta-analysis. *J. Gastroenterol. Hepatol.* **31**, 1823–1830 (2016).
57. Wei, J. T., Mott, L. A., Baron, J. A. & Sandler, R. S. Reported use of 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitors was not associated with reduced recurrence of colorectal adenomas. *Cancer Epidemiol. Biomark. Prev.* **14**, 1026–1027 (2005).
58. Lee, J. E. et al. Statin use and colorectal cancer risk according to molecular subtypes in two large prospective cohort studies. *Cancer Prev. Res.* **4**, 1808–1815 (2011).
59. Mo, H. & Elson, C. E. Studies of the isoprenoid-mediated inhibition of mevalonate synthesis applied to cancer chemotherapy and chemoprevention. *Exp. Biol. Med.* **229**, 567–585 (2004).
60. Wong, W. W., Dimitroulakos, J., Minden, M. D. & Penn, L. Z. HMG-CoA reductase inhibitors and the malignant cell: the statin family of drugs as triggers of tumor-specific apoptosis. *Leukemia* **16**, 508–519 (2002).
61. Kodach, L. L. et al. The effect of statins in colorectal cancer is mediated through the bone morphogenetic protein pathway. *Gastroenterology* **133**, 1272–1281 (2007).
62. Amirkhah, R. et al. MmCMS: mouse models' consensus molecular subtypes of colorectal cancer. *Br. J. Cancer* **128**, 1333–1343 (2023).
63. Ouahoud, S. et al. Statin use is associated with a reduced incidence of colorectal cancer expressing SMAD4. *Br. J. Cancer* **126**, 297–301 (2022).
64. DeStefano Shields, C. E. et al. Bacterial-driven inflammation and mutant BRAF expression combine to promote murine colon tumorigenesis that is sensitive to immune checkpoint therapy. *Cancer Discov.* **11**, 1792–1807 (2021).
65. Andersen, C. L., Jensen, J. L. & Ørntoft, T. F. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* **64**, 5245–5250 (2004).
66. Bustin, S. A. et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* **55**, 611–622 (2009).
67. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).

68. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
69. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
70. Smedley, D. et al. BioMart—biological queries made easy. *BMC Genomics* **10**, 22 (2009).
71. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omic* **16**, 284–287 (2012).
72. Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
73. Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–315 (2004).
74. Grossman, R. L. et al. Toward a shared vision for cancer genomic data. *New Engl. J. Med.* **375**, 1109–1112 (2016).
75. Colaprico, A. et al. TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* **44**, e71 (2016).
76. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
77. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
78. Reback, J. et al. pandas-dev/pandas: Pandas 1.3.1 (v1.3.1). Zenodo. <https://doi.org/10.5281/zenodo.5136416>. (2021).
79. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e1821 (2019).
80. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e3529 (2021).
81. Kassambara A ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.4.0, <https://rpkgs.datanovia.com/ggpubr/>. (2020).
82. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

## Acknowledgements

This work was supported by the Cancer Prevention Research Trust (CPRT), a CRUK program grant (A13083) and a NIHR/CRUK-funded ECOM grant [C325/A15575]. S.W. was funded by a PhD fellowship from the CPRT (grant code TM60003-CPRT), I.G. was funded by a grant from the CPRT (grant code RM60G0665-CPRT). P.R. was funded by a grant from the CPRT (grant code RM60G0817). C.T. was supported by a Wellcome Trust Research Career Re-entry Fellowship (210911/Z/18/Z, 'Exploiting transcription of repetitive DNA to study early events in colorectal cancer'). R.P.G. was supported through a Development Grant to A.R. and C.T. from the charity Hope Against Cancer (<https://www.hopeagainstcancer.org.uk/>, grant code RM60G0751). We thank Dr. Sara Galavotti and Professor Andreas Gescher for their critical reading of the manuscript. The authors acknowledge the help from the staff of the Division of Biomedical Services, Preclinical Research Facility, University of Leicester, for technical support and the care of

experimental animals. We thank the Advanced Imaging Facility (RRID:SCR\_020967) and the University of Leicester Core Biotechnology Services Electron Microscopy Facility at the University of Leicester for their support.

## Author contributions

A.R. and C.P. designed the experiments. P.R., R.P.G., and M.M. performed bioinformatics analysis, S.W., P.F., H.C., I.G., S.S.U., C.A., L.S., C.G., E.P., H.J., F.H., and S.G. performed the experiments. N.B.S. supported microarray and RNA-seq experiments. N.S.A. and A.S-I. performed electron microscopy. A.R., C.P., K.B., C.T., and R.F. analyzed the data. A.R. and P.R. wrote the manuscript. All authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42003-023-05331-x>.

**Correspondence** and requests for materials should be addressed to Alessandro Rufini.

**Peer review information** *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Silvia Belluti and George Inglis. A peer review file is available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



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

© The Author(s) 2023