# NEURO-IMMUNE CRYPT-ASSOCIATED CELLS DRIVE COLORECTAL CARCINOGENESIS VIA REST-MEDIATED PHENOTYPIC REPROGRAMMING: IMPLICATIONS FOR TUMORIGENESIS AND VIRAL SUSCEPTIBILITY

David Díaz-Carballo[1]*, Adrien Noa-Bolaño[1], Udo Rahner[1], Ali H. Acikelli[1], Sahitya Saka[2], Jacqueline Klein[1], Flevy D'Souza[1], Sascha Malak[1], Anne Höppner[1], Annabelle Kamitz[1], Carla Casula[1], Lalitha Kamada[3], Andrea Tannapfel[4], Jens Christmann[4], Sarah Albus[4], Enrico Fruth[4], Daniela Gerovska[5], Marcos J. Araúzo-Bravo[5,6,7], Metin Senkal[8], Crista Ochsenfarth[9] and Dirk Strumberg[1]

[1] Institute of Molecular Oncology and Experimental Therapeutics, Division of Hematology and Oncology, Ruhr University Bochum Medical School, Marien Hospital Herne, 44625 Herne, Germany.

[2] Department of Medical Sciences, School of Veterinary Medicine, University of Wisconsin-Madison, MI 53706, USA.

[3] Clinic of Pediatric Oncology, Hematology and Immunology, Düsseldorf University Hospital 40225, Düsseldorf, Germany.

[4] Institute of Pathology, Ruhr University Bochum, Medical School, Bürkle-de-la-Camp-Platz 1, 44789 Bochum, Germany.

[5] Group of Computational Biology and Systems Biomedicine. Biogipuzkoa Health Research, San Sebastian, Guipuzcoa, 20014, Spain.

[6] Basque Foundation for Science, IKERBASQUE, Bilbao, Spain.

[7] Department of Cell Biology and Histology, Faculty of Medicine and Nursing, University of Basque Country (UPV/EHU), Leioa, Spain.

[8] Department of General and Visceral Surgery, Marien Hospital Witten, Ruhr University Bochum, Medical School, 58452 Witten, Germany.

[9] Department of Anesthesia, Intensive Care, Pain and Palliative Medicine, Marien Hospital Herne, Ruhr University Bochum, Medical School, 44625 Herne, Germany.

*Corresponding author:

Dr. rer. medic. David Díaz-Carballo
Institute of Molecular Oncology and Experimental Therapeutics. Division of Hematology and Oncology.
Marien Hospital Herne. Clinic for Internal Medicine III.
Ruhr University Bochum, Medical School
Düngelstr. 33, 44623 Herne, Germany
Phone: 0049 2323 499 1092
Fax: 0049 2323 499 1049
david.diaz-carballo@marienhospital-herne.de

**Key words**: Neuroendocrine cells, EBV, tumor cell heterogeneity, cancer stem cells, HERVs.

**Conflict of interests:** The authors have no conflict of interests to disclose.

Running title: " THE NEUROIMMUNE NEXUS IN COLORECTAL CARCINOGENESIS: ROLES OF THE REST TRANSCRIPTION FACTOR AND EPSTEIN–BARR VIRUS IN TUMOR DEVELOPMENT"

**Abstract**

Colorectal cancer (CRC) incidence correlates anatomically with a shared gradient of lymphoid follicles and previously defined neuroendocrine cells. We now define the cellular basis for this observation by identifying a resident cell in the colonic crypts, the Neuro-Immune Crypt-Associated (NICA) cell, characterized by a unique triple phenotype: neuronal (ChgA, TH, peripherin, somatostatin, NSE), robust stemness markers (LGR5, CD133, Glut3, ASCL2, β-catenin), and a diverse array of innate immune molecules (TLR4, CD80, SDF-1). These markers (excluding ChgA) are retained in malignant CRC sections, demonstrating lineage persistence. Functioning as a pathogen gateway, NICA cells express viral entry receptors and the CXCR4 ligand SDF-1, rendering them susceptible to Epstein-Barr Virus (EBV) infection and subsequent immune cell recruitment. Crucially, we define a novel oncogenic intermediate: the BLEICS (B-Lymphocyte, EBV-Infected, Calamari-Shaped) cell, which exhibits a B cell-myeloid hybrid signature and high CXCR4. The fusion between BLEICS and crypt cells, facilitated by fusogenic Human Endogenous Retrovirus (HERV) envelope proteins, generates the malignant, hybrid CRC lineage. Furthermore, BLEICS cells actively engage cancer-associated fibroblasts (CAFs) via the CXCR4-SDF-1 axis to drive stromal expansion and crypt enlargement. In the resulting CRC cells, the overexpression of the transcription factor REST actively suppresses the neuronal identity (e.g., TH, ChgA), thereby maintaining the proliferative stemness phenotype (LGR5). The specific loss of ChgA in CRC is attributable to REST activity. We show that *REST* ablation restores ChgA expression and reinstates neuronal differentiation and induces growth arrest, an effect mirrored by pharmacological REST inhibition using valproic acid. This study establishes a new pathogenesis model where the increasing density of lymphoid follicles and NICA precursor cells toward the rectum creates an ideal niche for viral-driven, immune-cell fusion, providing mechanistic targets for REST-driven differentiation therapy and EBV vaccination strategies.

## Introduction

Colorectal cancer (CRC) remains one of the most prevalent and lethal malignancies worldwide, with incidence rates climbing even among younger adults. Traditionally, CRC pathogenesis has been associated with relentless epithelial renewal and mutations accruing in long-lived stem cells located at the crypt base. Yet, this view is evolving as increasing evidence points to additional cellular contributors and microenvironmental factors in tumor initiation.

Neuroendocrine (NE) cells are a specialized epithelial lineage characterized by the synthesis of neuronal markers and bioactive peptides that regulate endocrine and paracrine signaling[1,2]. These cells are found in nearly every organ, with the gastrointestinal and respiratory tracts exhibiting a high density of NE cells. In association, 70% of all neuroendocrine neoplasms (NENs) originate in the gut, while the respiratory tract accounts for about 25%[3]. In CRC, NE cells are detected in roughly 34% of cases, and their prevalence may be even higher in preneoplastic lesions, suggesting their significant role in carcinogenesis[4,5]. Notably, the anatomical distribution of CRC—most frequently diagnosed in the descending colon, sigmoid colon, and rectum—correlates with regions of highest NE cell and lymphatic follicle density. Such spatial patterns hint that the local microenvironment, enriched in both NE cells and immune components, may be involved in malignant transformation.

The colonic mucosa is among the most rapidly regenerating tissues in the human body, relying on the dynamic activity of the Lieberkühn crypts. Crypts are composed of discrete compartments: the crypt-base columnar (CBC) stem cell niche (marked by LGR5), an intermediate transit-amplifying zone (expressing markers such as ZNF27 and Bmi1), and a zone of differentiated mature cells[6–8]. However, emerging evidence indicates that stemness is not confined solely to the crypt base. Under stress conditions or following injury, nearly every cell type within the crypt can acquire facultative stem cell properties to maintain tissue integrity[9–11]. In alignment with the stem cell theory of cancer[10–14], such plasticity suggests that multiple cell types—including those outside the classical CBC stem cell compartment—may initiate malignancy.

In our work, we demonstrate for the first time that a subset of colonic NE cells exhibits a unique neuroimmune phenotype that we term Neuro-Immune Crypt-Associated (NICA) cells. These cells express a full complement of neuronal markers (chromogranin A, tyrosine hydroxylase, peripherin, somatostatin, and neuron-specific enolase) alongside markers associated with stemness (LGR5, CD133, Glut3, ASCL2, and β catenin) and the epithelial protein keratin 80 (KRT80). Strikingly, NICA cells also express innate immune molecules (TLR4, REG4, PigR, TNFαIP2, viperin, DEFA6, and SDF-1) as well as antigen-presenting proteins (CD22 and CD80), while lacking the pan-leukocyte marker CD45. Their preferential localization near follicle-associated epithelia—where lymphatic follicles are enriched—supports the notion that NICA cells integrate neural, regenerative, and immunological functions to maintain mucosal homeostasis.

A pivotal observation in our study is that tumor crypts display a marked loss of neuronal gene expression—approximately 70–75% reduction relative to normal tissue—concomitant with an upregulation of *REST*, a transcriptional repressor of neuronal pathways[15]. Despite this dedifferentiation, tumors retain neuronal, stemness markers and KRT80 expression, suggesting that the malignant cells are derived, at least in part, from an original NICA lineage. Moreover, we demonstrate that NICA cells express receptors (Glut1, EPHA2, CD80, and NgR1) commonly exploited by viruses, rendering these cells potential targets for oncogenic infections. Additionally, we document the emergence of a unique cell population—BLEICS (B-Lymphocyte, EBV-Infected, Calamari-Shaped) cells—that appears prominently in colorectal tumor tissues. These cells exhibit characteristic “calamari” morphology, express stathmin, and harbor Epstein–Barr virus (EBV) genome along with intact human endogenous retroviral (HERV) envelope genes. The latter may confer fusogenic properties that facilitate cell–cell fusion events and further contribute to tumor heterogeneity. In particular, BLEICS cells interact robustly with cancer-associated fibroblasts (CAFs) via a CXCR4–SDF-1 axis, driving stromal expansion—a hallmark of advanced tumors.

In summary, our work proposes an integrative model of CRC pathogenesis whereby NICA cells—at the nexus of neuronal, stem, and immune functions—undergo malignant transformation under the influence of viral infections and cell fusion events. The emergent BLEICS cells further modulate the tumor microenvironment by interacting with cancer-associated fibroblasts (CAFs) via the CXCR4–SDF 1 axis, leading to robust stromal expansion. These findings not only redefine the cellular origins of CRC but also unveil novel therapeutic targets, including antiviral approaches, fusion-inhibition strategies, and immunomodulatory treatments.

## Methods

### Ethical Considerations, Patient Material, and Cell Lines

This study was approved by the Ethics Committee of the Ruhr University Bochum, Medical Faculty (approval numbers: 15-5416, 17-6114, and 5235-15). All procedures were conducted in accordance with institutional guidelines and relevant regulations. Written informed consent was obtained from all participants prior to inclusion in the study. Patient samples were anonymized and coded to ensure confidentiality, and access was restricted to authorized research personnel. No subjects dropped out during the study period. Sex and gender data were not collected, as they were not considered pertinent to the study's objectives. No blinding procedures were implemented, as all analyses were based on anonymized samples and predefined gating strategies.

### Colorectal Carcinoma Cell Lines

The colorectal carcinoma cell lines HCT8 (HRT-18, CVCL_2478), HT29 (HTB-38, CVCL_0320), and CaCo2 (HTB-37, CVCL_0025) were obtained from the tumor bank of the Medical Faculty of the University of Duisburg-Essen. Cell lines were authenticated by the cell bank using standard short tandem repeat (STR) profiling. Cells were maintained in Dulbecco's Modified Eagle Medium (DMEM; PAN Biotech, Aidenbach, Germany) supplemented with 10% heat-inactivated fetal calf serum (FCS) and 3 µg/mL doxycycline. Cultures were incubated at 37 °C in a humidified atmosphere containing 5% $CO_2$. Cells were routinely tested for mycoplasma contamination using the MycoAlert Mycoplasma Detection Kit (Biozym, Hessisch Oldendorf, Germany), according to the manufacturer's instructions.

### Isolation of Primary Colorectal Crypts, Carcinoma Cells, and BLEICS Cells

Colorectal tissue samples from tumor and adjacent normal regions of the same patients were freshly collected immediately following surgical resection. Specimens were washed three times with pre-chilled Hank's Balanced Salt Solution (HBSS) supplemented with 1% penicillin–streptomycin–amphotericin B (Pan Biotech, Aidenbach, Germany), 25 µg/mL gentamycin, 25 µg/mL fluconazole, and 50 µg/mL metronidazole. The lamina propria mucosae was mechanically separated from the underlying submucosa using sterile forceps and fine scissors.

The epithelial mucosa was rinsed twice in fresh, pre-chilled HBSS and transferred into 50 mL conical tubes containing 25 mL of an enzymatic digestion cocktail composed of 1 mg/mL each of collagenase I, collagenase II, and dispase I in 1× HBSS. Tissue was finely minced within the enzymatic solution and incubated at 37 °C for 30 minutes on an end-over-end rotator. After digestion, samples were centrifuged at 500 × g for 10 minutes. Pellets were resuspended in fresh enzymatic cocktail and subjected to a second 30-minute incubation under identical conditions. The resulting cell suspensions were filtered through 500-µm mesh strainers and centrifuged again at 500 × g for 4 minutes at room temperature.

Crude isolates of regular crypts (RCs) and tumor-derived crypts (TCs) were recovered and washed three times in DMEM/F12 medium supplemented with 10% FCS and the same antibiotic/antimycotic cocktail (excluding fluconazole). Crypts were either cultured in DMEM/F12 medium containing 20% FCS and antibiotics to establish primary colorectal carcinoma cultures, or washed with PBS and immediately processed for RNA extraction and transcriptomic analysis.

Following isolation, tumor-derived crypts were cultured to promote expansion of epithelial cells. During early culture stages, non-epithelial cells (e.g., fibroblasts and immune cells) were gradually depleted through selective medium changes and wash steps. Upon passaging, epithelial tumor cells were preferentially retained due to their strong adherence and colony-forming properties, while loosely attached contaminating cells were removed. Continued passaging enriched the epithelial fraction, leading to the establishment of stable primary colorectal carcinoma cell populations. Independently adhering fibroblasts that proliferated during this process were expanded separately for use in co-culture and mechanistic assays.

### Characterization of Primary Colorectal Carcinoma Cell Lines and BLEICS Cells

In addition to commercial colorectal carcinoma cell lines, two primary tumor-derived cell lines—Col33 and Col124—were selected for further experimental analysis based on their distinct biological characteristics. Col33 is a rapidly proliferating culture capable of forming tubular, crypt-like structures in two-dimensional culture systems. Notably, these cells exhibit a radial organization centered on a distinctive polarity axis (see SI: CRC primary cells morphology and cellular organization). In contrast, Col124 is a slow-growing culture characterized by robust expression of E-Cadherin 17 and resistance to detachment by standard enzymatic dissociation methods.

BLEICS cells (B-lymphocyte, EBV-Infected, Calamari-Shaped) emerged spontaneously in fibroblast cultures within 5–10 days post-tumor isolation. These cells were maintained in DMEM/F12 medium supplemented with 10% fetal calf serum and 3 µg/mL penicillin–streptomycin–amphotericin B. BLEICS cells remained in suspension and were readily collected by harvesting the culture supernatant.

### Immunotechniques

All immunological techniques—including Western blotting, immunocytochemistry (ICC), Immunohistochemistry (IHC), and flow cytometry—were performed according to standard protocols as previously described[16].

### Western Blotting

Cells in the exponential growth phase were harvested for protein extraction. Monolayers were washed with PBS and detached using 0.05% trypsin (PAN Biotech, Aidenbach, Germany). The resulting cell suspensions were neutralized with culture medium and centrifuged at 300 × g for 5 minutes. Cell pellets were washed in cold PBS, centrifuged, and lysed in RIPA buffer (150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, and 50 mM Tris-HCl, pH 7.4) supplemented with a protease inhibitor cocktail (Roche Diagnostics GmbH, Basel, Switzerland). Lysis was carried out on ice for 30 minutes, followed by centrifugation at 14,000 × g at 4 °C for 20 minutes. Supernatants were collected, and protein concentrations determined using the Bradford assay.

Thirty micrograms of total protein per sample were resolved by SDS-PAGE using a 4–20% gradient gel (Biotrend, Cologne, Germany) in Tris–glycine buffer (0.025 M Tris-HCl, 0.192 M glycine, pH 8.5), then transferred onto 0.2 µm nitrocellulose membranes (Bio-Rad Laboratories, CA, USA). Membranes were blocked with 10% non-fat dry milk in PBST (0.05% Tween-20 in 1× PBS), washed, and incubated overnight at 4 °C with primary antibodies diluted in 1.25% non-fat milk in PBST, as recommended by the manufacturers (SI: Table 1: Antibodies). After three PBST washes, membranes were incubated with horseradish peroxidase-conjugated secondary antibodies diluted in 2.5% non-fat milk in PBST for 2 hours at room temperature. Immunoreactivity was visualized using Western Lightning Plus-ECL (PerkinElmer) and imaged using the ChemiDoc XRS+ system equipped with Image Lab Version 2.0.1 software (Bio-Rad Laboratories).

### Immunocytochemistry (ICC)

Cells were cultured in chamber slides to the appropriate density, washed with 1× PBS, fixed with 4% formaldehyde in PBS for 20 minutes, and stored at –20 °C until use. Prior to staining, slides were thawed, rinsed twice with PBS (5 minutes each), permeabilized with 0.1% Triton X-100 in PBS, and blocked for 2 hours at room temperature with 10% bovine serum albumin (BSA; AbD Serotec, London, UK) and 1:500 human FcR Blocking Reagent (Miltenyi Biotec, Gladbach, Germany) in PBST. Primary antibodies (SI: Table 1 Antibodies), diluted in 1% BSA in PBST, were applied overnight at 4 °C.

Following three PBST washes, cells were incubated with secondary antibodies diluted in 1% BSA in PBST for 2 hours at room temperature in the dark. After an additional PBST wash (5 minutes), cell

nuclei were counterstained with Hoechst 33258 (1 µg/mL in PBST) for 15 minutes. Slides were rinsed three times in PBST and mounted using Faramount Mounting Medium (Agilent Technologies, CA, USA). Samples were visualized using an Eclipse i-50 microscope (Nikon, Tokyo, Japan) and analyzed with NIS-Elements Advanced Research imaging software.

**Immunohistochemistry (**IHC**)**

Paraffin-embedded tissue sections (5 µm thickness) were baked overnight at 60 °C to ensure firm adhesion to microscope slides and to eliminate residual paraffin. Sections were deparaffinized by two 20-minute incubations in xylene substitute (Thermo Scientific, London, UK), followed by rehydration in a graded ethanol series (3×100%, 95%, 70%, and 50%; 25%, $dH_2O$, PBS, 5 minutes each). Antigen retrieval was performed by heating the sections in 10 mM sodium citrate buffer (pH 6.0; Sigma-Aldrich, Taufkirchen, Germany) at 95 °C for 30 minutes using a domestic vegetable steamer. After cooling, slides were washed twice with 1× PBS for 5 minutes and blocked for 2 hours at room temperature in 5% BSA (AbD Serotec, London, UK) and 1:500 human FcR Blocking Reagent (Miltenyi Biotec, Gladbach, Germany) prepared in PBST.

Primary antibodies (SI: Table 1 Antibodies) were diluted in 1% BSA in PBST according to the manufacturers' specifications and applied overnight at 4 °C. The following day, slides were washed three times with PBST (5 minutes each), and tissue sections were incubated for 2 hours at room temperature with conjugated secondary antibodies (Cell Signaling, Leiden, Netherlands) diluted in 1% BSA in PBST, following the manufacturers' instructions.

Nuclear staining with Hoechst 33258, mounting, and imaging were performed as described in the Immunocytochemistry section.

**Peripheral Blood Collection and Flow Cytometry Analysis**

Peripheral blood samples were collected from 102 patients diagnosed with colorectal cancers (mean age: 72 years), as well as from two control groups: young healthy individuals (<30 years) and elderly healthy donors (mean age: 83 years). To investigate systemic immune alterations, particularly in relation to BLEICS cell involvement, comprehensive flow cytometric profiling of peripheral blood leukocytes was performed. All cytometric parameters, including antibody panels, fluorochrome combinations, and gating strategies, were standardized across all study groups. Reagents and protocols were applied uniformly to ensure comparability of results.

For immune phenotyping, flow cytometric analyses were performed using established protocols. In brief, 50 µL of EDTA-anticoagulated whole blood was incubated for 20 minutes with fluorochrome-conjugated antibodies (SI: Table 1 Antibodies; BioLegend, Amsterdam, Netherlands). Following staining, samples were diluted to a final volume of 500 µL with VersaLyse solution (Beckman Coulter, Krefeld, Germany) and incubated at room temperature for 5 minutes to lyse erythrocytes.

Appropriate isotype-matched control antibodies (SI: Table 1 Antibodies) were employed for signal gating. Labeled cells were analyzed using a CytoFLEX Research Cytometer B5-R5-V5 (Beckman Coulter, Krefeld, Germany). For each sample, a minimum of $2 \times 10^4$ events were acquired to ensure sufficient statistical power. The main cell population was identified based on forward and side scatter (FSC/SSC) dot plot density, allowing exclusion of dead cells. Isotype control signals were used to define the threshold for negative staining.

**Histochemistry**

Fontana-Masson staining was performed for the histological detection of argentaffin cells using a validated kit (ScyTek, Utah, USA), following the manufacturer's protocol. After deparaffinization, tissue sections were immersed in silver nitrate solution pre-equilibrated with ammonium hydroxide and incubated at 58 °C until they developed a yellowish-brown coloration, typically after 30 minutes.

Slides were then rinsed three times with distilled water and treated with 0.2% gold chloride solution for 30 seconds. Following another three rinses in distilled water, sections were incubated in 5% sodium thiosulfate solution for 2 minutes to stabilize silver deposits. After additional rinsing, tissues

were counterstained with Fast Red solution for 5 minutes and subsequently mounted using Eukitt® mounting resin.

### Transmission Electron Microscopy (TEM) and Three-Dimensional Reconstruction

BLEICS cells were fixed in 2.5% glutaraldehyde (GA) in PBS for 3 hours at room temperature, washed three times for 20 minutes each with PBS, and post-fixed in 1% osmium tetroxide (pH 7.4) for 45 minutes. Residual osmium tetroxide was removed by three additional PBS washes (3 × 20 min). Samples were then pre-dehydrated in 30% ethanol (3 × 20 min) and subjected to a graded ethanol dehydration series (30%–100%), 5 minutes per step. Dehydrated samples were incubated in a 1:1 (v/v) mixture of propylene oxide and ethanol for 15 minutes, embedded in pure EPON®, followed by a 1:1 (v/v) mixture of EPON® and Araldite for 15 minutes, and polymerized by desiccation at 80 °C for 48 hours.

Resin-embedded blocks were sectioned into 60 nm ultrathin slices using a Leica Ultracut UCT ultramicrotome (Leica, Wetzlar, Germany). Sections were mounted on 200-mesh hexagonal copper grids (Leica) and contrasted using a Leica EM AC20 system with 0.5% aqueous uranyl acetate for 45 minutes at room temperature. Grids were further stained with 3% lead citrate and rinsed with Milli-Q water for 10 minutes. Samples were examined using a JEOL 2100 transmission electron microscope (JEOL, Tokyo, Japan) operating at 120 kV. Images were captured at magnifications ranging from 3,000× to 140,000× using ITEM® 5.0 software (Soft Imaging Systems, Cardiff, UK).

For three-dimensional reconstruction, semi-thick (0.5 µm) sections were generated using the same ultramicrotome and stained with 1% toluidine blue, azure II, methylene blue, and sodium tetraborate in distilled water. Sections were mounted on glass slides, air-dried, and cover slipped with a xylene substitute. Serial images from 50 consecutive sections were acquired at 1,000× magnification on an Eclipse i-50 microscope (Nikon, Tokyo, Japan), aligned using Adobe Photoshop CS4, and processed into animated reconstructions using JASC Animation Shop 3.01.

### Scanning Electron Microscopy (SEM)

BLEICS cells co-cultured with cancer-associated fibroblasts (CAFs) were seeded on Thermanox coverslips and grown to appropriate densities. After medium removal, samples were washed three times with PBS and fixed with 2.5% GA in PBS for 2 hours at room temperature. Fixed specimens were rinsed three times with Milli-Q water (15 minutes each), dehydrated via a graded ethanol series (10%–100%, 5 minutes per step), and air-dried in a fume hood. After sputter-coating with gold, samples were imaged using a Zeiss DSM 982 scanning electron microscope equipped with Gemini optics and the Digital Image Scanning System (DISS5; Point Electronic, Halle, Germany).

### Time-Lapse Videography

Live-cell imaging was performed using a TE2000 inverted microscope (Nikon, Tokyo, Japan) equipped with a PeCom environmental incubation system (5% $CO_2$, 37 °C) to maintain physiological culture conditions. Time-lapse image acquisition and processing were carried out using NIS-Elements BR software (Nikon Vision Systems).

### Total RNA Isolation and RT-qPCR

Cells were cultured in DMEM/F12 medium supplemented with 10% FCS (or 20% FCS for primary cultures) and harvested at approximately 70% confluency. Total RNA was extracted using TRIzol® reagent (Life Technologies, California, USA), followed by DNase I treatment (7 Kunitz units; Qiagen, Hilden, Germany) to remove genomic DNA contamination. RNA samples were further purified using RNeasy Mini columns (Qiagen) according to the manufacturer's protocol. RNA integrity was assessed using a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA).

Neuronal marker transcripts were quantified via RT-qPCR using custom-designed primers and probes (SI: Table 2: qPCR primers and probes), synthesized by IDT Inc. (Iowa, USA). Amplification of 100 ng of cDNA per reaction was performed in triplicate using 250 nmol primers and PrimeTime® Gene Expression Master Mix (IDT) on a CFX96™ Real-Time PCR Detection System (Bio-Rad Laboratories, CA,

USA). Relative gene expression levels were calculated using the comparative ΔCt method. For downstream analysis, only transcripts with ΔCt values below 29 cycles were considered reliably expressed.

**Transcriptome Analysis**

Total RNA (0.5 µg) was extracted from colorectal carcinoma cell lines (HCT8, HT29, CaCo2, Col33, and Col124), regular and tumor-derived crypts from three individual patients, BLEICS21 and BLEICS126 cells, the B-lymphoma cell line Namalwa, and peripheral B cells isolated from buffy coats of ten healthy donors. RNA samples were subjected to next-generation sequencing (NGS) on the Illumina platform (BGI Genomics, Hong Kong, China). Transcriptome data from this study was deposited in NCBI GEO under the accession number GSE303666.

Initial transcriptomic profiling was performed using the Dr. Tom platform (BGI Genomics), followed by additional bioinformatic analyses. Differential gene expression analysis was conducted using the DESeq2 package (v1.42.1), and gene set enrichment analysis (GSEA) was performed using the GSEA 4.3.3 desktop application (Broad Institute). Protein–protein interaction (PPI) networks and enrichment clusters were explored via the STRING database, while immune cell-type deconvolution was carried out using CIBERSORTx with the LM22 signature matrix.

For data visualization and enrichment mapping, Cytoscape was employed in conjunction with the EnrichmentMap plugin and associated analytic tools.

To investigate the expression of human endogenous retroviruses (HERVs), the genomic coordinates of all transcriptional peaks were intersected with a curated reference set comprising ~3,220 individual HERV loci annotated in the human genome. These loci were previously characterized using the ERVmap database, which provides high-resolution annotations of HERV family classifications, genomic positions, and transcriptional activity[17].

**CRISPR/Cas9 Ablation of Genes Encoding RE1-Silencing Transcription Factor (*REST*)**

CRISPR/Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated protein) technology was utilized to knock out all variants of the *REST* gene (NM_005612.5). The Alt-R™ CRISPR/Cas9 System and HPRT gRNA controls (IDT Technologies, Iowa, USA) were employed according to the manufacturer's recommendations. crRNAs targeting REST were specifically designed using the IDT CRISPR design tool (https://eu.idtdna.com/site/order/designtool/index/CRISPR_CUSTOM) and synthesized by the same provider (SI: REST CRISPR/Cas9 ablation).

Guide RNA (gRNA) sequences were aligned to the human genome using the Basic Local Alignment Search Tool (BLAST) to assess potential off-target effects. The crRNA:tracrRNA complexes were prepared at equimolar concentrations in Nuclease-Free Duplex buffer (IDT), incubated at 95 °C for 5 minutes, and subsequently cooled to room temperature.

Ribonucleoprotein (RNP) complexes were formed following IDT protocols by incubating 0.1 µM of the crRNA:tracrRNA complex with 1 U of Alt-R® S.p. HiFi Cas9 Nuclease V3 (IDT Technologies) for 20 minutes at room temperature. The resulting RNPs were immediately transfected into $HCT8^{WT}$ cells using the 4D-Nucleofector™ System (Lonza), applying pulse code EH-100.

Genome editing was confirmed using the EnGen™ Mutation Detection Kit (New England BioLabs, MA, USA). Genomic DNA was isolated with the NucleoSpin DNA RapidLyse Kit (Macherey-Nagel). Amplicons were generated via endpoint PCR using control primers. Wild-type and edited PCR products underwent heteroduplex formation (denaturation at 95 °C for 5 minutes, followed by annealing from 95 °C to 25 °C at a ramp rate of 0.2 °C/s). Heteroduplexes were then treated with EnGen T7 Endonuclease I at 37 °C for 15 minutes. DNA fragments were separated on a 2.0% agarose gel, stained with ethidium bromide, and visualized using the ChemiDoc™ XRS+ Imaging System (Bio-Rad, CA, USA).

CRISPR/Cas9-edited cells were cloned using limiting dilution. Cells were diluted to a concentration of 100 cells/mL and seeded into 96-well plates at one cell per well, with single-cell occupancy confirmed

by field microscopy. Wells containing a single cell that exhibited growth were selected, and the cells were recovered by trypsinization (100 µL). These were subsequently transferred to 24-well plates for expansion and then to 25 $cm^2$ flasks for further cultivation and expansion. Clonal populations were analyzed by qPCR to assess changes in REST mRNA expression, and REST protein levels were evaluated by Western blotting. Clones lacking REST expression were considered successfully mutated and selected for downstream analyses.

**Generation of Organoids from Colorectal Carcinoma Cells**

Organoids are three-dimensional cellular assemblies capable of self-organization and self-renewal, designed to replicate the complex growth patterns and cellular diversity of tumors. To generate organoids from colorectal carcinoma (CRC) cells, we followed a previously described protocol with minor modifications[16]. Colorectal carcinoma (CRC) cells were embedded in the basal membrane matrix Geltrex™ (Thermo Fisher Scientific, MA, USA), forming domes containing 3,000 cells in 10 µL volumes on culture slides. The embedded cells were subsequently maintained in organoid medium (SI: Table 3 Colonic Organoid Medium), with media changes performed every three days. Organoids were used approximately one week after initiation to evaluate the impact of REST ablation and the effects of REST inhibitors on their growth.

**Epigenetic Modulation of the Transcription Factor REST in CRC Cells Using Specific Inhibitors**

To induce a semi-neuronal phenotype in colorectal cancer (CRC) cells and thereby explore latent lineage characteristics, the HCT8 cell line and Col33 CRC cells were treated with subtoxic concentrations of key epigenetic modulators. The cytotoxicity of the REST inhibitor drugs X5050 and GSK 126 was initially assessed in organoids derived from Col33 cells (see SI: Toxicity of REST inhibitors X5050 and GSK126). Valproic acid was excluded from testing due to its well-documented low toxicity in culture, showing no impact on cell growth at concentrations up to 100 µg/mL. The experimental aim was to partially reverse oncogenic epigenetic reprogramming and assess transcriptional changes linked to neural gene regulation.

Cells were incubated with the histone deacetylase (HDAC) inhibitor valproic acid (30 µg/mL) and the EZH2 methyltransferase inhibitor GSK126 (10 µg/mL) for 72 hours. Valproic acid was selected due to its well-established ability to alter chromatin accessibility by inhibiting HDAC activity, thereby facilitating the expression of silenced neuronal genes. GSK126 targets EZH2—the catalytic component of the polycomb repressive complex 2 (PRC2)—which catalyzes trimethylation of histone H3 on lysine 27 (H3K27me3), a mark associated with gene silencing. Inhibiting EZH2 can derepress genes involved in neuronal differentiation, including targets of REST.

Following treatment, total RNA was extracted according to previously described protocols. Gene expression analysis was performed, focusing on the expression levels of REST-regulated and neurogenesis-related genes. This strategy allowed evaluation of how epigenetic remodeling affects REST functionality and downstream gene networks, potentially identifying regulatory mechanisms relevant to CRC differentiation and therapeutic susceptibility.

**Statistic and Reproducibility**

Statistical analysis of immune profiling data was performed using paired, two-tailed Student's *t*-tests, with significance defined as $p < 0.05$. Graphical representation and statistical computations were conducted using Microsoft Excel 2016 and OriginLab 2022 (v9.9.0.225; OriginLab Corp., MA, USA). Western blot quantification was carried out by densitometric analysis, with signal intensities normalized to housekeeping protein levels.

Immunocytochemistry (ICC) and immunohistochemistry (IHC) data were evaluated qualitatively for descriptive purposes and were not subjected to statistical testing. All experimental results reflect findings from at least three independent biological replicates.

## Results

### Figure 1

### Argentaffin Cells in Colonic Epithelium: Unveiling Neuronal and Stemness Phenotypes

The presence of enterochromaffin (argentaffin) cells in the gastrointestinal tract has been recognized for decades[2,18]. In the colon, these cells are located within crypts but also appear in the lamina propria and submucosa, forming part of the subepithelial neuroendocrine complex (**Figure 1A-Ia** & **Ib**). Within crypts, argentaffin cells typically display a balalaika shape, with elongated extensions protruding into the crypt lumen. They exhibit no specific spatial distribution in regular crypts.

These cells demonstrate neuronal characteristics, expressing chromogranin A (ChgA), tyrosine hydroxylase (TH) (**Figure 1A-IIa** & **IIb**), peripherin (**Figure 1A-IIIa** & **IIIb**), somatostatin (**Figure 1A-Iva** & **IVb**), neuron-specific enolase (NSE) (**Figure 1A-Va** & **Vb**), PYY (**Figure 1A-Via** & **VIb**), keratin 80 (KRT80) (**Figure 1A-VIIa** & **VIIb**), and EPH receptor A2 (EPHA2) (**Figure 1A-VIIIa** & **VIIIb**). Notably, EPHA2 functions as an entry receptor for both EBV and KSHV gamma-herpesviruses[19], which have oncogenic potential. These cells exist in two distinct states—an "activated" form, expressing ChgA, and a "quiescent" form, which lacks ChgA expression. This dual status is observed for other markers in balalaika-shaped cells as well (data not shown).

Exploring the stemness characteristics of these cells reveals they express LGR5 (**Figure 1B-Ia/Ib**), CD133 (**Figure 1B-IIa** & **IIb**), Glu3/SCLA3 (**Figure 1B-IIIa** & **IIIb**), ASCL2 (**Figure 1B-Iva** & **IV**b), and β-catenin (**Figure 1B-Va** & **Vb**), the latter being closely linked to stemness regulation and colorectal carcinoma biology[20]. DCLK1 in healthy tissues marks tuft cells in the gut and it is associated with a distinct subpopulation of stem-like cells that drive tumor progression and recurrence. Balalaika-shaped cells and particular cells localized to the *Lamina propria* express this protein. In addition, DCLK1 and ChgA double positive cells are expressed in distinct cell of the Plexus myentericus. Interestingly, lymphatic follicles cells are also positive for DCLK1 (SI: DCLK1 in subepithelial tissues of colon). However, some of these markers are not exclusive to colonic stem cells; they also appear in tissues undergoing regeneration or carcinogenesis[21].

Interestingly, some lamina propria cells express markers similar to those found in crypt-located argentaffin cells. This raises the question of whether balalaika-shaped cells originate directly from $LGR5^+$ crypt stem cells or migrate from the subepithelial compartments. Supporting this hypothesis, we observed neuronal phenotype cells drilling into colonic crypts, evidenced by their hyperbolic nuclear deformation during epithelial accommodation (**Figures 1C-Ia**, **1C-IIa-c** & **1C-IIIa-c**).

### Figure 2

### Neuro-Immune Crypt-Associated (NICA) Cells: Bridging Innate and Adaptive Immune Systems

Following the characterization of the neuronal phenotype of argentaffin balalaika-shaped cells, we investigated their distribution in the colonic epithelium and proximal subepithelial compartments. **Figures 2A-Ia-c** illustrate their presence in the epithelial lining and their potential interaction with follicle-associated epithelium (FAE). Notably, a gradient of distribution was observed, with a greater concentration of balalaika-shaped cells near FAE, suggesting a possible functional link to subepithelial lymphoid structures (**Figure 2A-Ia**).

### Immune Features of Balalaika-Shaped Cells

One of the first indications that balalaika-shaped argentaffin cells might have immunological properties was their expression of glycoprotein 2 (Gp2), a bacterial transcytotic receptor typically found in Microfold (M) cells, which are key players in mucosal immunity[22]. This initially led us to hypothesize that balalaika-shaped cells could represent a specialized M cell variant. Further

supporting this hypothesis, these cells were rarely observed to form pocket-like spaces in their basolateral region, housing adjacent immune cells (SI: NICAS basolateral pocket). However, unlike typical M cells, Gp2 expression in balalaika-shaped cells extends throughout the entire cell body, and they secrete a soluble form of Gp2 into the crypt lumen, forming a protective layer over the epithelial surface (SI: NICAS basolateral pocket).

To further investigate whether balalaika-shaped cells contribute to innate immunity, we analyzed their expression of additional immune markers. These cells were specifically labeled by lectin peanut agglutinin (PNA) (**Figure 2B-IIa** & **IIb**), which binds the Thomsen-Friedenreich (T) antigen, a molecule associated with neuraminidase-producing viral infections and considered oncodevelopmental in the human colon[23,24]. In addition, Ulex Europaeus Agglutinin 1 (UEA-1), commonly used to identify M cells in rodents, weakly labeled balalaika-shaped cells (**Figure 2B-IIIa** & **IIIb**). The glycoprotein zymogen granule protein 16 (ZG16), known to bind to various pathogens and prevent their penetration into the intestinal barrier[25], was also detected (**Figure 2B-Iva** & **IVb**).

Balalaika-shaped cells exhibit Toll-Like Receptor 4 (TLR4) (**Figure 2B-Va** & **Vb**) expression, which facilitates pathogen recognition and inflammatory cascades[26]. These cells also express REG4 (**Figure 2B-V**ia & **VIb**), a protein with bactericidal properties that mediates immune cell phagocytosis[27]. Another key immunological factor detected in these cells is PigR (Polymeric Immunoglobulin Receptor) (**Figure 2B-VIIa** & **VIIb**), which plays an essential role in mucosal immunity by transcytosing IgA and IgM. The inflammatory mediator TNFαIP2 (Tumor Necrosis Factor Alpha-Induced Protein 2), which regulates immune responses, was also present (**Figure 2B-VIIIa** & **VIIIb**). Furthermore, ICAM-1 (CD54) (**Figure 2B-IXa/IXb**) was detected, a molecule upregulated in response to TNF-alpha and IL-1, and functioning as a viral receptor for rhinoviruses and influenza[28].

Two additional antiviral and antibacterial proteins—viperin (**Figure 2B-Xa** & **Xb**) and DEFA6 (**Figure 2B-Xia** & **XIb**)—were expressed by balalaika-shaped cells. Viperin inhibits a wide range of DNA and RNA viruses[29], whereas DEFA6 provides broad antimicrobial activity against bacterial, viral, and fungal infections[30]. Additionally, SDF-1 (Stromal Cell-Derived Factor 1/CXCL12) (**Figure 2B-XIIa** & **XIIb**), responsible for immune cell recruitment during inflammation, was expressed.

Interestingly, some ChgA-derived bioactive peptides are known to exert antipathogen activity, further supporting a role for balalaika-shaped cells in innate immunity[31]. Based on their immune marker expression and neuronal phenotype, we classified these cells as a distinct neuroimmune cell type, which we now refer to as Neuro-Immune Crypt-Associated (NICA) cells.

### Adaptive Immune Properties of NICA Cells

To determine whether NICA cells possess characteristics of adaptive immune cells, we examined key immunological markers. CD45, a leukocyte common antigen (LCA) widely expressed by cells of hematopoietic origin, was absent in NICA cells (**Figure 2C-Ia** & **Ib**). However, CD45-positive immune cells were observed within crypt structures, suggesting potential interactions between NICA cells and leukocytes.

Despite lacking CD45 expression, NICA cells exhibited CD22 expression (**Figure 2C-IIa** & **IIb**). Although CD22 is primarily found on B-lymphocytes, it is also present on antigen-presenting cells (APCs), such as macrophages and dendritic cells, particularly in tumor microenvironments[32,33]. Additionally, CD80 (**Figure 2C-IIIa** & **IIIb**), a co-stimulatory molecule typically expressed by APCs, was detected in NICA cells within crypts and subepithelial regions. CD80 expression is stimulated by microbial infections[34] and intracellular parasites such as Toxoplasma gondii[35]. Moreover, CD80 serves as a receptor for clade B adenoviruses[36], some of which have potential oncogenic properties[37–39].

Collectively, NICA cells express immune components found in both innate and adaptive immune cells, although they appear more functionally aligned with innate immunity. Their neuronal phenotype suggests dual roles in mucosal defense and immune regulation. Notably, their potential interactions with viruses, including oncogenic adenoviruses, highlight their importance in host-pathogen dynamics.

Figure 3

**Colon Carcinoma Progression Leads to a Loss of Neuronal Identity and Acquisition of Immune Cell Traits**

Transcriptomic analysis revealed a loss of neuronal identity in tumor crypts compared to regular crypts. Gene Set Enrichment Analysis (GSEA) demonstrated a negative enrichment of multiple neuronal-related pathways in tumor crypts (**Figure 3A-I**). Differential expression analysis showed a significant reduction (~72%) in neuronal gene expression in tumor crypts versus regular crypts. A similar pattern was observed when comparing primary and commercial colon cancer cell cultures with normal crypts (**Figure 3A-II**, **3A-III** & SI: Neuronal genes in crypts), indicating the suppression of the neuronal phenotype in tumor cells.

***REST* Inactivation Reveals Neuronal Cell Identity in Colorectal Carcinoma Models**

The transition from a neuronal to a tumoral phenotype appears to be driven by both genetic and epigenetic mechanisms that influence differentiation pathways. One proposed driver is the transcription factor REST, a key regulator that silences neuronal gene expression. Transcriptomic analysis demonstrated REST upregulation in tumor crypts compared to normal crypts (**Figure 3**, **3A-IV**). Furthermore, analysis of a publicly available dataset investigating neuronal gene repression following REST inhibition in metastatic small-cell lung cancer (MSCLC, GSE40695 CREEDSID GENE) revealed that approximately 76% of REST-dependent neuronal genes identified in tumor crypts were similarly downregulated. These findings suggest that tumor crypts suppress neuronal gene expression during malignant progression in a REST-dependent manner.

Functional interrogation of REST's role through pharmacological inhibition with valproic acid (VPA), a histone deacetylase inhibitor, and genetic ablation via CRISPR/Cas9 in HCT8 colon carcinoma cells and primary colorectal tumor cultures consistently resulted in transcriptional activation of neuronal genes across independent biological replicates. Treatment with VPA led to upregulation of several neuronal genes commonly observed in CRC models (**Figures 4F-I** & **4G-I**). Specifically, VPA enhanced the expression of ChAT, ChgA, BDNF, neurturin, tenascin C, and the cancer stem cell markers DCLK1 and LGR5 (**Figure 3A-V**). However, REST ablation via CRISPR/Cas9 revealed that truly REST-dependent genes include ChAT, ChgA, synaptophysin, TBR2, stathmin, and LGR5. Notably, DCLK1 expression was found to be downregulated in HCT8$^{\Delta REST}$ colon carcinoma cells (**Figure 3A-VI**). These data robustly validate REST's repressive activity in the tumoral setting. However, despite significant mRNA upregulation, quantitative proteomic analyses revealed a significant reduction some of the corresponding neuronal protein levels, indicating a post-transcriptional blockade. This discordance between transcript and protein abundance underscores the complexity of REST-mediated regulation and suggests that its suppression alone is insufficient to fully derepress neuronal differentiation.

CRISPR/Cas9-mediated knockout of REST in HCT8 colorectal carcinoma cells resulted in complete loss of REST protein, confirmed by Western blot. This genetic disruption led to increased transcription of several REST target genes, including ChgA, TH, ChAT, Stathmin, TBR2, and LGR5 (**Figure 3A-VII**). Despite elevated mRNA levels, corresponding protein levels were reduced, indicating a disconnect between transcription and translation. Upon treatment with 100 µg/mL of valproic acid (VPA) however, edited cells showed restored protein expression that matched transcript levels. In contrast, VPA treatment in non-edited cells led to decreased protein levels of the same targets (SI: Effect of VPA on HCT8$^{\Delta REST}$ neuronal expression). REST ablation in the colorectal carcinoma cell line HCT8 significantly impairs organoid formation. Wild-type cells exhibit robust organoid generation (**Figure 3A-VIIIa-c**), whereas REST-deficient cells show a marked reduction in this capacity. This impaired phenotype is accompanied by a downregulation of the key stemness-associated transcription factor LGR5 (**Figure 3A-VII**), suggesting that REST may contribute to organoid formation through regulation of LGR5 expression. Consistently, pharmacological inhibition of REST using specific inhibitors

recapitulates this effect (see SI: Toxicity of REST inhibitors X5050 and GSK126), further supporting its role in maintaining stemness. Long-term treatment of primary colorectal cancer (CRC) cell cultures with valproic acid results in a marked inhibition of cell proliferation, potentially attributable to the induction of cellular senescence. Given its well-established clinical use over several decades, valproic acid emerges as a promising candidate for therapeutic intervention in the management of colorectal carcinomas (**Figure 3A-IX**). These findings highlight the existence of a multilayered regulatory architecture in colorectal cancer, wherein REST operates as one node within a broader network that governs cell identity and fate decisions.

**Acquisition of an Immune Cell Signature**

Despite retaining some stemness and neuronal traits, colorectal carcinoma (CRC) cells exhibited immune-related gene expression patterns characteristic of immune cells (SI: Immune significant genes TC). Tumor crypts gained a gene signature typical of cells from germinal centers in lymphoid follicles (**Figure 3A-IV**). This aligns with previously observed tumor infiltration by stathmin-positive cells from germinal centers in CRC paraffin sections (**Figures 4E-IIa-d**).

Using the CIBERSORTx tool, immune gene expression analysis revealed a profound remodeling of the immune microenvironment in tumor-derived colorectal crypts compared to regular crypts (**Figure 3A-V**). Immune cell deconvolution analysis identified a decline in plasma cells, $CD8^+$ T cells, activated NK cells, and M2 macrophages, along with increased infiltration by monocytes, M0 macrophages, regulatory T cells (Tregs), and $CD4^+$ memory T cells. These alterations suggest a shift toward an immunosuppressive and tumor-promoting immune environment.

**HERV Expression and Tumor Evolution**

To evaluate involvement of paleoviruses, we examined the differential expression of human endogenous retroviruses (HERVs) using ERVmap. HERV expression analysis (**Figure 5B-I**) revealed cell-type-specific regulation, with regular crypt cells exhibiting tight clustering, while tumor crypts displayed greater dispersion, indicating ongoing tumor evolution. Interestingly, isolated CRC cells were phylogenetically distant from their cell type of origin.

Structural analysis of differentially expressed HERVs (DE-HERVs) in tumor crypts versus regular crypts provided insights into their immunosuppressive and fusogenic potential. A total of 111 DE-HERVs were identified, with 78 upregulated and 33 downregulated elements (**Figure 3B-II**). Among these, certain HERVs displayed distinct expression patterns (**Figure 3B-III**).

Consistent with retroviral genomic architecture, most DE-HERVs retained partially conserved segments of the gag (CA, NC, MA) and pol (RT, IN) regions, suggesting a relatively stable subgenomic structure (SI: HERVs transcriptome data TC). **Table 1** summarizes all significantly modulated HERVs reported by Vargiu et al.[40] to possess an ISD, including annotations of retroviral genes, ISD sequence, best-matching reference elements, and super group classifications.

It is well established that human endogenous retroviruses (HERVs) can undergo transactivation following exogenous viral infection, such as with Epstein-Barr virus (EBV). In our laboratory, we have further demonstrated that the mere introduction of mRNA encoding viral structural proteins, without full viral infection, is sufficient to induce transcriptional activation of multiple HERV loci (see SI: HERV Transactivation by Viral Structural Proteins). From a translational perspective, these findings highlight the potential of HERV env proteins as biomarkers of viral cofactor activity, modulators of tumor–immune dynamics, and targets for immunotherapeutic intervention. Given that env proteins contain immunosuppressive domains (ISDs) capable of dampening host immune responses, and possess fusogenic properties that may promote syncytia formation and genomic instability, their re-expression in cancer cells may contribute to immune escape, altered cell–cell interactions, and enhanced tumor plasticity. Moreover, HERV activation may serve as a surrogate marker for chromatin accessibility or epigenetic deregulation, offering mechanistic insight into the transcriptional vulnerability of malignant cells under viral perturbation.

**Expression of HERV-Derived Gene Segments with Fusogenic and Immunosuppressive Properties**

Special focus was given to envelope (*env*) fragments, particularly the surface unit (SU) and transmembrane (TM) domains, which are critical for fusogenic activity. Among the upregulated DE-HERVs, 41 contained a TM domain, 25 an SU domain, and 19 both domains, suggesting potential functional envelope protein encoding. In contrast, downregulated DE-HERVs included 14 with a TM domain, 6 with an SU domain, and 3 carrying both domains. While fusogenic capacity appears structurally preserved in both groups, the predominance of upregulated HERVs suggests a more active fusogenic role in tumor crypts.

The analysis of immunosuppressive domains (ISDs) in ENV proteins revealed 16 upregulated and 5 downregulated DE-HERVs carrying identifiable immune-modulating sequences (see **Table 1**). Sequence logos demonstrated high ISD conservation in upregulated elements, whereas downregulated elements exhibited greater sequence variability, characterized by multiple amino acid substitutions. These findings suggest that ISD functionality may be differentially preserved based on the transcriptional activity of individual HERV loci, potentially influencing immune modulation in the colorectal tumor microenvironment.

**Figure 4**

**Colorectal Carcinomas Mirror the Neuroimmune Phenotype of NICA Cells**

Histopathological analysis of early-stage colorectal carcinomas (CRCs) revealed a localized proliferation of ChgA$^+$ cells, particularly near the follicle-associated epithelium (FAE). However, the abundance of these cells declined over time, suggesting a transition from carcinoid tumors to adenocarcinomas, the predominant histopathological classification of CRCs. **Figures 4A-I-VIII** highlight numerically amplified NICA cells, identified by ChgA expression in early CRC development. Notably, these cells did not express Lipopolysaccharide Binding Protein (LBP), an acute-phase immunological marker associated with gram-negative bacterial infections, primarily expressed by B-lymphocytes.

**Retention of Neuronal Characteristics in CRCs**

Colorectal carcinomas retain several neuronal features observed in NICA cells, including the expression of tyrosine hydroxylase (TH) (**Figure 4B-I**), peripherin (**Figure 4B-II**), somatostatin (**Figure 4B-III**), neuron-specific enolase (NSE) (**Figure 4B-IV**), and peptide YY (PYY) (**Figure 4B-V**). Additionally, stathmin, a protein crucial for neural progenitor maintenance[41] and hematopoietic stem cell function[42], was highly expressed in CRCs (**Figure 4B-VI**). In contrast, NICA cells were stathmin-negative, and other crypt cells did not clearly express this protein. DCLK1, a neuronal marker, which is also widely recognized as a stem cell marker is expressed in CRCs (**Figures 4B-VII** & **VIII**).

**Stemness Markers and Cellular Fusion Events**

CRCs exhibited a stemness phenotype, expressing markers such as LGR5 (**Figure 4C-I**), CD133 (**Figure 4C-II**), Glut3 (**Figure 4C-III**), ASCL2 (**Figure 4C-IV**), β-catenin (**Figure 4C-V**), and stathmin (**Figure 4C-VI**). Given that regular crypt cells do not express stathmin, tumor cells may have acquired it via fusion with a different stathmin-expressing cell type. This hypothesis is supported by abundant stathmin expression in leukocyte infiltrates, as shown in **Figure 4C-VI**.

**Immunological Profile of CRCs**

Beyond neuronal and stemness traits, CRCs exhibited an immune phenotype, as indicated by weak MHC-II expression (**Figure 4D-I**). It is notable that MHC-II expression is highly inducible by IFNγ in CRC

models. Furthermore, CRCs expressed multiple proteins detected in NICA cells, including ZG16 (**Figure 4D-II**), TLR4 (**Figure 4D-III**), PigR (**Figure 4D-IV**), TNFAαIP2 (**Figure 4D-V**), CD54 (**Figure 4D-VI**), viperin (**Figure 4D-VII**), DEFA6 (**Figure 4D-VIII**), CD80 (**Figure 4D-XI**), CD22 (**Figure 4D-X**), REG4 (**Figure 4D-XI**), NGR1 (**Figure 4D-XII**), and SDF-1 (**Figure 4D-XIII**). Notably, CD45 expression was nearly absent in CRC sections (**Figure 4D-XIV**), consistent with observations in NICA cells. Additionally, CRCs exhibited high levels of KRT80 expression (**Figure 4D-XIV**).

**Leukocyte Infiltration and Fusion Events in CRCs**

Given that NICA cells do not express stathmin, analysis of CRC paraffin sections identified a stathmin-positive B-lymphocyte population in the FAE, which was also observed infiltrating colonic crypts. These cells appear to undergo cell-cell fusion with crypt cells, particularly in early tumor development stages (**Figures 4E-Ia** & **Ib**, **Figures 4E-IIa-d**).

**Validation in CRC Cell Cultures**

To confirm these immunohistochemical findings, we extended our investigation to commercial and primary CRC cell cultures, detecting a distinct neuroimmune protein expression profile (**Figures 4F-I** & **4F-II)**. Transcriptomic analysis validated the neuronal background and stemness characteristics of CRCs in commercial cell lines, including HCT8, HT29, and CaCo2, as well as two primary CRC lines, Col33 and Col124 (**Figure 4G-I**).

<u>Figure 5</u>

**Impaired immune status in the blood of CRC patients**

To explore systemic immune alterations in colorectal cancer (CRC) patients, particularly in the context of BLEICS cell involvement, we performed comprehensive flow cytometric analysis of peripheral blood leukocytes from 102 CRC patients and compared them to young healthy individuals (<30 years) and an elderly control group. All cytometric parameters, reagents, and protocols were standardized across groups.

As shown in **Figure 5A-I**, CRC patients exhibited a notable increase in monocyte counts alongside mild lymphopenia. The most consistent and significant change within the lymphoid compartment was a reduction in $CD4^+$ T helper cells. This decrease was statistically robust and marked a defining difference between CRC patients and healthy controls. **Figure 5A-II** shows that B cell counts were also decreased in both elderly individuals and CRC patients relative to the younger cohort, while $CD8^+$ cytotoxic T cell levels remained relatively stable across all groups. Natural killer (NK) cells were significantly elevated in CRC patients (**Figure 5A-III**), and their reduced expression of the inhibitory receptor CD158 (**Figure 5A-IV**) suggests a potentially enhanced cytotoxic phenotype—particularly in the context of MHC-I downregulation observed in tumor crypts. However, this compensatory NK response appears insufficient to eliminate tumor growth, implying the presence of tumor-intrinsic or microenvironmental escape mechanisms.

Checkpoint molecule profiling revealed significantly reduced CTLA-4 expression in both $CD4^+$ and $CD8^+$ T cells from CRC patients (**Figure 5A-V and 5A-VI**), which may limit the effectiveness of CTLA-4–directed immunotherapies in this population. In contrast, PD-1 expression was markedly elevated, particularly on $CD4^+$ cells (**Figure 5A-V**), indicating a state of chronic activation or functional exhaustion.

Chemokine receptor analysis revealed nuanced shifts across the groups. As shown in **Figure 5A-VII**, CCR5 expression in $CD4^+$ and $CD8^+$ cells increased with age but was lower in CRC patients than in their age-matched counterparts. CXCR3 expression was elevated in $CD4^+$ T cells of CRC patients compared to the younger group, and slightly increased in $CD8^+$ cells. Notably, CXCR4 expression was significantly elevated in both $CD4^+$ and $CD8^+$ cells in CRC patients (**Figure 5A-VIII**). Given that SDF-1—the ligand for CXCR4—is highly expressed by tumor-associated fibroblasts, this axis likely facilitates

lymphocyte recruitment into the tumor microenvironment. However, the persistent depletion of $CD4^+$ T cells in peripheral blood, without proportionate accumulation in tumor tissues, suggests that these cells may undergo attrition post-migration—possibly due to exhaustion, suppression, or apoptosis induced by stromal or tumor-derived signals, including retroviral elements.

Activation markers added further context: HLA-DR expression was elevated on $CD4^+$ T cells in both CRC patients and elderly individuals (**Figure 5A-IX**), indicating immune activation. CD28 expression was diminished in both $CD4^+$ and $CD8^+$ subsets of elderly individuals (**Figure 5A-X**), reflecting an age-associated trend rather than a CRC-specific effect.

Although absolute cell counts remained within reference ranges, these consistent inter-group differences—especially the selective depletion of $CD4^+$ T cells—strongly suggest tumor-associated systemic immune perturbation that may compromise antitumor responses or predispose to infection.

## Figure 6

### BLEICS cells exhibit EBV-driven reprogramming, fibroblast-dependent growth, and expression of immunosuppressive HERV elements

During the establishment of primary colorectal carcinoma (CRC) cultures, we identified a recurring population of ~5 µm lymphocyte-like cells in 17% of samples. These cells emerged alongside cancer-associated fibroblasts (CAFs) and displayed a unique squid-like morphology and dynamic motility (SI: Videographies 1–8). While roughly half of these populations failed to proliferate beyond early passages, the remaining clones persisted and became immortalized. All derived tumors had been histologically classified as adenocarcinomas, and such cells did not appear in adjacent healthy tissue cultures from the same patients—ruling out contamination or primary lymphoma[43]. We termed these cells BLEICS (B-lymphocyte, EBV-Infected, Calamari-Shaped), reflecting their morphology, viral status, and B cell origin.

Morphological characterization revealed petal-like apical structures (**Figure 6A-Ia–c)**, illustrated schematically in Figure 6A-Id and confirmed by scanning and transmission electron microscopy (**Figure 6A-IIa–d**). TEM images showed prominent lipid droplets accumulated behind the apical pole, suggesting a putative pathogen-trapping function. In culture, BLEICS cells tightly associated with fibroblasts, frequently dividing on their surfaces (**Figure 6A-IIIa–d**; SI: Videographies 9–11 and SI: Commentaries on Videography) or forming dense aggregates reminiscent of hematologic malignancies.

Phenotypically, BLEICS cells were predominantly $CD45^+$ (**Figure 6A-IVa**) and exhibited a B cell profile on FACS (SI: BLEICS Phenotype by FACS), though CD19 (**Figure 6A-IVb**) and MHC-II (**Figure 6A-IVc**) expression varied. Neurotrophin-3 (NT3) was consistently expressed (**Figure 6A-IVd)**, prompting a search for overlapping features with mucosal M cells[44,45]. Indeed, BLEICS cells expressed M cell markers such as Gp2, Spi-B, and uromodulin (**Figure 6A-Va–c**), and bound UEA-1, supporting partial functional convergence.

Serial-section SEM and 3D reconstructions revealed that BLEICS-like cells could localize within fibroblast polykaryons (**Figure 6B-Ia–d**; SI: 3D reconstruction from TEM), and videography evidence (**Figure 6B-IIa–l**; Videographies 12–15, SI: Commentaries on Videography) supported a model in which BLEICS cells migrate toward perinuclear fibroblast regions, deform host nuclei, and divide in close proximity—suggesting transient fusion events.

Quantitative PCR confirmed the presence of EBV latency genes in BLEICS cells, and cytogenetic analysis revealed chromosomal instability with near-diploid but aberrant karyotypes in ~40% of metaphases (SI: Cytogenetic I & II). BLEICS 126 harbored a pathogenic MUTYH mutation (SI: Genetic Mutations Analysis), and both cell lines exhibited somatic immunoglobulin gene rearrangements—monoclonal in BLEICS 21 and biclonal in BLEICS 126—alongside Ig production (SI: Clonality and Ig Expression). These findings point to an EBV-driven malignization process.

Global transcriptomic analysis demonstrated that the BLEICS cell lines (BLEICS 21 and BLEICS 126) clustered closely together, yet differed markedly from both normal B cells and the classic lymphoma line, Namalwa (**Figure 6C-Ia**). CIBERSORTx deconvolution of the bulk transcriptomes (**Figure 6C-Ib**) revealed a heterogeneous and hybrid immune signature. This deconvolution indicated that BLEICS cells retained core memory B cell features while simultaneously expressing significant transcriptional profiles consistent with Plasma Cells (a terminally differentiated B-cell lineage) and Activated Dendritic Cells (DCs). Furthermore, the signature also included a component consistent with M1 Macrophages (pro-inflammatory myeloid cells), solidifying the acquisition of myeloid/Antigen-Presenting Cell (APC) characteristics. This simultaneous expression of markers from both B-cell effector and myeloid APC lineages suggests a profound state of cellular plasticity or the presence of a B-myeloid hybrid subpopulation, potentially resulting from B–DC fusion. Quantitatively, BLEICS 126 displayed a less pronounced activation profile compared to BLEICS 21, suggesting a reduced transcriptional commitment or an earlier stage of differentiation within the hybrid lineage.

Functional pathway analysis demonstrated that the BLEICS transcriptome is highly enriched for genes associated with intracellular pathogen response signatures, including those linked to Toxoplasma gondii, Mycobacterium tuberculosis, and Leishmania. This high level of generalized immune activation is corroborated by the enrichment of specific viral interaction pathways, particularly those related to Epstein-Barr Virus (EBV) infection, Kaposi sarcoma-associated herpesvirus (KSHV) infection, and Viral interaction with cytokine and cytokine receptor signaling (**Figure 6C-Ic**). This persistent state of stimulation suggests a chronic immune activation. This robust transcriptional activation may reflect either pathogen mimicry driven by the cells' activated B-myeloid hybrid phenotype or epigenetic imprinting resulting from the persistent EBV infection known to drive B-cell transformation.

The hierarchical clustering heatmap reveals two distinct and inversely regulated gene expression profiles: one corresponding to the Cancer-Associated Fibroblasts (CAFs 21 and 126) and the other to the BLEICS cells (BLEICS 21 and 126) isolated from the same tumor specimens. The two CAF samples cluster tightly together, as do the two BLEICS samples. These CAFs express a robust signature of activated myofibroblasts and extracellular matrix components, including key genes such as FLNC, TAGLN, FBN2, and FHL1, which are hallmarks of a contractile and matrix-remodeling phenotype (**Figure 6C-Id**, lower red cluster). This signature is further supported by the high expression of genes like PALLO, FSTL1, and COL6A3. Conversely, the BLEICS samples form a separate, tight cluster characterized by the upregulation of a central gene set (e.g., MXRA7, CSRP1, FILIP1), and a strong downregulation of the myofibroblast signature, underscoring the fundamental difference in the transcriptional identity between the two cell types.

To characterize the neuronal lineage features of EBV-infected B-lymphocyte derivatives, we performed comparative qPCR profiling of BLEICS 21 and BLEICS 126 cells across a curated panel of neurodevelopmental, neuroendocrine, and epithelial-mesenchymal markers. As shown in **Figure 6D-I**, both cell types exhibited expression of canonical neuronal genes, including Nestin, MAP2, Beta-3 Tubulin, NeuroD2, ASCL1, and BDNF, indicating a conserved neuroectodermal signature. Additionally, transcripts associated with synaptic function and neurotransmitter signaling -such as CHAT, Synaptophysin, GDNF, and NTRK1/2/3- were detected in both populations, suggesting partial acquisition of neuronal-like traits.

Despite this shared background, distinct expression patterns emerged between the two BLEICS lines. BLEICS 21 cells showed elevated levels of c-Myc, CD24, CD133, and Oct4, consistent with a more undifferentiated or progenitor-like phenotype. Upregulation of HES1, HES5, and Notch1 further supports active Notch signaling, commonly associated with neural stem cell maintenance. Moreover, expression of Snail, Vimentin, and N-Cadherin suggests mesenchymal features and potential epithelial-to-mesenchymal transition (EMT) activity.

In contrast, BLEICS 126 cells displayed higher expression of neuroendocrine differentiation markers, including Chromogranin A, Somatostatin, SCG5, and PSD95, as well as cortical lineage regulators such

as Pax6, TBR1, and TBR2. The presence of E-Cadherin, Claudin 5, and Occludin indicates stronger epithelial characteristics, distinguishing BLEICS 126 as a more differentiated and lineage-restricted population.

Together, these data reveal that BLEICS 21 and BLEICS 126 cells share a neuronal-like transcriptional foundation but diverge in their differentiation trajectories—BLEICS 21 favoring a progenitor-like, mesenchymal profile, and BLEICS 126 adopting a neuroendocrine-epithelial identity. These findings suggest functional heterogeneity among EBV-infected B-cell derivatives and provide a molecular basis for their differential stromal interactions.

To investigate the stromal-modulatory potential of EBV-infected B-lymphocytes (BLEICS 21 cells), we first assessed EBV gene expression within these cells. As shown in **Figure 6D-II**, quantitative PCR revealed robust transcription of EBV-associated genes —including VCAp18, LMP1, and EBNA-1—confirming active viral programs in BLEICS 21 cells. Growth kinetics further demonstrated that BLEICS 21 cells stimulate fibroblast proliferation, as evidenced by increased MTT absorbance over a 28-day period (**Figure 6D-III**). To evaluate the cellular consequences of this interaction, we performed immunocytochemical staining following a three-day co-culture. CAFs were labeled with α-SMA to confirm their myofibroblastic identity, Ki67 to assess cell cycle activity, and DAPI to visualize nuclear morphology. As illustrated in **Figure 6D-IV**, co-cultured CAFs exhibited a pronounced increase in nuclear Ki67 signal within α-SMA-positive cells, indicating enhanced proliferative engagement. In contrast, CAF monocultures showed sparse Ki67 positivity, consistent with low basal proliferation. These ICC findings corroborate the stimulatory effect of BLEICS 21 cells on CAF proliferation and activation, reinforcing their role as direct stromal modulators within the colorectal tumor microenvironment.

To validate the neuro-stem-immune hybrid phenotype of BLEICS cells at the protein level, we performed immunoblotting across key lineage and signaling markers. As shown in **Figure 6E-Ia**, BLEICS 21 and BLEICS 126 cells exhibited distinct protein expression profiles consistent with their transcriptional signatures (**Figure 6D-I**). BLEICS 21 cells showed elevated levels of Stathmin, LGR5, GLUT1, ASCL1, and c-Myc, supporting a progenitor-like and metabolically active phenotype. In contrast, BLEICS 126 cells expressed higher levels of Chromogranin A, NSE, TH, and S100A4, indicative of neuroendocrine differentiation and stress response. Densitometric quantification of selected markers is presented in Figure 6E-Ic, confirming differential abundance across the two cell lines. CAF protein expression was more heterogeneous but revealed consistent detection of stromal and immune-interactive markers. As shown in **Figure 6E-Ib**, CAFs expressed β-Catenin, SOX9, EPHA2, CCR5, and Viperin, with variable intensity across experimental conditions. Quantitative analysis of CAF responses under cytokine and receptor stimulation is summarized in **Figure 6F-Ic**, demonstrating context-dependent modulation of immune and epithelial signaling components. Together, these protein-level findings reinforce the transcriptional data and highlight the functional divergence between BLEICS 21 and BLEICS 126 cells, as well as the dynamic responsiveness of CAFs to BLEICS-derived cues.

**Differentially Expressed HERVs with Intact Immunosuppressive Domains (ISDs) in BLEICS cells**

To assess retroviral involvement, we analyzed differentially expressed human endogenous retroviruses (HERVs) using ERVmap. 164 DE-HERVs were identified between BLEICS and normal B cells. Among upregulated elements, 74 were Class I ENV proteins, 20 Class II, and 4 Class III. Of these, 24 retained SU domains, 56 TM domains, and 21 encoded both. Downregulated elements included 49 Class I, 14 Class II, and 3 Class III HERVs, with 18 possessing both SU and TM domains (**Figure 6F-I**).

The analysis of HERV expression profiles focused on elements containing an Immunosuppressive Domain (ISD), as previously characterized by Vargiu et al.[40] Importantly, a total of 36 differentially

expressed (DE) HERVs were found to retain intact ISDs—specifically, 17 were upregulated and 19 were downregulated.

These ISDs were observed to be highly conserved, exhibiting only minor amino acid substitutions (see **Tables 1**). This high degree of preservation, along with the retention of both fusogenic and immunosuppressive potential across these differentially expressed HERVs, suggests that these domains may remain functionally active and underscores their likely relevance to BLEICS cell biology. Specifically, **Table 2** provides a comprehensive list of all significantly upregulated and downregulated HERVs that contain an ISD. For each differentially expressed HERV, the table includes key descriptive annotations to facilitate interpretation, detailing the element's subgenomic structure, the precise ISD sequence itself, the best-matching reference elements, and its corresponding HERV supergroup classification.

This focused presentation highlights HERV elements with a potential functional role in modulating the host's immune response via the ISD, offering insight into their possible contributions to the studied biological process.

Altogether, the global transcriptomic, deconvolution, and mechanistic analyses support a model in which the BLEICS cells represent a unique, pathologically reprogrammed B-cell population that has acquired myeloid-like functions. Specifically, CIBERSORTx deconvolution reveals a stable B cell-myeloid hybrid subpopulation defined by the co-expression of terminal B-cell effector markers (Plasma Cells/Memory B cells) with Antigen-Presenting Cell (APC) signatures (Activated Dendritic Cells and M1 Macrophages). This striking phenotypic convergence suggests a B-DC or B-Macrophage fusion event, a mechanism notably supported by the high expression of Human Endogenous Retroviruses (HERVs) that encode known fusogenic envelope proteins.

**Figure 7**

**Tumoral crypt expansion is driven by CAF-BLEICS interactions and shaped by early NICA cell accumulation**

Aberrant crypt architecture is a common histological feature in colorectal cancer (CRC). Upon enzymatic dissociation of fresh tissues, tumor crypts were found to be up to three times larger than normal crypts (**Figure 7A-I** & **II**). This enlargement was strongly associated with stromal proliferation, particularly the expansion of cancer-associated fibroblasts (CAFs). As demonstrated in **Figure 3D-II**, this fibroblast overgrowth is driven by direct interaction with EBV-infected BLEICS cells, likely mediated by stathmin induction. These interactions contribute to progressive crypt deformation and expansion (**Figure 7A-III**).

**Proposed Model for BLEICS Cell-Mediated Colorectal Carcinoma Pathogenesis: An Evolutionary Hybrid Lineage**

We propose a multi-stage model where Colorectal Carcinoma (CRC) pathogenesis is initiated by the formation of BLEICS cells through a unique heterotypic fusion event driven by chronic viral infection in the tumor-initiating niche (**Figure 7B**).

1. NICA Cell Gatekeeping and Immune Recruitment

The process begins with Neuro-Immune Crypt-Associated (NICA) cells accumulating at the base of colonic crypts. The terminus (origin or established cell type) of the NICA cell is currently unknown, as no description exists for a cell of this kind. While provisionally considered neuroendocrine cells (due to ChgA expression), NICA cells exhibit a unique triple phenotype (neuronal, stemness, and immune) and possess the molecular machinery for innate immune functions. These NICA cells act as gateways for pathogen internalization, functionally mimicking M cells. Given their neuronal-like characteristics (e.g., TH, somatostatin) and broad receptor expression, they are plausible targets for neurotropic herpesviruses like Cytomegalovirus (CMV) (theoretically) and Epstein-Barr Virus (EBV). Upon

pathogen internalization, NICA cells express the chemokine SDF−1, establishing a gradient that attracts CXCR4-positive immune cells.

2. Viral-Driven Fusion and BLEICS Formation

The SDF-1 gradient recruits dendritic cells (DCs) and B-lymphocytes. The B-lymphocytes become EBV-infected, express monoclonal Igs, and upregulate Stathmin, an EBV-triggered activation marker. IHC shows Stathmin$^+$/CD45$^+$ cells migrating from lymphoid follicles. The subsequent generation of BLEICS cells involves a complex fusion event, which is mechanistically facilitated by the high expression of EBV-activated fusogenic Human Endogenous Retroviruses (HERVs) (including HERV-K). This fusion is supported by the B cell/myeloid hybrid signature of BLEICS cells observed via CIBERSORTx. Crucially, IHC visualizes Stathmin$^+$ CD45$^+$ immune cells fusing directly with the colonic epithelium, indicating a B-DC-Epithelial cell heterotypic fusion as the likely source of the highly plastic BLEICS hybrid.

3. Spatiotemporal Stromal Remodeling and Malignant Progression

The BLEICS cells, characterized by high CXCR4 expression, promote the permanent growth and proliferation of fibroblasts, driving their differentiation into muscle-type Cancer-Associated Fibroblasts (CAFs). This interaction results in the spatiotemporal remodeling and expansion of the stroma, contributing to a stroma-rich phenotype, which is clinically associated with poor outcomes in CRC. The final CRC tumor cells inherit and express key molecular features, including the neuronal marker Chromogranin A (ChgA), stem cell markers (LGR5, CD133, ASCL2), and immune markers (CD80, CD22). The observation that the majority of these markers are found in the original NICA cells strongly indicates that the CRC malignancy is the evolutionary product of this initial NICA-BLEICS cell transformation process, which also generates an immunosuppressive niche through viral contributions to immune senescence.

**Concluding Synthesis**: **CRC as an Evolutionary Hybrid Lineage**

The compelling convergence of molecular, phenotypic, and clinical data establishes a conclusive model for Colorectal Carcinoma (CRC) pathogenesis as a cellular descendant and evolutionary product of the NICA-BLEICS hybrid transformation. The process initiates with the Neuro-Immune Crypt-Associated (NICA) cell —a stem-like cell of neuro-endocrine origin

The compelling convergence of molecular, phenotypic, and clinical data establishes a conclusive model for Colorectal Carcinoma (CRC) pathogenesis as a cellular descendant and evolutionary product of the NICA-BLEICS hybrid transformation. The process initiates with the Neuro-Immune Crypt-Associated (NICA) cell—a functionally novel cell type that, while previously considered to be a neuroendocrine cell, uniquely expresses a neuronal phenotype (including ChgA and other neuronal molecules) and stemness. This cell is naturally provided with an innate immune molecular machinery. This cell acts as a pathogen gateway, similar to the M-cells, and a focus for chronic herpesvirus infection, particularly EBV (which is known to trigger the expression of Stathmin in lymphocytes), but without ruling out other neurotropic viruses like CMV. This viral presence drives the formation of the BLEICS intermediate via a B cell–DC–Epithelial cell heterotypic fusion, which is mechanistically facilitated by EBV-activated fusogenic HERVs, a well-known process. The resulting BLEICS hybrid simultaneously promotes malignancy through two critical mechanisms: first, the spatiotemporal remodeling of the stroma by inducing muscle-type Cancer-Associated Fibroblasts, which drives the stroma-rich phenotype associated with poor clinical outcomes in CRC; and second, the chronic viral presence provides a pro-senescence, dysfunctional local immune environment, which is also observed at the cellular immune composition and activation in patients' blood. Crucially, the malignant CRC tumor cells retain the entire hybrid lineage phenotype, expressing the majority of markers found in the original NICA cells, including neuronal (ChgA), stem cell (LGR5, ASCL2), and immune (CD80, CD22) components, firmly establishing CRC as a highly plastic, Neuro-Immune-Epithelial hybrid tumor resulting from viral-driven cell fusion.

## Discussion

For decades, pathologists have observed a shift in the evolution of colon tumors from a neuroendocrine to an epithelial phenotype[4,5]. Gene-expression studies across diverse tumor histologies reveal a persistent neuronal background[46]. These observations imply that many cancers may originate from cells bearing at least a neuroepithelial identity[16]. In addition, nearly all tumors display immunological traits, believed to arise through fusion events between effector immune cells and nascent tumor tissues[47]. In colorectal cancer, the highest incidence[48]—from the left colic flexure to the distal rectum— coincides with anatomical gradients of endocrine cells[49] and lymphoid follicles[50], suggesting that regions rich in both cell types are preferential sites for malignant transformation.

In this study, we identified a distinct population of argentaffin, balalaika-shaped cells distributed throughout the colonic crypt epithelium and subepithelial layers. These cells exhibit both neuronal and immunological features and have been designated Neuroimmune Crypt-Associated (NICA) cells. NICA cells co-express canonical neuronal markers—including tyrosine hydroxylase, peripherin, somatostatin, neuron-specific enolase, peptide YY, EPHA2, and chromogranin A—alongside antimicrobial proteins (Gp2, ZG16, and DEFA6) and the antiviral effector viperin. Their diverse repertoire of immunomodulatory molecules, such as TLR4, REG4, pIgR, TNFαIP2, CD22, CD54, CD80, and SDF-1, combined with their frequent localization near lymphoid follicles, positions them as key mediators of both innate and adaptive immune responses. A potential mechanism for the fusion of NICA cells with lymphocytes or other CXCR4-expressing cells may involve the context-dependent expression of SDF-1 by NICA cells. Intriguingly, NICA cells also express stemness-associated proteins, including CD133, LGR5, SLC3A3, ASCL1/ASCL2, β-catenin, and DCLK1—markers typically linked to neuronal development and progenitor cell identity—highlighting their dual neurogenic and stem-like character. Furthermore, early colorectal lesions show focal accumulation of NICA cells at or near crypt bases, implicating them in the earliest stages of malignant transformation.

As colorectal tumors progress, they shed overt neuroendocrine features and adopt a more epithelial phenotype[51], yet retain a broad spectrum of neuronal markers. This transition is driven in part by overexpression of the transcriptional repressor REST, which silences neuronal genes in non-neuronal cells[52]. In fact, REST is markedly upregulated in tumor crypts, where neuronal gene expression can decrease by up to 75%. It appears that in CRCs, REST is upregulated as a strategic adaptation, actively silencing neuronal identity in tumoral crypts. Since neuronal differentiation entails irreversible cell cycle exit, REST helps cancer cells evade terminal differentiation and maintain proliferation. This suggests tumor cells deliberately erase lineage programs to sustain growth, with REST acting as a gatekeeper against differentiation pathways that threaten tumor fitness.

Our findings reveal a previously unrecognized dual-layered control by REST over both transcription and translation. The discovery that HDAC inhibition via VPA selectively restores protein synthesis in REST-deficient cells uncovers a novel REST-dependent translational checkpoint, adding depth to our understanding of lineage regulation. Epigenetic modulation of REST acetylation patterns induces a pleiotropic activation of neuronal genes at subtoxic concentrations. Precise genetic ablation of the REST gene in colorectal carcinoma models using genome editing has revealed several of its downstream targets—chromogranin A, stathmin, choline acetyltransferase, synaptophysin and stemness markers LGR5 and DCLK1—with DCLK1 showing particularly stringent REST-dependent regulation, inducing a partial retro-differentiation. This indicates that, CRC cells possess a suppressed neuronal/neuroendocrine identity. TH re-expression, central to catecholamine biosynthesis, implies reactivation of a silenced lineage rather than acquisition of new traits. These findings position REST as a key regulator of cell identity in CRC. The selective recovery of ChgA and TH also points to a poised chromatin state, capable of reactivating lineage programs under permissive conditions.

However, it is unlikely that REST alone governs these phenotypic and genotypic shifts. Other transcription factors are likely involved in orchestrating the neuronal-to-epithelial plasticity. Indeed, most cancer cells, regardless of histological subtype, retain a latent neuronal-like transcriptional background, suggesting a conserved neurobiological imprint across malignancies. It remains to be determined whether this imprint constitutes a defining feature of cancer stem cell biology or reflects a residual trace of embryonic developmental programs. Therapeutically, targeting REST or its epigenetic network may reprogram tumor cells toward differentiation, undermining their proliferative advantage. In this context, the potential cytostatic activity of valproic acid in colorectal cancers warrants renewed investigation[52]. Clinically utilized for several decades, VPA possesses a well-established safety profile and is tolerated at relatively high doses, which are capable of achieving therapeutic plasma concentrations—potentially sufficient to elicit cytostatic effects.

Notably, oncoviruses like EBV and KSHV can hijack REST's co-repressors, potentially reactivating REST and silencing neuronal genes[53]. This may lead to loss of neuronal identity, cell cycle re-entry, and even malignant transformation. In parallel, we observed upregulation of HERV-derived envelope genes in tumor crypts. These protein-coding genes retain intact fusogenic and immunosuppressive domains[55], potentially contributing to cellular heterogeneity and facilitating immune evasion.

We initially sought to determine whether the distribution and quantity of peripheral immune cells correlate with malignancy, and whether such patterns might suggest underlying infectious processes. Peripheral immune profiling of colorectal carcinoma (CRC) patients revealed significant monocytosis, accompanied by subtly reduced B cell counts, which may imply a possible association with Epstein–Barr virus (EBV) infection. Unexpectedly, $CD4^+$ T cell populations were markedly diminished compared to controls without a history of malignancy. These lymphocytes also exhibited pronounced overexpression of CXCR4, a chemokine receptor known to mediate cellular migration. Although the precise mechanisms remain unclear, it is plausible that CXCR4-driven chemotaxis contributes to the relocation of these cells toward tumor sites[56]. One particularly noteworthy observation is the significant downregulation of CTLA-4 in both $CD4^+$ and $CD8^+$ T cell subsets. Given CTLA-4's critical role in immune checkpoint regulation, this reduction may indicate impaired immunosurveillance. These findings support the rationale for evaluating checkpoint blockade immunotherapies targeting CTLA-4 in CRC patients, especially in light of emerging evidence suggesting their potential efficacy in select subgroups. Collectively, these findings suggest that CRC patients may be at an immunological disadvantage. While this may not pose an immediate threat from an immunoprotective standpoint[57], it could reflect a selective decline in immune competence, warranting further investigation.

Interestingly, in a substantial fraction of primary colorectal cancer specimens, we observed the emergence of a distinct lymphoid population in culture exhibiting a squid-like morphology. These cells feature apical, petal-like projections clearly identifiable by both light and electron microscopy, and a posterior "neck" region laden with lipid droplets, which may attenuate microbial entry. This population displays chromosomal aberrations, demonstrates immortal growth *in vitro*, and expresses both B-cell surface markers and monoclonal immunoglobulins. Notably, EBV-derived genes are consistently detected despite the absence of viral particles on ultrastructural examination, indicating a latent herpesvirus infection. We have termed these cells BLEICS (B-Lymphocyte, EBV-Infected, Calamari-Shaped). Intriguingly, transcriptional profiling of BLEICS reveals signatures of activated dendritic cells, suggesting either profound immaturity or fusion-derived hybridization. Their gene expression reflects features of "trained immunity,"[58] including memory-like responses to viral, bacterial, and parasitic stimuli—although this infectious transcriptome may reflect mimicked signals rather than active infection.

BLEICS express high levels of CXCR4 and avidly adhere to and fuse with SDF-1–secreting fibroblasts, thereby markedly promoting their proliferation. The activation of fibroblasts by BLEICS cells may

provide mechanistic insight into the widely cited notion that cancer resembles a wound that fails to heal. Fusion between BLEICS and fibroblasts is likely facilitated by the overexpression of nearly intact human endogenous retrovirus (HERV) envelope proteins, which confer fusogenic potential and immunoevasive properties via intact immunosuppressive domains[55]. This interaction with fibroblasts may drive stromal expansion—a known correlate of tumor malignancy[59].

The presence of BLEICS-like cells in colorectal tumors may account for the frequent emergence of lymphoma-like malignancies observed in freshly xenografted tumors in immunodeficient mice, rather than the preservation of their original histological features, as reported in other tumor types[60]. Furthermore, stathmin-positive lymphocytes infiltrate emerging colorectal tumors, where subepithelial lymphoid follicles appear to serve as a reservoir for these stathmin-positive cells, underscoring a prominent tumor–follicle axis in colorectal neoplasia. In this regard, one of the major contributors to tumor cell heterogeneity is cell-cell fusion[61]. This fascinating and complex biological event plays a significant role in cancer progression by generating hybrid cells with novel malignant properties, such as increased metastatic potential and resistance to chemotherapy and radiation[62]. Evidence suggests that cancer cells can acquire markers typically restricted to immune cells. Therefore, it is plausible that immune cell characteristics observed in cancer cells arise from fusion events between immune cells and epithelial cells during tumor evolution. For example, it has been reported that primary tumors may originate from the fusion of myeloid cells with tissue-resident cells undergoing malignant process. One such case involved a female patient with renal cell carcinoma who had received a bone marrow transplant (BMT) from her cancer-free six-year-old brother, providing direct evidence of this event[63]. Moreover, leukocyte–cancer cell hybrids have been implicated in the metastatic process, indicating an immune gain by cancer cells[64,65].

BLEICS cells harbor EBV genes and express stathmin, a protein known to be inducible by EBV[66]. Similarly, EBV is capable of transactivating HERV-derived envelope genes[67]. Tumor crypts overexpress several HERV envelope proteins with intact immunosuppressive domains and fusogenic capacity, while downregulated HERV sequences remain minimal, collectively favoring expression of intact fusogens.

This study redefines the cellular landscape of colonic malignant transformation by unveiling two previously underappreciated contributors—NICA cells and BLEICS—drivers of colorectal carcinogenesis via distinct but converging mechanisms: epithelial transformation and stromal dysregulation. First, the neuroimmune identity of crypt-resident cells—historically misclassified as neuroendocrine—is clarified, positioning NICA cells as the plausible cellular origin of colorectal carcinoma. This paradigm shift challenges longstanding assumptions and necessitates a reevaluation of early oncogenic events in the colonic epithelium. The characterization of EBV-infected BLEICS cells, and their stromal remodeling capacity introduces a second axis of transformation. The disruptive activity of BLEIC cells on fibroblasts closely resembles (or recapitulates) the characteristics of pathological fibroproliferative conditions with known oncogenic potential. Notably, keloids—quasi-neoplastic lesions marked by aberrant fibroblast expansion and B-cell infiltration—serve as a compelling analog. Metadata analyses further reinforce this link, revealing a strong correlation between keloid formation and malignant transformation[68]. Similarly, cirrhosis exemplifies a fibroproliferative state that frequently precedes carcinoma, with B-cells implicated in disease progression.

Finally, it is necessary to re-evaluate the role of Epstein-Barr virus (EBV) in the malignant transformation[69] of colon tissues and advocate for the serious implementation of protective vaccination strategies[70]. In our view, such interventions could significantly reduce the incidence of colorectal tumors associated with EBV activity. Even acknowledging that EBV may require co-

infection with other helper viruses to complete the malignant transformation process, interrupting even one critical link in this chain could reduce the incidence of this tumor type.

**Data availability**

All data supporting this study as well as primary cells used are reasonably available upon request. Transcriptome data from this study was deposited in NCBI GEO under the accession number GSE303666.

**Acknowledgements**

This research was partly supported by a grant awarded by the Ruhr University Bochum under the FoRUM Program (AZ F860R-2016), as well as by institutional funds from the Marien Hospital Herne. We would like to thank Mrs. Kerstin Heise from the Institute of Cell Biology, University of Essen, for her support in the analysis of the variable regions of the immunoglobulin genes in BLEIC cells. We are also grateful to Mrs. Iris Over from the Institute of Human Genetics, Ruhr University Bochum, for her invaluable assistance with the karyotyping analysis. Furthermore, we thank Dr. Ulrik Stervbo from the Center for Translational Medicine with a focus on Immunology and Transplantation, Marien Hospital Herne, for his help in the phenotyping of BLEIC cells. We regret if we failed to mention relevant studies in the field published by others.

**Contributions**

David Díaz-Carballo (concept, study design and manuscript preparation). Adrien Noa-Bolaño, transcriptome analysis and analysis of HERVs sequences, establishment of primary cell cultures, bioinformatics, ICC, IHC, FACS, WBs, qPCR, and support in article preparation. Udo Rahner, ICC, IHC qPCR, WBs. Ali H. Acikelli, qPCR, WBs, Fontana-Masson staining, bioinformatics, FACS, and support in article preparation. Sahitya Saka, ICC, WBs. Jacqueline Klein, CRISPR/Cas9 experiments design and conduction, IHC, karyotyping, ICC. Sascha Malak, Critical reading. Anne Höppner, qPCR. Annabelle Kamitz, 3D reconstruction of images and graphic representation. Carla Casula, primary cell culture and organoids. Lalitha Kamada, transcriptome analysis, bioinformatics. Andrea Tannapfel, Histopathological analysis. Jens Christmann, NGS, mutation analysis. Sarah Albus, IHC. Enrico Fruth, SEM and TEM. Daniela Gerovska, transcriptome analysis of HERVs, bioinformatics. Marcos J. Araúzo-Bravo, transcriptome analysis of HERVs, bioinformatics. Metin Senkal, sample collection. Crista Ochsenfarth, molecular biology experiments. Andrea Tannapfel, histopathological diagnosis. Dirk Strumberg, analysis of the data and approval of the final version.

**Corresponding author**

Correspondence to David Díaz-Carballo.

## References

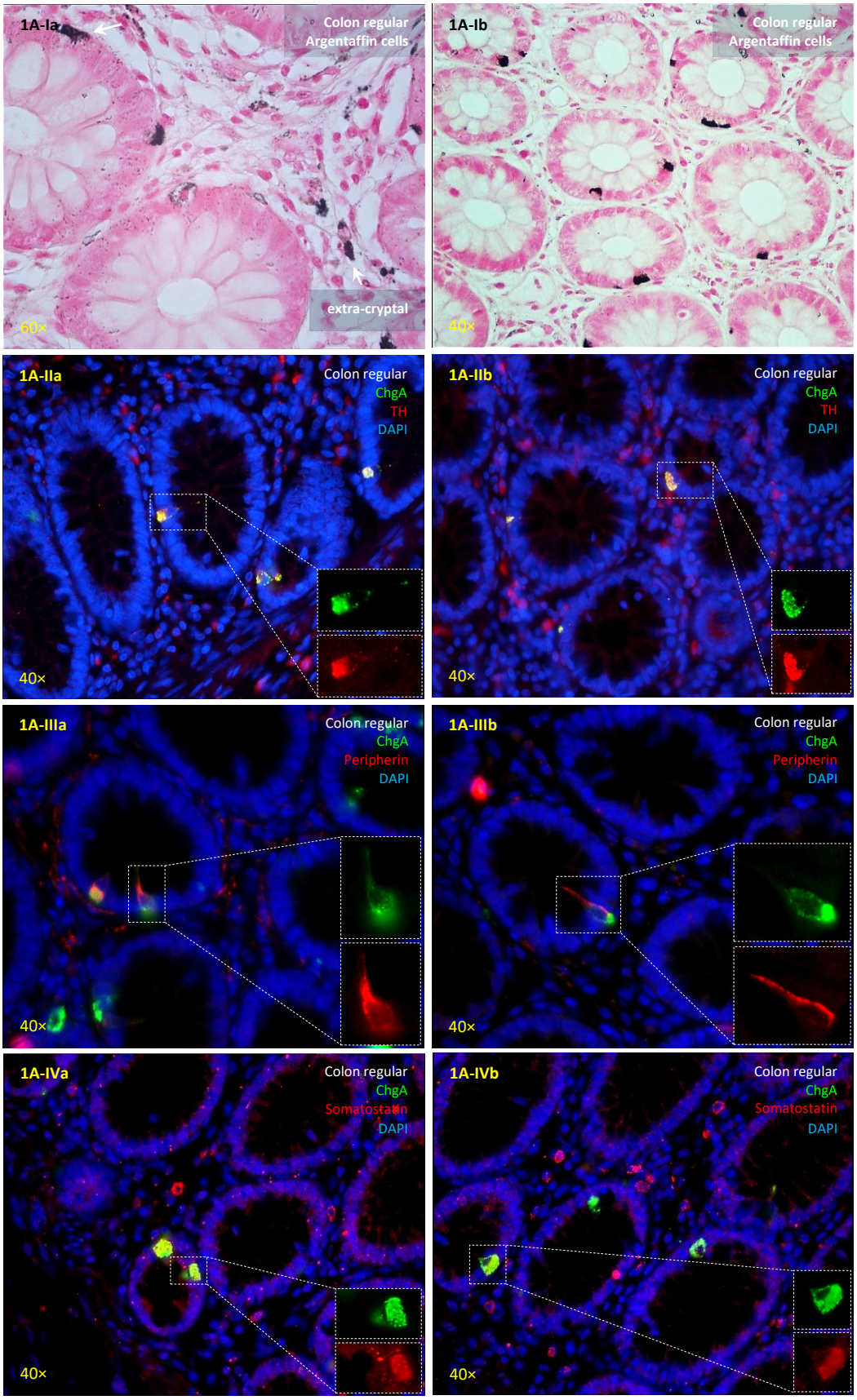

1. Atanga, R., Singh, V. & In, J. G. Intestinal Enteroendocrine Cells: Present and Future Druggable Targets. *Int. J. Mol. Sci.* **24**, (2023).
2. Gunawardene, A. R., Corfe, B. M. & Staton, C. A. Classification and functions of enteroendocrine cells of the lower gastrointestinal tract. *Int. J. Exp. Pathol.* **92**, 219–231 (2011).
3. Klöppel, G. Neuroendocrine Neoplasms: Dichotomy, Origin and Classifications. *Visc. Med.* **33**, 324–330 (2017).
4. Gulubova, M. & Vlaykova, T. Chromogranin A-, serotonin-, synaptophysin- and vascular endothelial growth factor-positive endocrine cells and the prognosis of colorectal cancer: an immunohistochemical and ultrastructural study. *J. Gastroenterol. Hepatol.* **23**, 1574–1585 (2008).
5. Ulich, T. R., Cheng, L., Glover, H., Yang, K. & Lewin, K. J. A colonic adenocarcinoma with argentaffin cells an immunoperoxidase study demonstrating the presence of numerous neuroendocrine products. *Cancer* **51**, 1483–1489 (1983).
6. Gehart, H. & Clevers, H. Tales from the crypt: new insights into intestinal stem cells. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 19–34 (2019).
7. Xie, G. *et al.* Zinc finger protein 277 is an intestinal transit-amplifying cell marker and colon cancer oncogene. *JCI Insight* **7**, (2022).
8. Yan, K. S. *et al.* The intestinal stem cell markers Bmi1 and Lgr5 identify two functionally distinct populations. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 466–471 (2012).
9. Huels, D. J. & Sansom, O. J. Stem vs non-stem cell origin of colorectal cancer. *Br. J. Cancer* **113**, 1–5 (2015).
10. Meyer, A. R., Brown, M. E., McGrath, P. S. & Dempsey, P. J. Injury-Induced Cellular Plasticity Drives Intestinal Regeneration. *Cell. Mol. Gastroenterol. Hepatol.* **13**, 843–856 (2022).
11. Verhagen, M. P. *et al.* Non-stem cell lineages as an alternative origin of intestinal tumorigenesis in the context of inflammation. *Nat. Genet.* **56**, 1456–1467 (2024).
12. Gharib, E. & Robichaud, G. A. From Crypts to Cancer: A Holistic Perspective on Colorectal Carcinogenesis and Therapeutic Strategies. *Int. J. Mol. Sci.* **25**, (2024).
13. Gustafsson, J. K. & Johansson, M. E. V. The role of goblet cells and mucus in intestinal homeostasis. *Nat. Rev. Gastroenterol. Hepatol.* **19**, 785–803 (2022).
14. Wang, Y., Song, W., Yu, S., Liu, Y. & Chen, Y.-G. Intestinal cellular heterogeneity and disease development revealed by single-cell technology. *Cell Regen. Lond. Engl.* **11**, 26 (2022).
15. Abrajano, J. J. *et al.* REST and CoREST Modulate Neuronal Subtype Specification, Maturation and Maintenance. *PLoS ONE* **4**, e7936 (2009).
16. Díaz-Carballo, D. *et al.* The neuroepithelial origin of ovarian carcinomas explained through an epithelial-mesenchymal-ectodermal transition enhanced by cisplatin. *Sci. Rep.* **14**, 29286 (2024).
17. Tokuyama, M. *et al.* ERVmap analysis reveals genome-wide transcription of human endogenous retroviruses. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 12565–12572 (2018).
18. SCHOFIELD, G. The argentaffin cells of the small intestine of the guinea-pig. *Acta Anat. (Basel)* **11**, 414–430 (1951).
19. Su, C. *et al.* Molecular basis of EphA2 recognition by gHgL from gammaherpesviruses. *Nat. Commun.* **11**, 5964 (2020).
20. Ahmed, E. M. *et al.* The Expression of Stem Cell Marker LGR5 and Its Coexpression with B-Catenin in Sporadic Colorectal Carcinoma and Adenoma: A Comparative Immunohistochemical Study. *Med. Kaunas Lith.* **59**, (2023).
21. Kim, H. *et al.* Differential epithelial and stromal LGR5 expression in ovarian carcinogenesis. *Sci. Rep.* **12**, 11200 (2022).
22. Ohno, H. Intestinal M cells. *J. Biochem. (Tokyo)* **159**, 151–160 (2016).
23. Yu, L.-G. The oncofetal Thomsen-Friedenreich carbohydrate antigen in cancer progression. *Glycoconj. J.* **24**, 411–420 (2007).
24. Yuan, M. *et al.* Comparison of T-antigen expression in normal, premalignant, and malignant human colonic tissue using lectin and antibody immunohistochemistry. *Cancer Res.* **46**, 4841–4847 (1986).
25. Bergström, J. H. *et al.* Gram-positive bacteria are held at a distance in the colon mucus by the lectin-like protein ZG16. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 13833–13838 (2016).
26. Ciesielska, A., Matyjek, M. & Kwiatkowska, K. TLR4 and CD14 trafficking and its influence on LPS-induced pro-inflammatory signaling. *Cell. Mol. Life Sci. CMLS* **78**, 1233–1261 (2021).
27. Wan, X., Wang, W., Zhu, J. & Xiao, Y. Antibacterial peptide Reg4 ameliorates Pseudomonas aeruginosa-induced pulmonary inflammation and fibrosis. *Microbiol. Spectr.* **12**, e0390523 (2024).

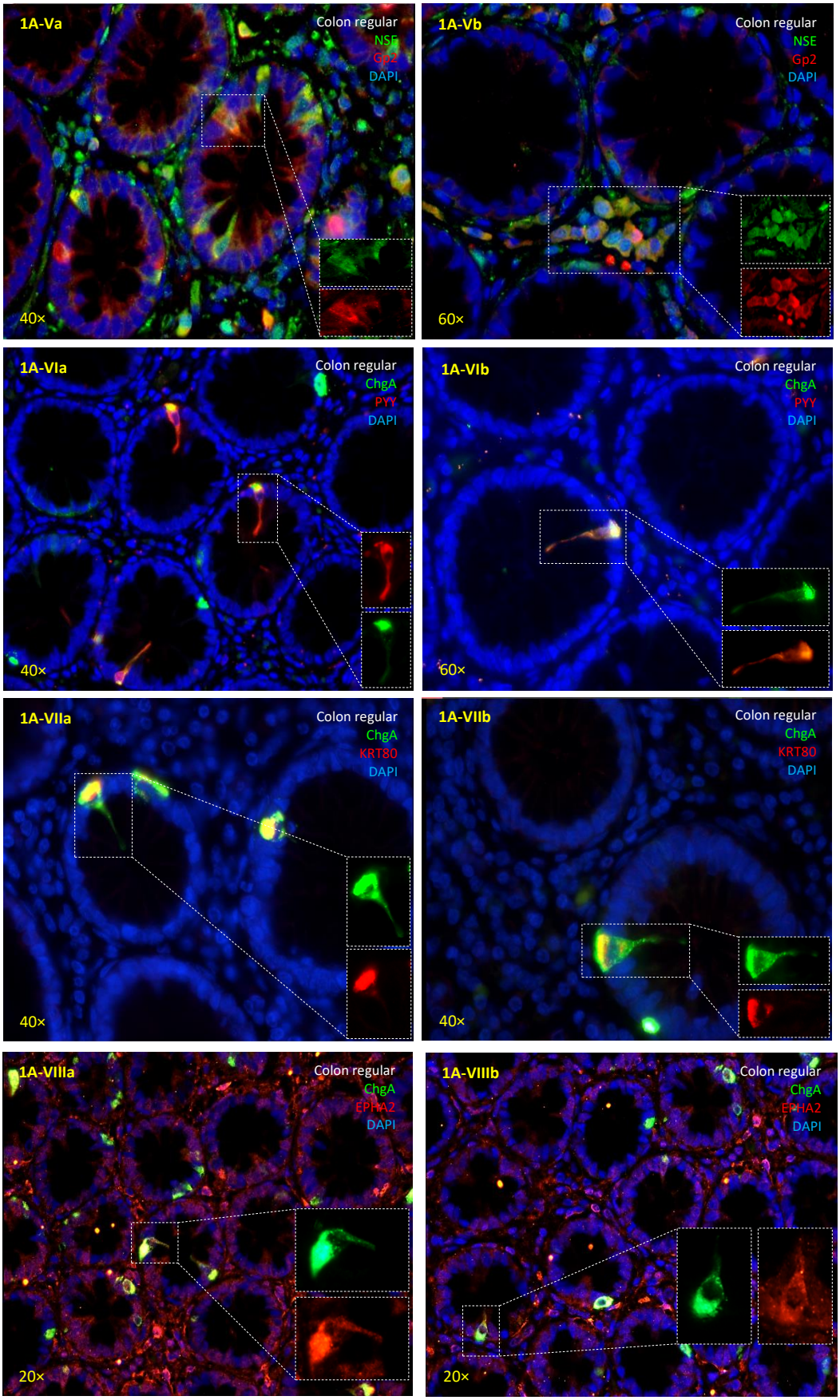
28. Shukla, S. D. *et al.* The main rhinovirus respiratory tract adhesion site (ICAM-1) is upregulated in smokers and patients with chronic airflow limitation (CAL). *Respir. Res.* **18**, 6 (2017).
29. Mattijssen, S. & Pruijn, G. J. M. Viperin, a key player in the antiviral response. *Microbes Infect.* **14**, 419–426 (2012).
30. Chairatana, P. & Nolan, E. M. Molecular basis for self-assembly of a human host-defense peptide that entraps bacterial pathogens. *J. Am. Chem. Soc.* **136**, 13267–13276 (2014).
31. D'amico, M. A., Ghinassi, B., Izzicupo, P., Manzoli, L. & Di Baldassarre, A. Biological function and clinical relevance of chromogranin A and derived peptides. *Endocr. Connect.* **3**, R45-54 (2014).
32. Clark, E. A. & Giltiay, N. V. CD22: A Regulator of Innate and Adaptive B Cell Responses and Autoimmunity. *Front. Immunol.* **9**, 2235 (2018).
33. Zastrow, A. *et al.* Regulation of Immune Function by CD22 on Antigen-Presenting Cells. *J. Immunol.* **212**, 1135_6075-1135_6075 (2024).
34. Nolan, A. *et al.* Differential role for CD80 and CD86 in the regulation of the innate immune response in murine polymicrobial sepsis. *PloS One* **4**, e6600 (2009).
35. Subauste, C. S., Waal Malefyt, R. & Fuh, F. Role of CD80 (B7.1) and CD86 (B7.2) in the Immune Response to an Intracellular Pathogen. *J. Immunol.* **160**, 1831–1840 (1998).
36. Hall, K., Blair Zajdel, M. E. & Blair, G. E. Defining the role of CD46, CD80 and CD86 in mediating adenovirus type 3 fiber interactions with host cells. *Virology* **392**, 222–229 (2009).
37. Graham, F. L. Transformation by and Oncogenicity of Human Adenoviruses. in *The Adenoviruses* (ed. Ginsberg, H. S.) 339–398 (Springer US, Boston, MA, 1984). doi:10.1007/978-1-4684-7935-5_9.
38. McDougall, J. K., Chen, L. B., Dunn, A. R. & Gallimore, P. H. Adenovirus — Integration and Oncogenicity. in *Integration and Excision of DNA Molecules* (eds Hofschneider, P. H. & Starlinger, P.) 59–70 (Springer Berlin Heidelberg, Berlin, Heidelberg, 1978). doi:10.1007/978-3-642-81203-3_9.
39. Poreba, E., Broniarczyk, J. K. & Gozdzicka-Jozefiak, A. Epigenetic mechanisms in virus-induced tumorigenesis. *Clin. Epigenetics* **2**, 233–247 (2011).
40. Vargiu, L. *et al.* Classification and characterization of human endogenous retroviruses; mosaic forms are common. *Retrovirology* **13**, (2016).
41. Lin, M.-J. & Lee, S.-J. Stathmin-like 4 is critical for the maintenance of neural progenitor cells in dorsal midbrain of zebrafish larvae. *Sci. Rep.* **6**, 36188 (2016).
42. Chiquetto Paracatu, L. *et al.* The Role of Stathmin-1 in the Regulation of Hematopoietic Stem Cells. *Blood* **140**, 5734–5735 (2022).
43. Hollie, N. & Asakrah, S. MALT lymphoma of the colon: a clinicopathological review. *J. Clin. Pathol.* **73**, 378–383 (2020).
44. Chalazonitis, A. Neurotrophin-3 as an essential signal for the developing nervous system. *Mol. Neurobiol.* **12**, 39–53 (1996).
45. Jepson, M. A. & Clark, M. A. The role of M cells in Salmonella infection. *Microbes Infect.* **3**, 1183–1190 (2001).
46. Regan, J. L. *et al.* Identification of a neural development gene expression signature in colon cancer stem cells reveals a role for EGR2 in tumorigenesis. *iScience* **25**, 104498 (2022).
47. Li, M., Basile, J. R., Mallya, S. & Lin, Y.-L. The impact and outcomes of cancer-macrophage fusion. *BMC Cancer* **23**, 497 (2023).
48. Md., A., Sheherin, T., Saikat, B. & Kabir, S. Clinicopathological and Histopathological Patterns of Colorectal Cancer: A Cross-Sectional Study at the National Institute of Cancer Research and Hospital. *Int. J. Clin. Oncol. Cancer Res.* **9**, 63–72 (2024).
49. Ogimi, T. *et al.* Distribution of Neuroendocrine Marker-Positive Cells in Colorectal Cancer Tissue and Normal Mucosal Tissue: Consideration of Histogenesis of Neuroendocrine Cancer. *Oncology* **97**, 294–300 (2019).
50. Langman, J. M. & Rowland, R. The number and distribution of lymphoid follicles in the human large intestine. *J. Anat.* **149**, 189–194 (1986).
51. Chen, Y., Liang, Y., Cao, L., Dong, X. & Sun, D. Neuroendocrine differentiation: a risk fellow in colorectal cancer. *World J. Surg. Oncol.* **21**, (2023).
52. Abrajano, J. J. *et al.* REST and CoREST Modulate Neuronal Subtype Specification, Maturation and Maintenance. *PLoS ONE* **4**, e7936 (2009).
53. Giordano, F. *et al.* Valproic acid inhibits cell growth in both MCF-7 and MDA-MB231 cells by triggering different responses in a cell type-specific manner. *J. Transl. Med.* **21**, 165 (2023).
54. Pillai, V. S. *et al.* REST Is Restless in Neuronal and Non-Neuronal Virus Infections: An In Silico Analysis-Based Perspective. *Viruses* **17**, 234 (2025).

55. Grandi, N. & Tramontano, E. HERV Envelope Proteins: Physiological Role and Pathogenic Potential in Cancer and Autoimmunity. *Frontiers in microbiology* vol. 9 462 (2018).
56. Wald, O. *et al.* CD4+CXCR4highCD69+ T Cells Accumulate in Lung Adenocarcinoma. *J. Immunol.* **177**, 6983–6990 (2006).
57. Péron, J. *et al.* CD4 lymphopenia to identify end-of-life metastatic cancer patients. *European journal of cancer (Oxford, England: 1990)* vol. 49 1080–1089 (2013).
58. Dagenais, A., Villalba-Guerrero, C. & Olivier, M. Trained immunity: A "new" weapon in the fight against infectious diseases. *Front. Immunol.* **14**, (2023).
59. Smit, M. A. *et al.* Tumour-stroma ratio outperforms tumour budding as biomarker in colon cancer: a cohort study. *Int. J. Colorectal Dis.* **36**, 2729–2737 (2021).
60. Chen, K., Ahmed, S., Adeyi, O., Dick, J. E. & Ghanekar, A. Human Solid Tumor Xenografts in Immunodeficient Mice Are Vulnerable to Lymphomagenesis Associated with Epstein-Barr Virus. *PLoS ONE* **7**, e39294 (2012).
61. Duelli, D. & Lazebnik, Y. Cell-to-cell fusion as a link between viruses and cancer. *Nature reviews. Cancer* vol. 7 968–976 (2007).
62. Shultes, P. V., Weaver, D. T., Tadele, D. S., Barker-Clarke, R. J. & Scott, J. G. Cell-cell fusion in cancer: The next cancer hallmark? *Int. J. Biochem. Cell Biol.* **175**, 106649 (2024).
63. Yilmaz, Y., Lazova, R., Qumsiyeh, M., Cooper, D. & Pawelek, J. Donor Y chromosome in renal carcinoma cells of a female BMT recipient: visualization of putative BMT-tumor hybrids by FISH. *Bone Marrow Transplant.* **35**, 1021–1024 (2005).
64. Pawelek, J. M. Fusion of bone marrow-derived cells with cancer cells: metastasis as a secondary disease in cancer. *Chin. J. Cancer* **33**, 133–139 (2014).
65. Laberge, G. S., Duvall, E., Haedicke, K. & Pawelek, J. Leukocyte¯Cancer Cell Fusion-Genesis of a Deadly Journey. *Cells* **8**, (2019).
66. Baik, S. Y. *et al.* Identification of stathmin 1 expression induced by Epstein–Barr virus in human B lymphocytes. *Cell Prolif.* **40**, 268–281 (2007).
67. Evans, E. F., Saraph, A. & Tokuyama, M. Transactivation of Human Endogenous Retroviruses by Viruses. *Viruses* **16**, 1649 (2024).
68. Ying-Yi Lu, Hung-Pin Tu, Chieh-Hsin Wu, Chien-Hui Hong, Kuo-Chia Yang, Hui-Ju Yang, Kee-Lung Chang, and Chih-Hung Lee. Risk of cancer development in patients with keloids. *Scientific Reports* **11**, 9390 (2021).
69. Bauer, M., Jasinski-Bergner, S., Mandelboim, O., Wickenhauser, C. & Seliger, B. Epstein–Barr Virus—Associated Malignancies and Immune Escape: The Role of the Tumor Microenvironment and Tumor Cell Evasion Strategies. *Cancers* **13**, 5189 (2021).
70. Cai, J. *et al.* Prophylactic and Therapeutic EBV Vaccines: Major Scientific Obstacles, Historical Progress, and Future Direction. *Vaccines* **9**, 1290 (2021).

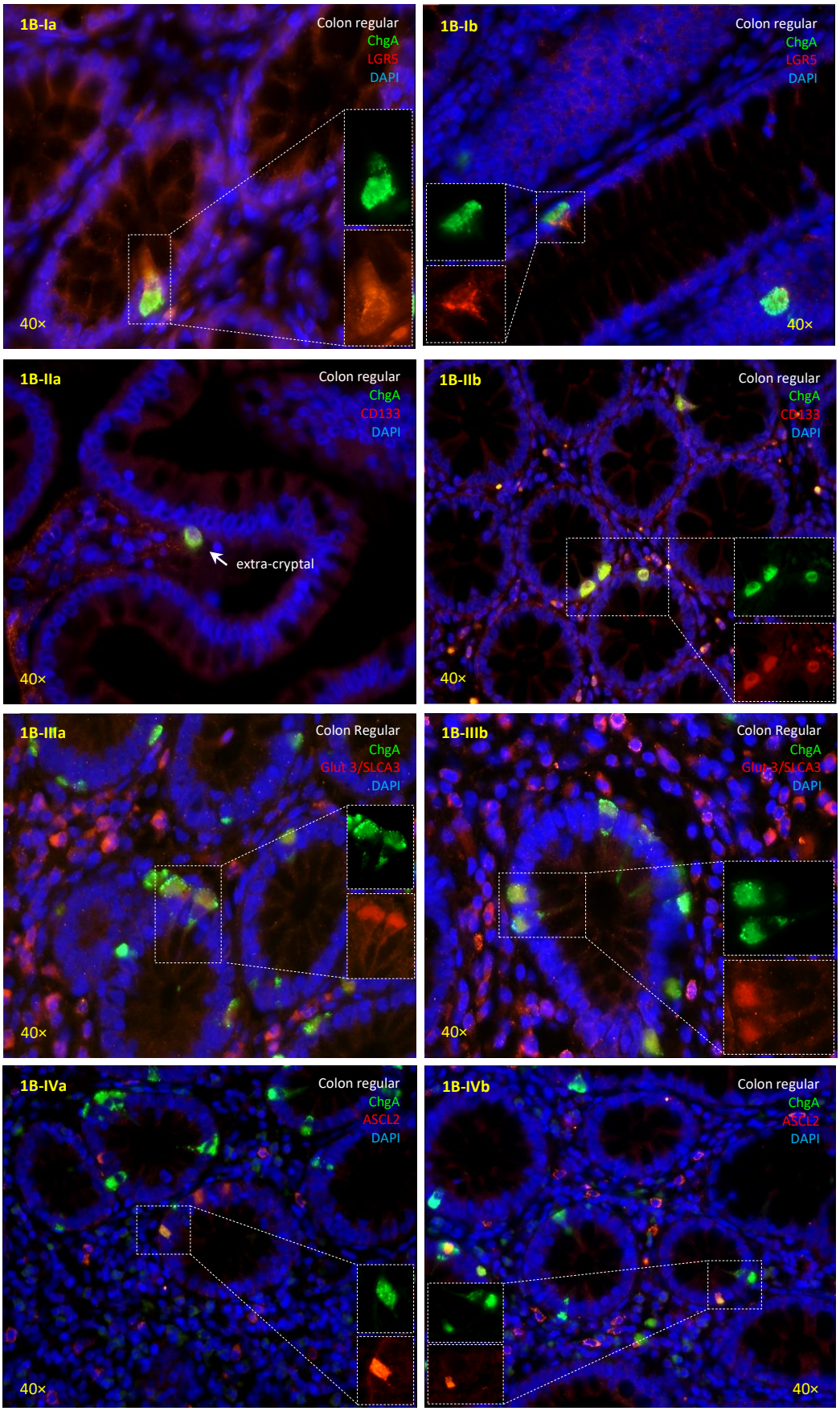
**Figure 1**:

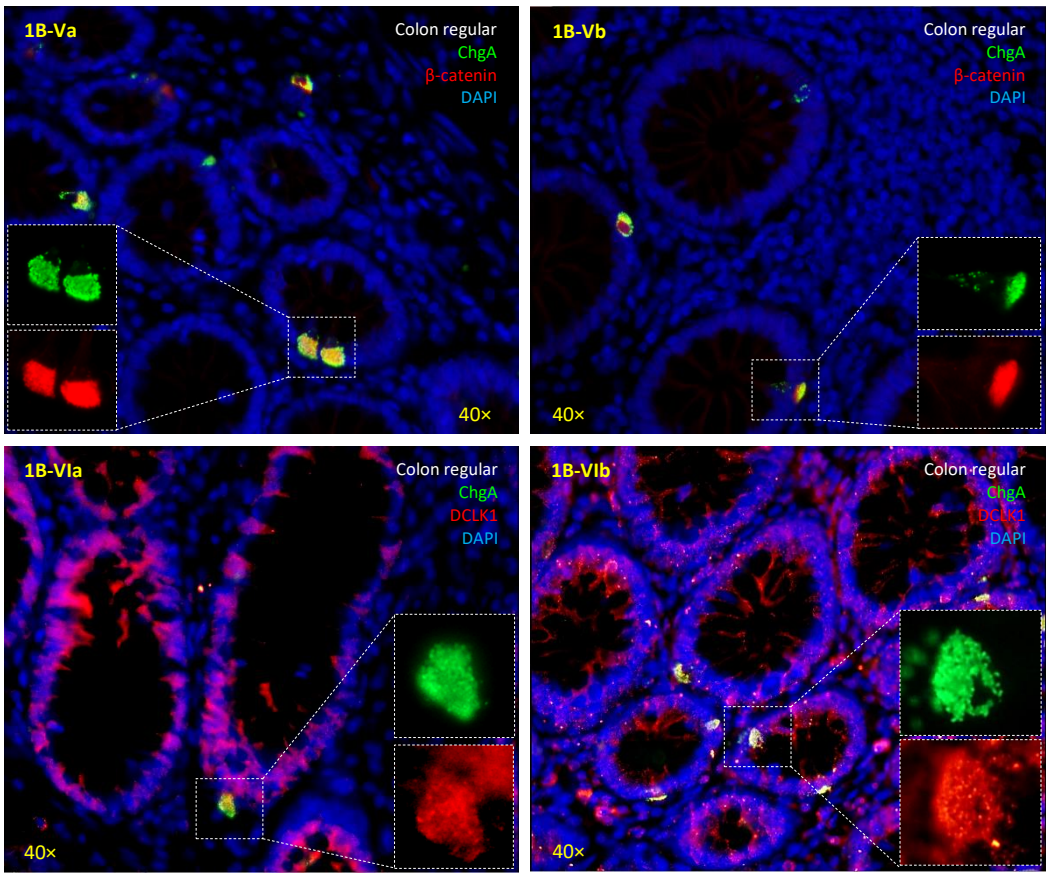
**Panel A**: Argentaffin cells with neuronal phenotype are observed in colonic crypts and the subepithelial *Lamina propria* regions



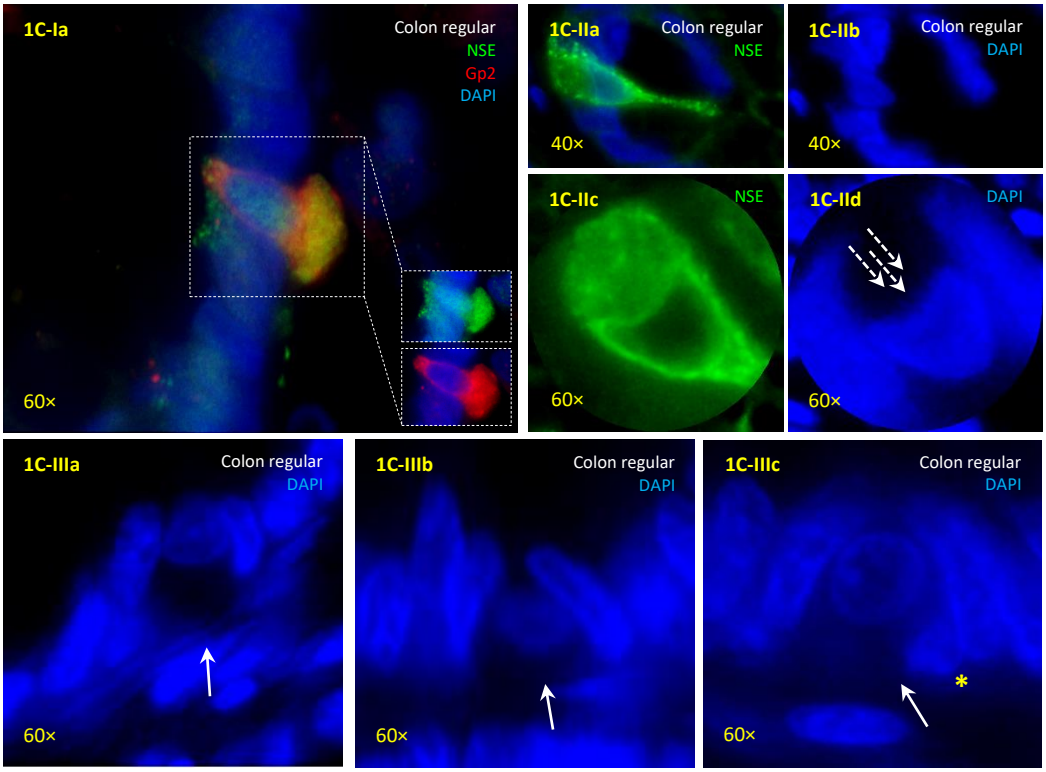

1A-Va
Colon regular
NSE
Gp2
DAPI
40×
1A-Vb
Colon regular
NSE
Gp2
DAPI
60×
1A-VIa
Colon regular
ChgA
PYY
DAPI
40×
1A-VIb
Colon regular
ChgA
PYY
DAPI
60×
1A-VIIa
Colon regular
ChgA
KRT80
DAPI
40×
1A-VIIb
Colon regular
ChgA
KRT80
DAPI
40×
1A-VIIIa
Colon regular
ChgA
EPHA2
DAPI
20×
1A-VIIIb
Colon regular
ChgA
EPHA2
DAPI
20×

**Panel B**: Argentaffin, Balalaika-shaped cells, which are associated to colonic crypts revealed a neuronal and stemness phenotype.



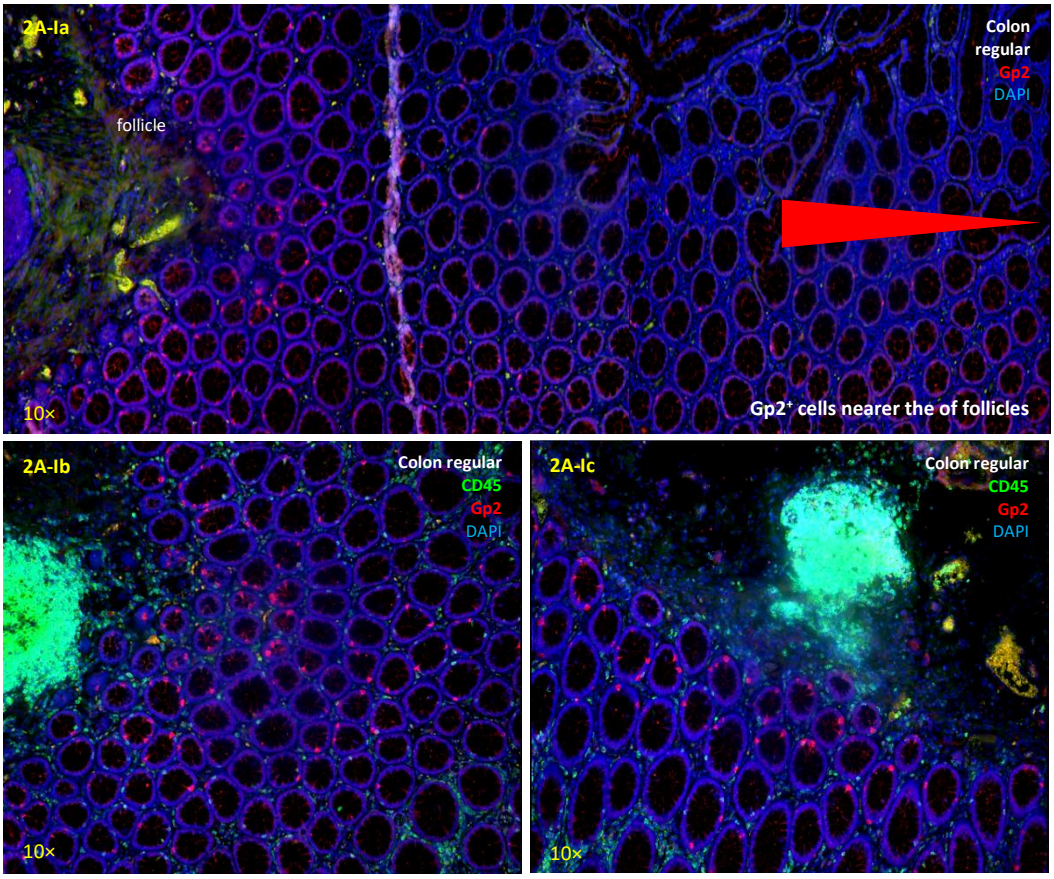



**<u>Panel C</u>**: Cells with both neuronal and stemness phenotype localized to the *Lamina propria* accommodate into the crypt structures

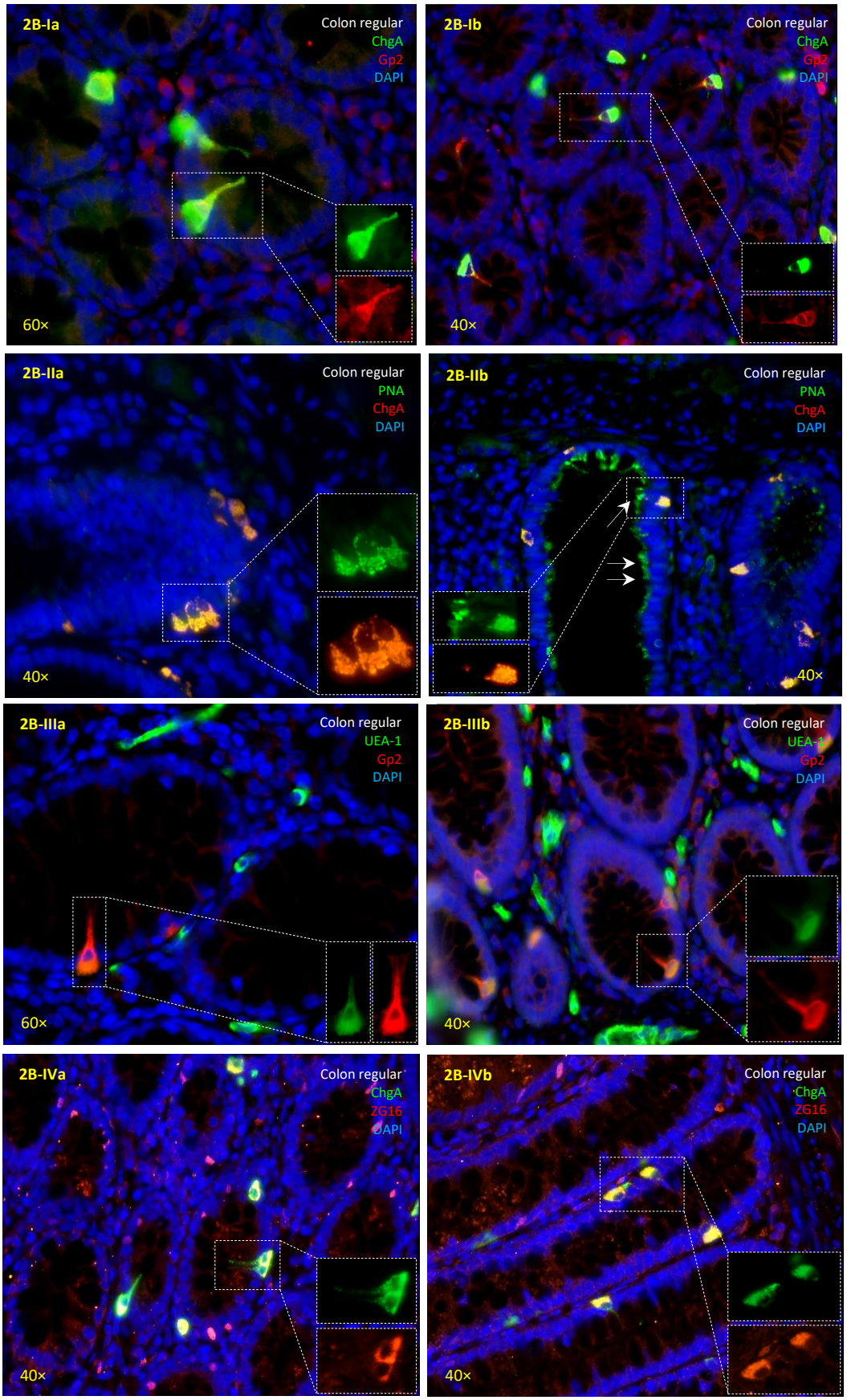


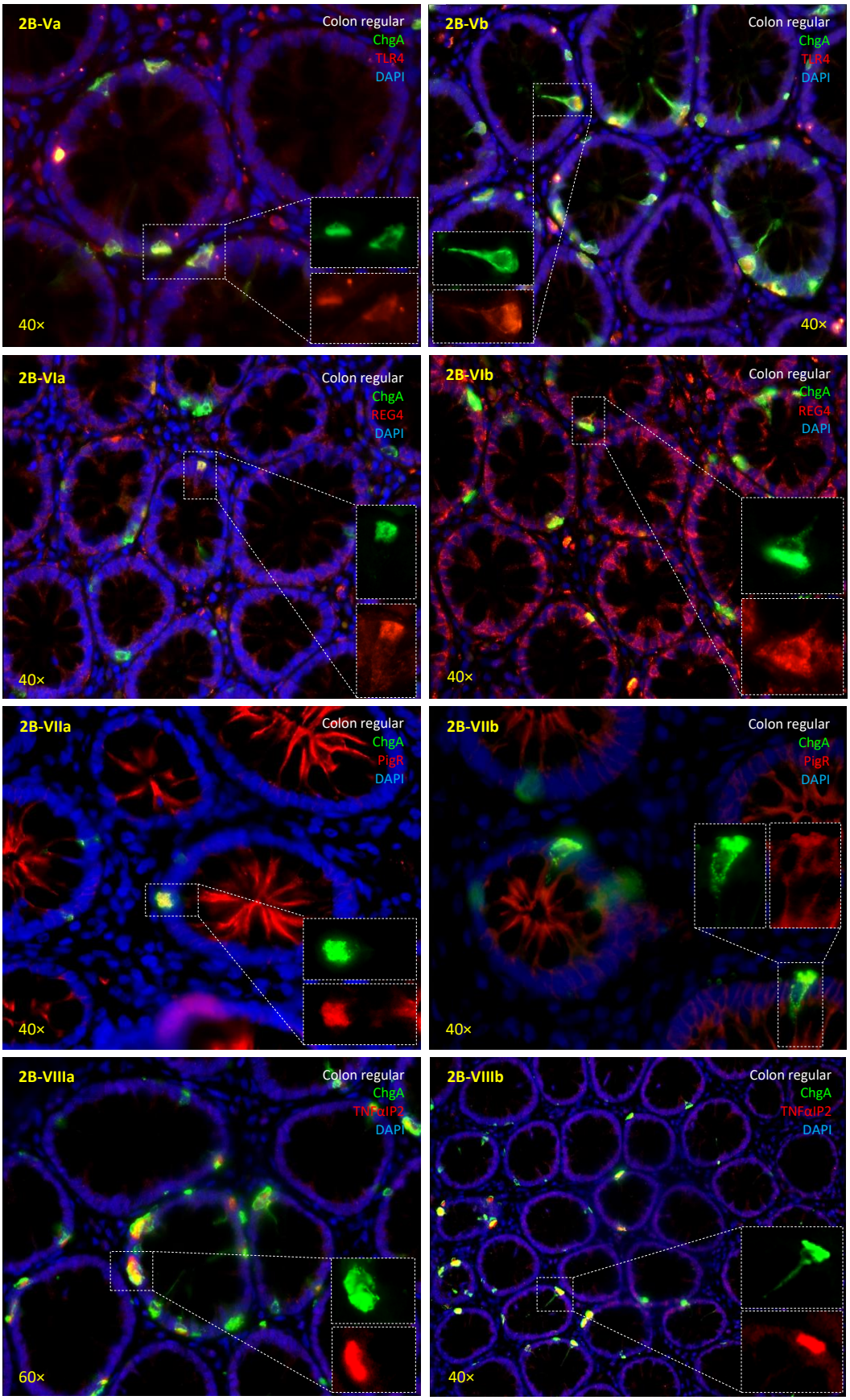
**<u>Figure 1</u>**: **Characterization of an argentaffin cell type with stemness phenotype in human regular colonic tissues**. **Panel A**: Argentaffin cells with a neuronal phenotype are observed in colonic crypts and the subepithelial *Lamina propria regions*. In the colonic crypts as well as in the subepithelial regions, diverse argentaffin cells are labeled using the Fontana-Masson staining (**Figures 1A**-**Ia** & **Ib**). The cells localized to the crypts are pyramidal or Balalaika-shaped, characterized by the presence of intracellular granules, mainly polarized in their basal areas. The Balalaika-shaped cells revealed a neuronal phenotype through the expression of Chromogranin A (ChgA) and Tyrosine hydroxylase (TH) (**Figures 1A**-**IIa** & **IIb**), and the type III intermediate filament Peripherin, which is mainly expressed in neurons of the PNS (**Figures 1A**-**IIIa** & **IIIb**). Somatostatin, a peptide hormone that plays a crucial role in regulating the endocrine system, is also expressed by Balalaika-shaped cells (**Figures 1A**-**Iva** & **IVb**). Neuron-specific enolase (NSE), which is predominantly found in neurons and neuroendocrine cells, is expressed by Balalaika-shaped cells (**Figures 1A**-**Va** & **Vb**). The hormone PYY, which is primarily secreted by L-type endocrine cells, is also expressed by Balalaika-shaped cells (**Figures 1A**-**Via** & **VIb**). The structural protein KRT80, known to be characteristic of epithelial cells, is particularly expressed by Balalaika-shaped cells and therefore associated with this neuronal cell type (**Figures 1A**-**VIIa** & **VIIb**). The Ephrin type-A receptor 2 (EPHA2), which plays important roles in neuronal development and serves as an entry receptor for EBV, is also expressed by these cells (**Figures 1A**-**VIIIa** & **VIIIb**). **Panel B**: Argentaffin, Balalaika-shaped cells associated with colonic crypts revealed a neuronal and stemness phenotype. The well-known markers for adult stem cells, LGR5 (**Figures 1B**-**Ia** & **Ib**) and CD133 (**Figures 1B**-**IIa** & **IIb**), are also expressed by Balalaika-shaped cells. GLUT3, a glucose transporter identified as a specific marker for embryonic stem cells, is expressed by Balalaika-shaped cells (**Figures 1B**-**IIIa** & **IIIb**). The transcription factor ASCL2, which plays a role in the regulation of stem cell maintenance and differentiation, particularly in the intestinal epithelium, is also expressed in these cells (**Figures 1B**-**IVa** & **IVb**). The key component of the Wnt signaling pathway, β-catenin, which is crucial for maintaining stem cell pluripotency and regulating cell fate decisions, is also expressed by Balalaika-shaped cells found in colonic crypts (**Figures 1B**-**Va** & **Vb**). DCLK1 is associated with self-renewal and pluripotency in CRCs. Some Balalaika-shaped cells as well localized to the Lamina propria express DCLK1 (**Figures 1B**-**VIa** & **VIb**). **Panel C**: Cells with both a neuronal and stemness phenotype localized to the *Lamina propria* accommodate into the crypt structures (**Figure 1C**-**Ia**). It appears that the cells settle into the crypts, as judged by the morphological movement shapes of their nuclei, which seem to adopt a hyperbolic shape and their reorganization in the colonic crypts (**Figures 1C**-**IIa**-**d**). Some of the cells already accommodated into the crypts' line (**Figures 1C**-**IIIa**-**c**) allow other cells to dock in the basolateral pole (**Figure 1C**-**IIIc**, yellow asterisk and SI: NICAS basolateral Pocket). IHC analysis was performed using paraffin sections from regular colon tissues. Experiments are representative of several analyses.

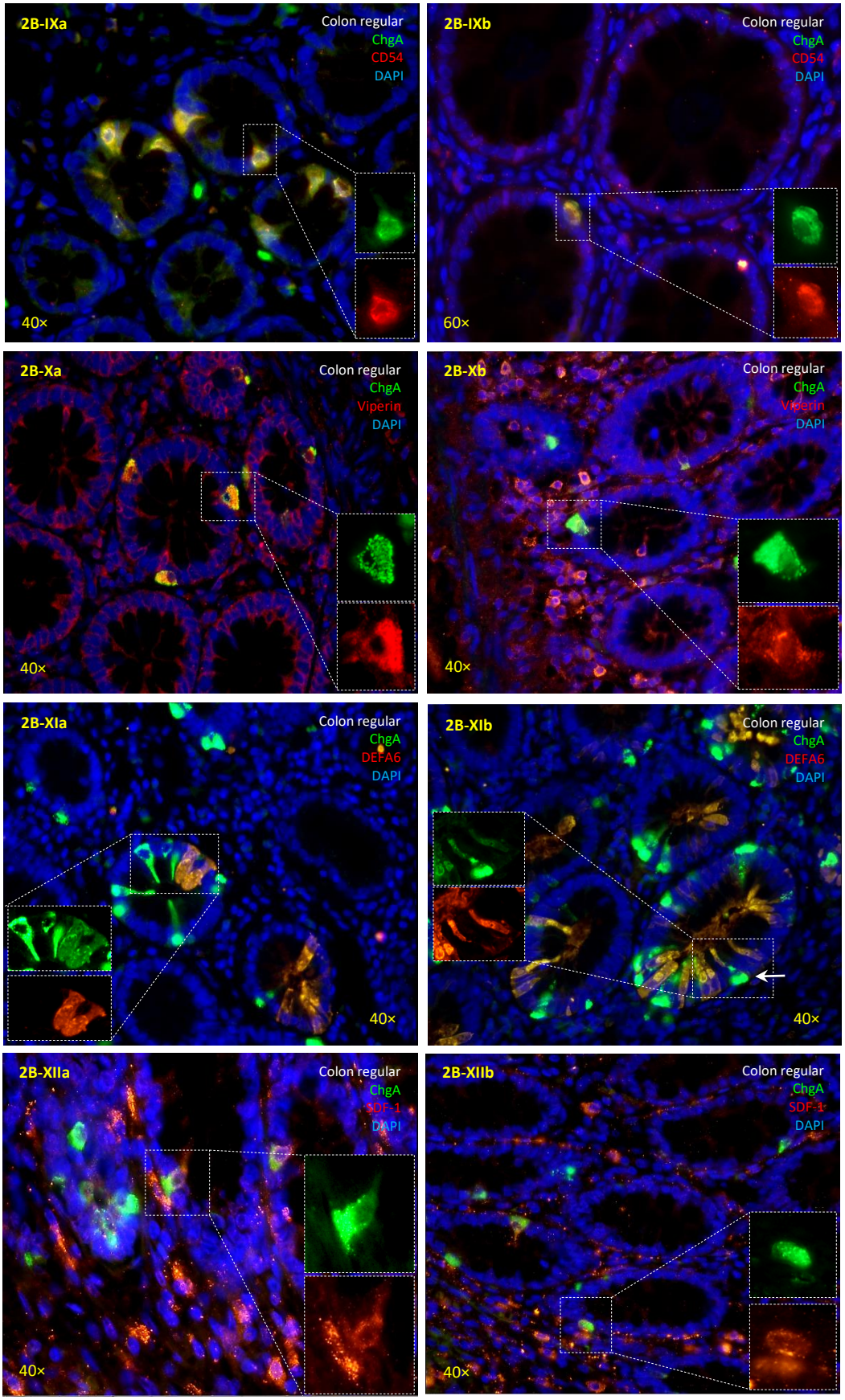
**Panel A:** The interconnection of NICAS cells to the follicle-associated epithelium (FAE)



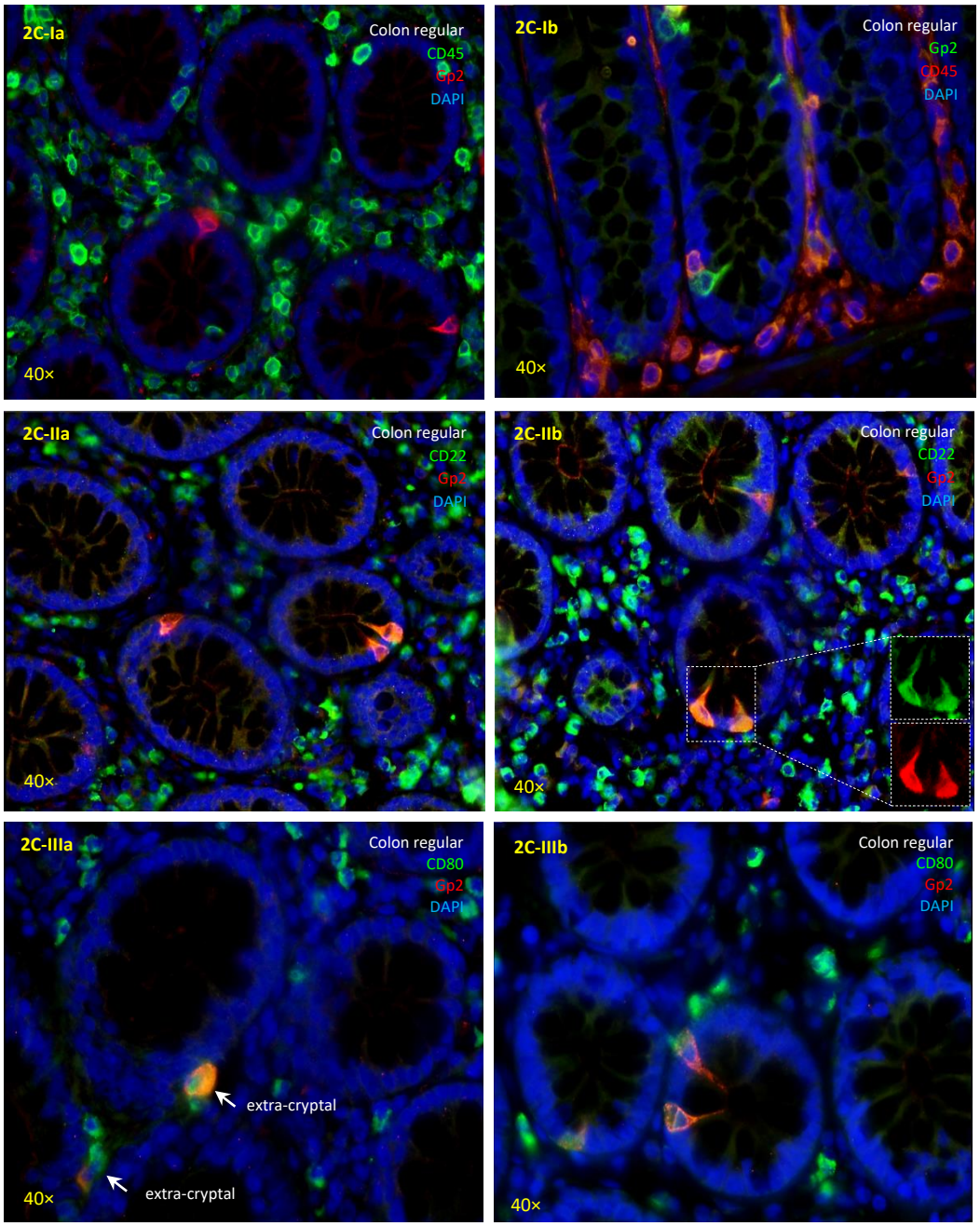
**Panel B**: Signature of the innate immunity NICAS found in regular colon tissues




2B-Va
Colon regular
ChgA
TLR4
DAPI
40×
2B-Vb
Colon regular
ChgA
TLR4
DAPI
40×
2B-VIa
Colon regular
ChgA
REG4
DAPI
40×
2B-VIb
Colon regular
ChgA
REG4
DAPI
40×
2B-VIIa
Colon regular
ChgA
PigR
DAPI
40×
2B-VIIb
Colon regular
ChgA
PigR
DAPI
40×
2B-VIIIa
Colon regular
ChgA
TNFαIP2
DAPI
60×
2B-VIIIb
Colon regular
ChgA
TNFαIP2
DAPI
40×


2B-IXa
Colon regular
ChgA
CD54
DAPI
40×
2B-IXb
Colon regular
ChgA
CD54
DAPI
60×
2B-Xa
Colon regular
ChgA
Viperin
DAPI
40×
2B-Xb
Colon regular
ChgA
Viperin
DAPI
40×
2B-XIa
Colon regular
ChgA
DEFA6
DAPI
40×
2B-XIb
Colon regular
ChgA
DEFA6
DAPI
40×
2B-XIIa
Colon regular
ChgA
SDF-1
DAPI
40×
2B-XIIb
Colon regular
ChgA
SDF-1
DAPI
40×

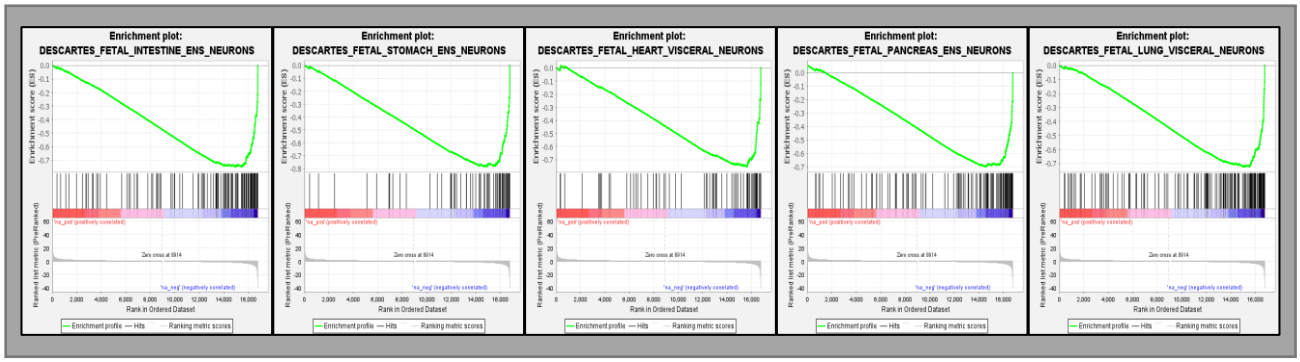
**Panel C**: Signature of the adaptive immunity in NICAS cells in regular colon tissues



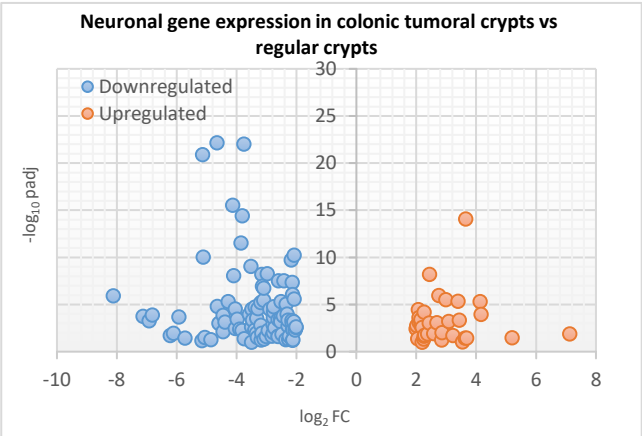
**<u>Figure 2</u>**: **A The Neuro-Immune Crypt-Associated** (**NICA**) **cell type bring together the innate and acquired immune systems**. **Panel A**: The interconnection of NICAS cells to the follicle-associated epithelium (FAE). Immunohistochemical examination revealed that Balalaika-shaped argentaffin cells are abundantly observed near the subepithelial follicles of regular colon tissues in a gradient (red wedge, **Figures 2A-Ia**, **2A-Ib/Ic**). **Panel B**: Signature of the innate immunity in NICAS cells found in regular colon tissues. IHC analysis of balalaika-shaped argentaffin cells shows that they are recognized by the FITC-conjugated lectins PNA (**Figures 2B-Ia** & **Ib**) and UEA-1 (**Figures 2B-IIa/IIb**), which specifically bind to β-D-Gal(1-3)-D-galNAc and L-fucose residues, respectively. In addition, balalaika-shaped cells express Gp2 (ZAP75), which acts as a receptor specific to type-I-piliated bacteria (**Figures 2B-IIIa** & **IIIb**), and ZG16 (Zymogen Granule Protein 16), which binds different carbohydrate moieties as well as peptidoglycans (**Figures 2B-Iva** & **IVb**). Furthermore, balalaika-shaped argentaffin cells express TLR4, a pivotal receptor for recognizing LPS from gram-negative bacteria which is mainly expressed by cells of myeloid origin (**Figures 2B-Va** & **Vb**). REG4 (Regenerating Family Member 4), a protein involved in inflammatory responses, is also detected in balalaika-shaped cells (**Figures 2B-VIa** & **VIb**). PigR (Polymeric Immunoglobulin Receptor), a transmembrane protein that facilitates the transcytosis of polymeric immunoglobulins across mucosal epithelia, is found in the apical portion of enterocytes as well as in the body of balalaika-shaped cells (**Figures 2B-VIIa** & **VIIb**). TNFaIP2 (Tumor Necrosis Factor Alpha-Induced Protein 2), involved in the cellular response to inflammation, is specifically expressed in balalaika-shaped cells (Figure 2B-VIIIa & VIIIb). CD54 (Intercellular Adhesion Molecule 1, or ICAM-1), which is primarily considered a part of the innate immune system, is also observed in these cells (**Figures 2B-IXa** & **IXb**). The antiviral protein viperin (**Figures 2B-Xa** & **Xb**) and the antibacterial protein DEFA6 (**Figure 2B-Xia** & **XIb** -white arrow) are expressed by Balalaika-shaped cells. The inflammatory chemokine SDF-1 (CXCL12), which is the ligand for fusin (CXCR4), is also expressed by balalaika argentaffin cells (**Figures 2C-XIIa** & **XIIb**). **Panel C**: NICAS cells display markers of adaptive immunity. Despite these immunological phenotypes, these cells do not express CD45, a pan-leukocyte marker characteristic of differentiated hematopoietic cells (**Figures 2C-Ia** & **Ib**). NICA cells express CD22, which is located on the surface of mature and immature B cells and functions as an inhibitory receptor for the B cell receptor (**Figures 2C-IIa** & **IIb**) and CD80, a member of the B7 family that is principally expressed in antigen-presenting cells (**Figures 2C-IIIa/IIIb**). The IHC analysis was performed using paraffin sections from regular colon tissues. The figures are representative of several analyses.

**Panel A**: Loss of neuronal features and immune modulation in tumor-derived crypts

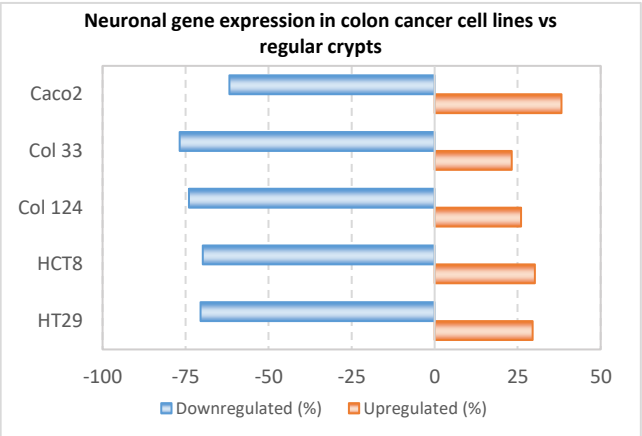
**3A-I**



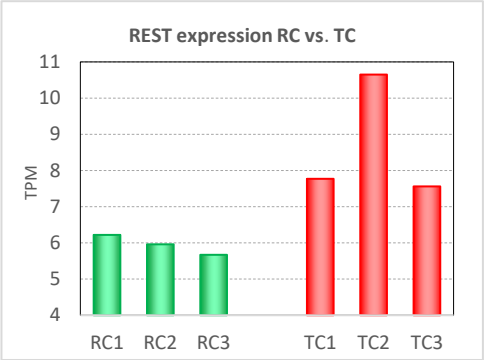

**3A-II**



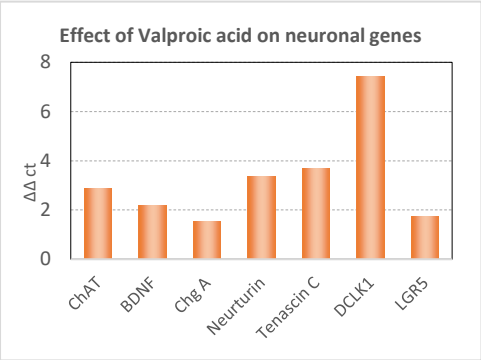

**3A-III**



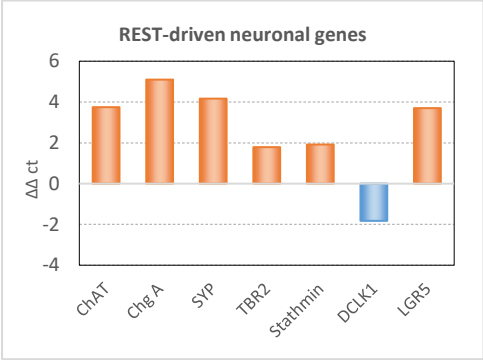

**3A-IV**



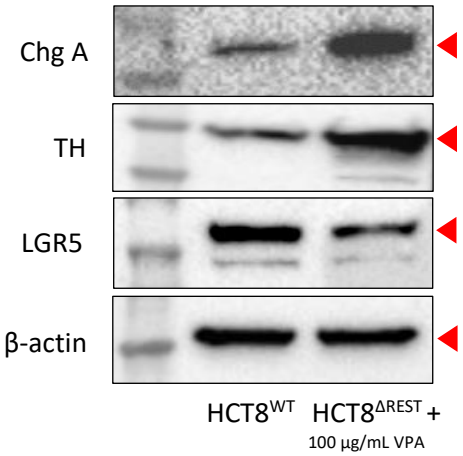

**3A-V**



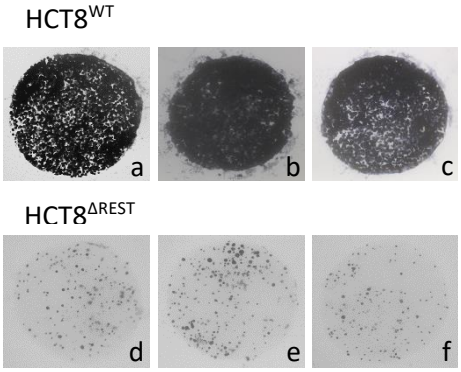

**3A-VI**



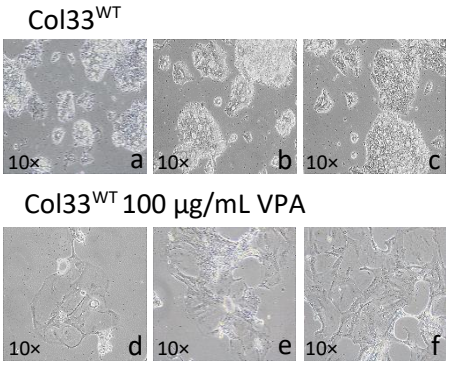

**3A-VII**



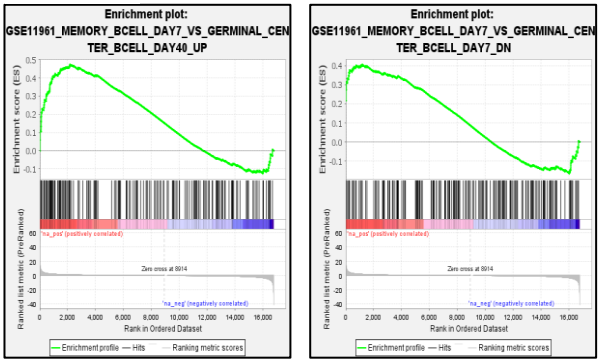

**3A-VIII**



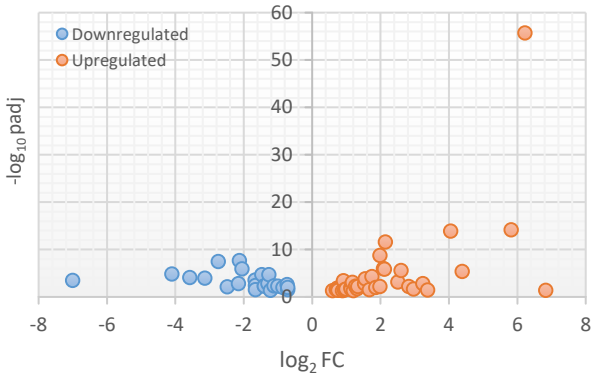

**3A-IX**



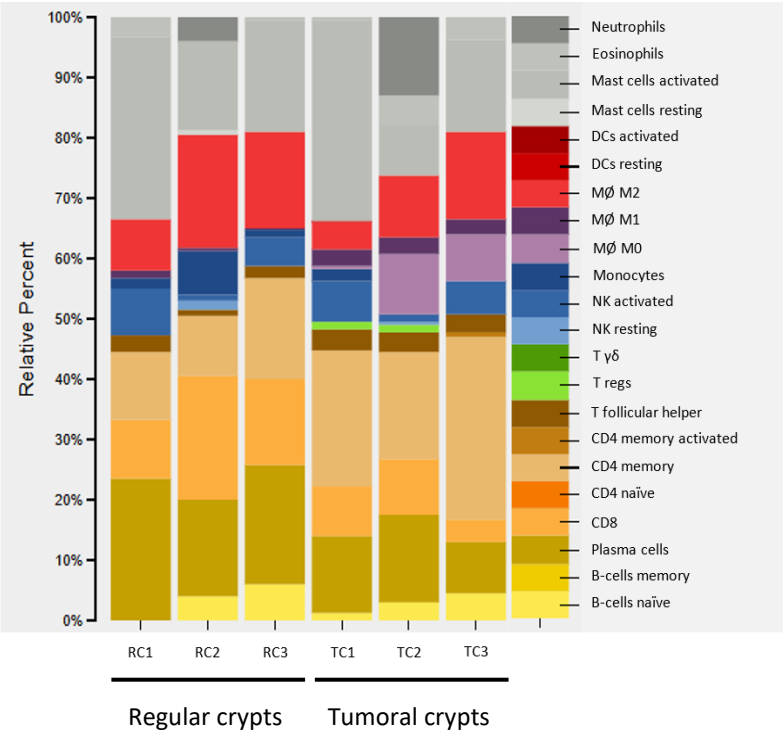
## 3A-VII




**Germinal Center B-lymphocyte related genes**




## 3A-VIII



**Panel B**: Regulation of human endogenous retroviruses (HERVs) in tumor-derived crypts

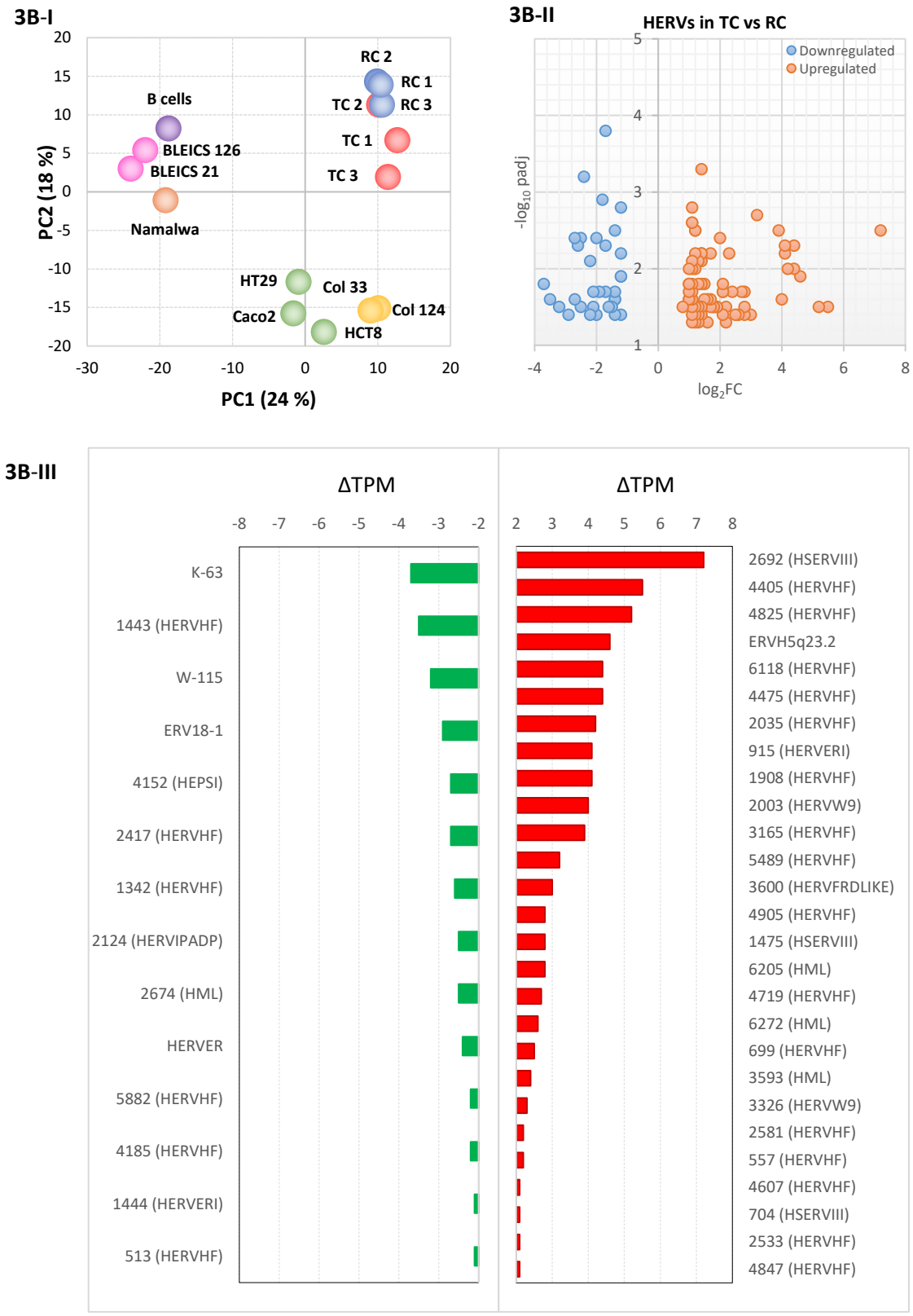
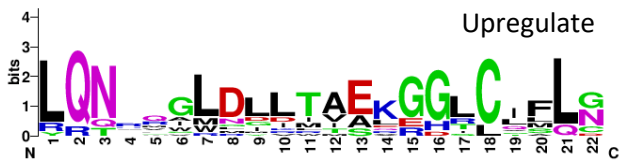


**Table 1**: Differentially expressed HERVs within colorectal crypts presenting ISD

| HERV ID | -LOG$_{10}$ (PVAL) | ΔTPM | SUBGENES | ISD | BESTREFRV | SUPER GROUP |
|---|---|---|---|---|---|---|
| 3542 | 1.8 | 1.1 | 5LTR PBS CA NC Prot RT IN SU TM PPT 3LTR | YQNRLvLDYLLv | Human endogenous retroviral DNA (4-1), complete re 92% -98 | HERVERI |
| 3593 | 1.7 | 2.4 | 5LTR CA NC RT IN SU TM PPT 3LTR | RQTVtWMGDkIMSLeHrLQMQC | Human endogenous retrovirus HERV-K10. 60% -2255 | HML |
| 6073 | 2 | 1 | 5LTR PBS IN SU TM PPT 3LTR | RQTVIWMrDRIiSLEHRLQmqC | Human endogenous retrovirus HERV-K(I) DNA, complete 64% -2205 | HML |
| 1805 | 1.5 | 1.9 | CA NC RT IN SU TM PPT 3LTR | LRQTVIWMGDRI | Human endogenous retrovirus HERV-K10. 61% -2110 | HML |
| 2048 | 1.6 | 1.1 | NC Prot RT SU TM PPT | LRQSVIWLGDwV | Homo sapiens tandemly repeated human endogenous re 62% -3890 | HML |
| 704 | 1.7 | 2.1 | 5LTR PBS NC Prot RT IN TM PPT 3LTR | LQNzmALnivTAAzGGTCAlLG | | HSERVIII |
| 510 | 1.8 | 1 | 5LTR PBS CA NC Prot RT IN SU TM 3LTR | LQNRwGLDLimAEKRdLCLsLG | Homo sapiens RGH2 gene, retrovirus-like element. 57% -1199 | HERVHF |
| 3165 | 2.5 | 3.9 | CA Prot RT IN TM 3LTR | LQNRRGLDLLtTEKGGsCLsLG | Homo sapiens human endogenous rv HML6 gag,pol,env 50% 307 | HERVHF |
| 906 | 1.3 | 1.2 | 5LTR PBS MA CA NC Prot RT IN SU TM PPT 3LTR | LQNRqsLDLLTAEKGGLCIFLN | Homo sapiens RGH2 gene, retrovirus-like element. 83% -22 | HERVHF |
| 5480 | 2.4 | 2 | 5LTR PBS MA CA NC Prot RT IN SU TM PPT 3LTR | LQNRqGLnLLTAEKGGLCIFLN | Homo sapiens RGH2 gene, retrovirus-like element. 80% -48 | HERVHF |
| 765 | 1.7 | 1.1 | 5LTR PBS NC Prot RT IN TM PPT 3LTR | LQNqmALDiiTtARGGTCslLG | | HSERVIII |
| 4475 | 2.3 | 4.4 | 5LTR PBS MA CA NC Prot RT IN TM PPT 3LTR | LQNLqGLDLLTAEERgLCIFLN | Homo sapiens RGH2 gene, retrovirus-like element. 73% -463 | HERVHF |
| 2518 | 1.3 | 1.1 | CA NC Prot RT IN TM PPT | LQNhRGLDLLTvEKGGLCtFLG | Human endogenous retrovirus, complete genome. 68% -1409 | HERVHF |
| 3698 | 2.8 | 1.1 | IN SU TM PPT 3LTR | LQNhRGLDLLTAEKGGLCIFLN | Homo sapiens RGH2 gene, retrovirus-like element. 77% -4627 | HERVHF |
| 4489 | 2.5 | 1.2 | 5LTR PBS MA CA NC Prot RT IN SU TM 3LTR | LQNgqGrDLLTAEKGGLCIFLN | Human endogenous retrovirus, complete genome. 84% -159 | HERVHF |
| 3016 | 1.5 | 1.9 | 5LTR PBS CA NC Prot RT IN SU TM PPT 3LTR | LQNcqGLDmLmAAQGGlCLALd | | HERVFRDLIKE |
| 2643 | 1.6 | -1.4 | 5LTR PBS CA IN TM PPT 3LTR | YQNRLALDYLLA | Human endogenous retroviral DNA (4-1), complete re 68% -1020 | HERVERI |
| 868 | 2.4 | -2 | 5LTR PBS MA CA NC Prot RT IN TM PPT 3LTR | wgNRiALDmLLA | Homo sapiens human endogenous retrovirus HERV-P4.6 73% -177 | HERVIPADP |
| 2124 | 1.5 | -2.5 | 5LTR PBS MA CA NC IN TM | weNRiALniiLA | Homo sapiens human endogenous retrovirus HERV-P4.6 75% -784 | HERVIPADP |
| 2754 | 2.8 | -1.2 | CA NC Prot RT SU TM PPT | LQNRzGLDLLTAEKGGLCtFLG | Human endogenous retrovirus pHE.1 (ERV9). 85% -503 | HERVW9 |
| 2674 | 2.4 | -2.5 | Prot IN SU TM PPT | hQTVpzMGDzIiNLEHRiQIQC | Human endogenous retrovirus HERV-K(II) DNA, comple 64% -5004 | HML |



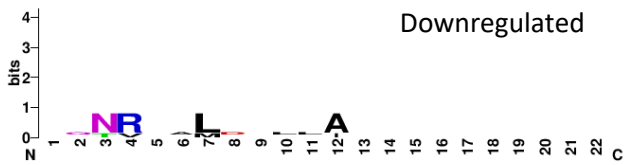
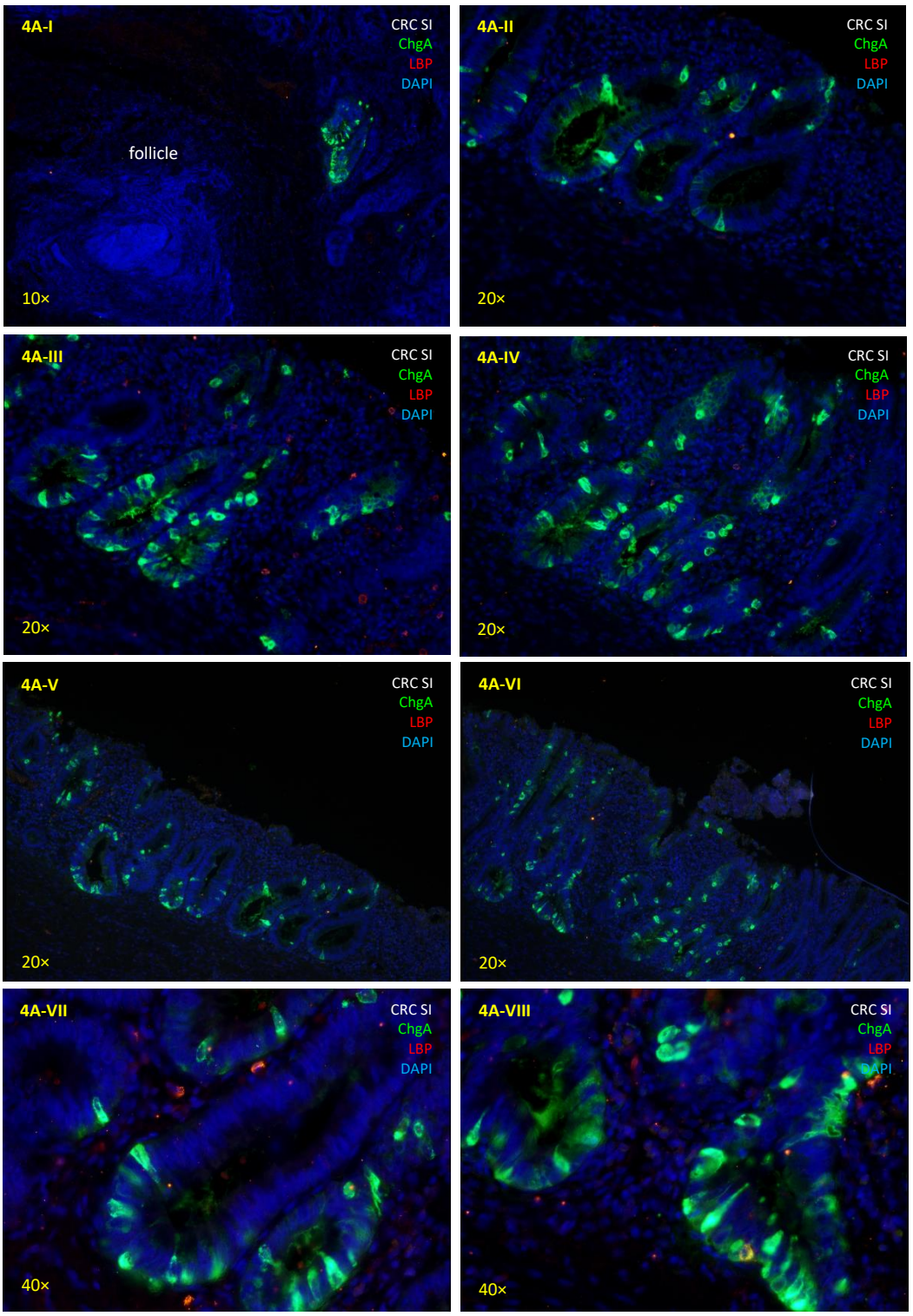
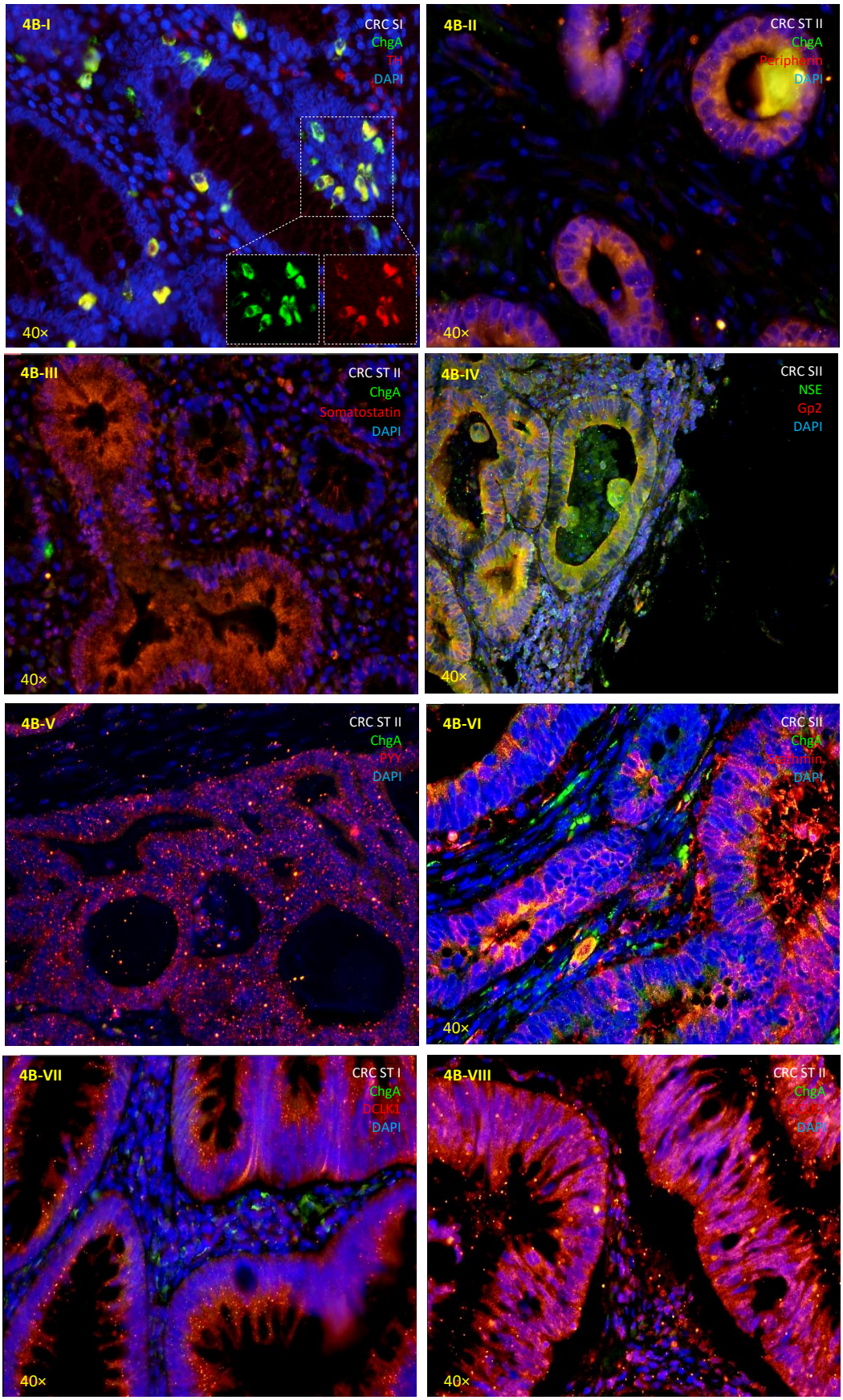



**Figure 3**: **Transcriptomic profiling of colorectal crypts revealing neuronal identity loss, immune remodeling, and HERVs regulation. Panel A**: **Loss of neuronal features and immune modulation in tumor crypts:** Transcriptomic analysis of colorectal crypts revealed major shifts in neuronal and immune-related signatures. Gene Set Enrichment Analysis (GSEA) showed strong negative enrichment of multiple neuronal pathways in tumor crypts (TC) vs. regular crypts (RC) (**Figure 3A**-**I**). Differential gene expression (DEG) analysis confirmed this finding. Intersecting DEGs with a custom neuronal gene set -compiled by merging multiple curated neuronal signatures-identified several significantly downregulated genes involved in neuronal differentiation, neurotransmission, synaptic architecture, and axonal function, including *CHGA, PHOX2B, SLC6A2, GABRG2, SNAP25, NRXN1, RELN, GAP43, NEFM,* and *NLGN1* (**Figure 3A**-**II**). These genes were also repressed in colorectal cancer (CRC) cell lines relative to crypt-derived samples (**Figure 3A**-**III**). In addition, STRING network analysis (see SI: Neuronal genes in crypts) showed a coordinated downregulation of interconnected neuronal gene modules, reinforcing the loss of neuronal features in tumor crypts. The transcription factor REST was found to be upregulated in the tumoral crypts **(Figure 3A**-**IV**). Its inhibition by Valproic acid (30 µg/ml) in the HCT8 colon carcinoma cell line induced the upregulation of several neuron-related genes, including ChAT, ChgA, BDNF, neurturin, tenascin C, as well as the cancer stem cell-related markers DCLK1 and LGR5 (**Figure 3A**-**V**). Ablation of REST in the HCT8 cell line using CRISPR/Cas9 confirmed that the expression of ChAT, ChgA, synaptophysin, TBR2, stathmin, and LGR5 is genuinely REST-dependent. In contrast, DCLK1 was found to be downregulated in HCT8$^{\Delta REST}$ cells (**Figure 3A**-**VI**). Among the neuronal proteins investigated, ChgA and TH was upregulated in the REST-mutated model and treated with 100 µg/mL VPA for 72 hours, consistent with the RNA results. Relevantly, ablation of REST induces a reduction of the pivotal transcription factor LGR5 at the protein level (**Figure 3A**-**VII**). In the colorectal carcinoma cell line HCT8, while wild type cells exhibit robust organoid generation (**Figure 3A**-**VIIIa-c**), REST-deficient edited cells show a marked impairment in this process (**Figure 3A**-**VIIId-f**). Continuous exposure of primary colorectal carcinoma cultures to valproic acid (VPA) at a subtoxic concentration of 100 µg/mL for over 14 days induces growth arrest and triggers pronounced morphological alterations. The treated cells exhibit a hypertrophic/senescent phenotype, as evidenced by their flattened morphology (**Figure 3A**-**IXa-d**). GSEA also revealed enrichment of germinal center B-cell-related pathways in TC. DEGs were intersected with a germinal center B-cell gene set to explore this, identifying a distinct set of upregulated and downregulated genes (**Figure 3A**-**VII, SI**: Immune significant genes TC). Immune cell deconvolution using CIBERSORTx and the LM22 signature matrix revealed compositional differences between RC and TC, including altered levels of B-cell subsets, plasma cells, $CD8^+$ T cells, and $CD4^+$ memory T cells (**Figure 3A**-**VIII**), indicating immune landscape remodeling in CRC. **Panel B**: **Regulation of human endogenous retroviruses** (**HERVs**) **in tumor crypts:** Principal component analysis (PCA) based on HERV expression distinguished TC, RC, and CRC cell lines into separate clusters (**Figure 3B**-**I**). DEG analysis between TC and RC identified multiple upregulated and downregulated HERV loci (**Figure 3B**-**II**). The most differentially expressed HERVs are visualized as bar graphs (**Figure 3B**-**III**), reflecting changes in transcripts per million (TPM). **Table 1** summarizes all significantly modulated HERVs reported by Vargiu et al.[40] to possess an ISD, including annotations of retroviral genes, ISD sequence, best-matching reference elements, and super group classifications. Additionally, a sequence logo was generated (https://weblogo.berkeley.edu/logo.cgi) for the ISD sequences of the upregulated (left) and the downregulated (right) HERVs, and placed at the bottom of the table.

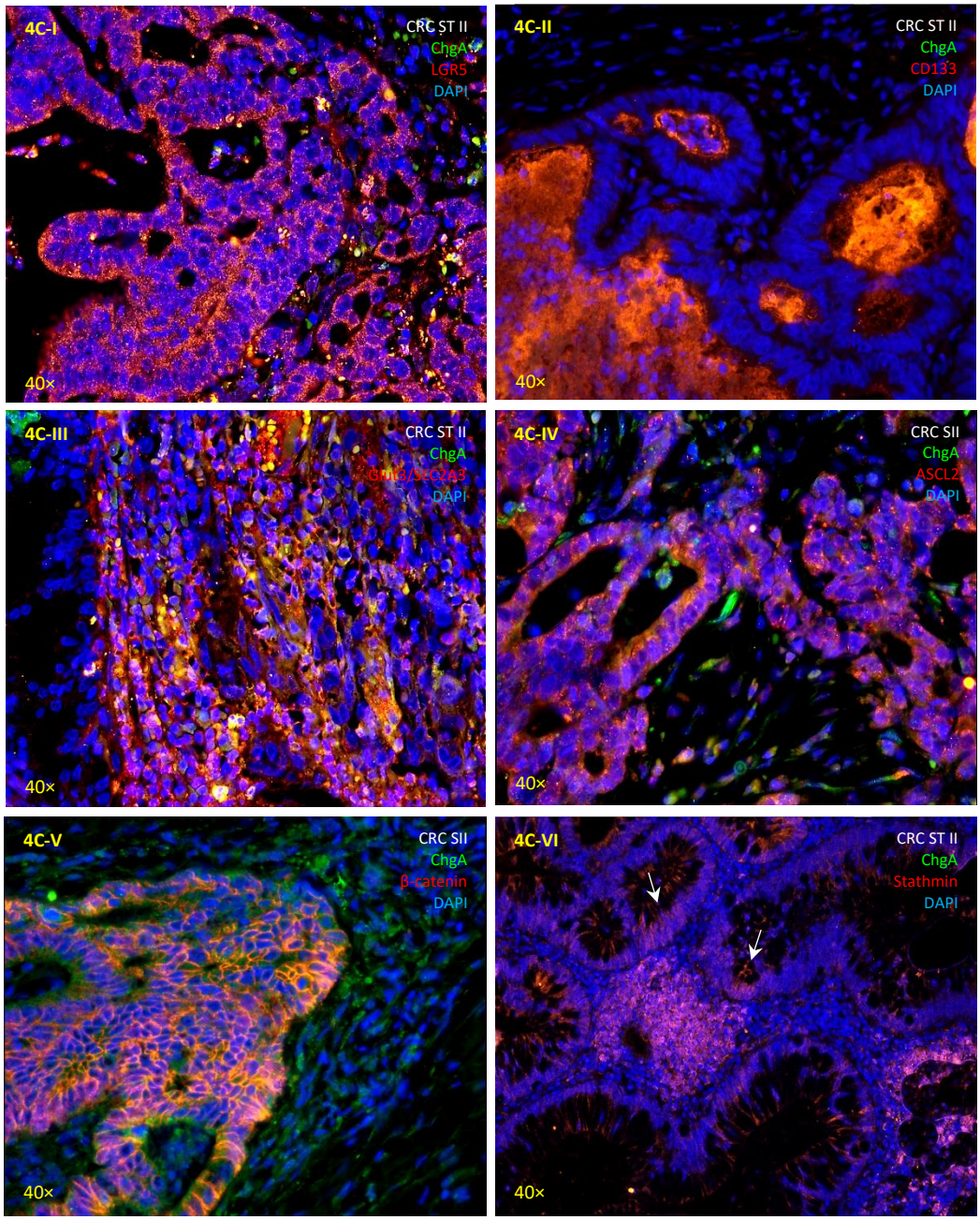
**Panel A**: Chromogranin A expression in early CRC development



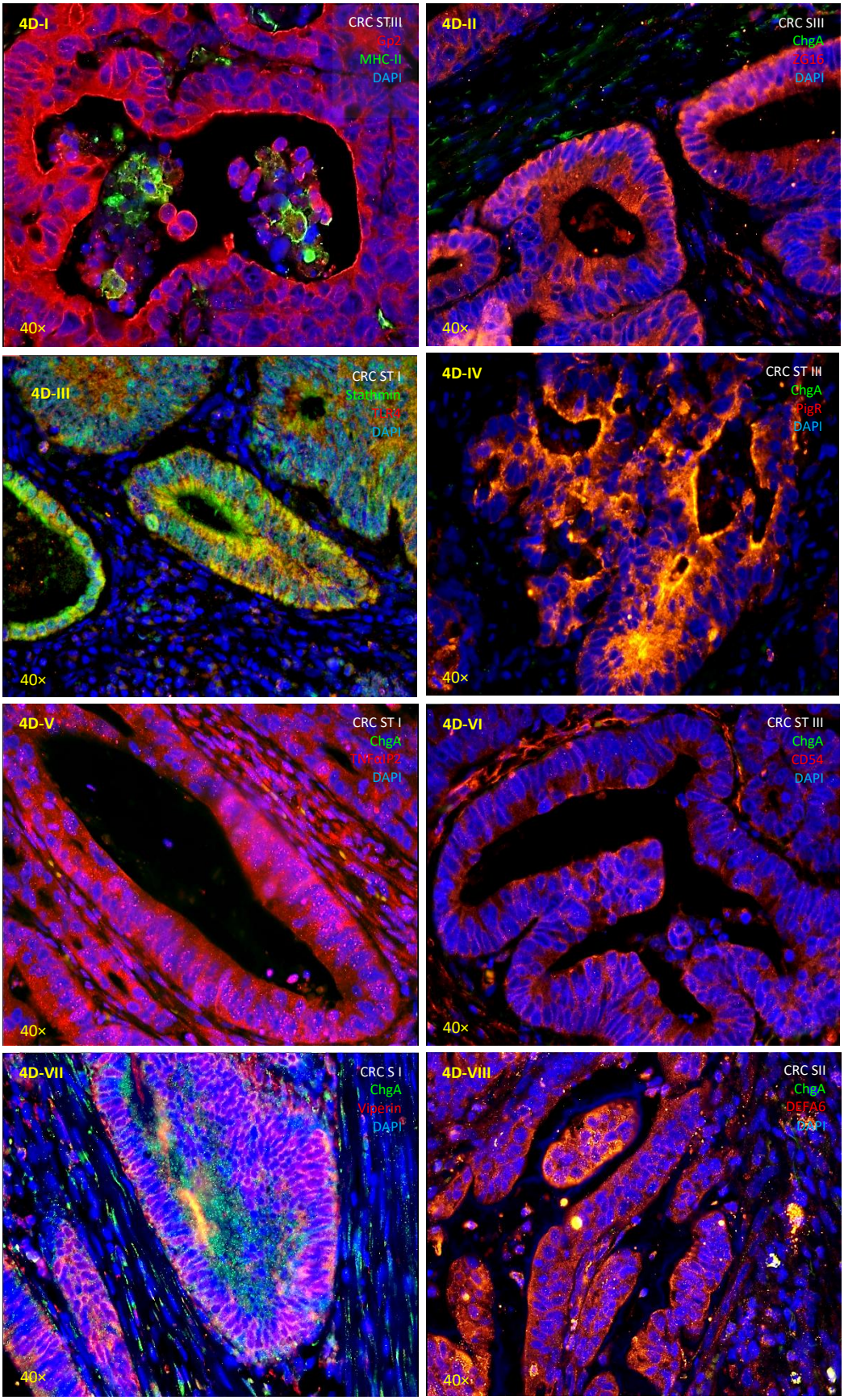
**Panel B**: Residual expression of neuronal-related markers in CRCs

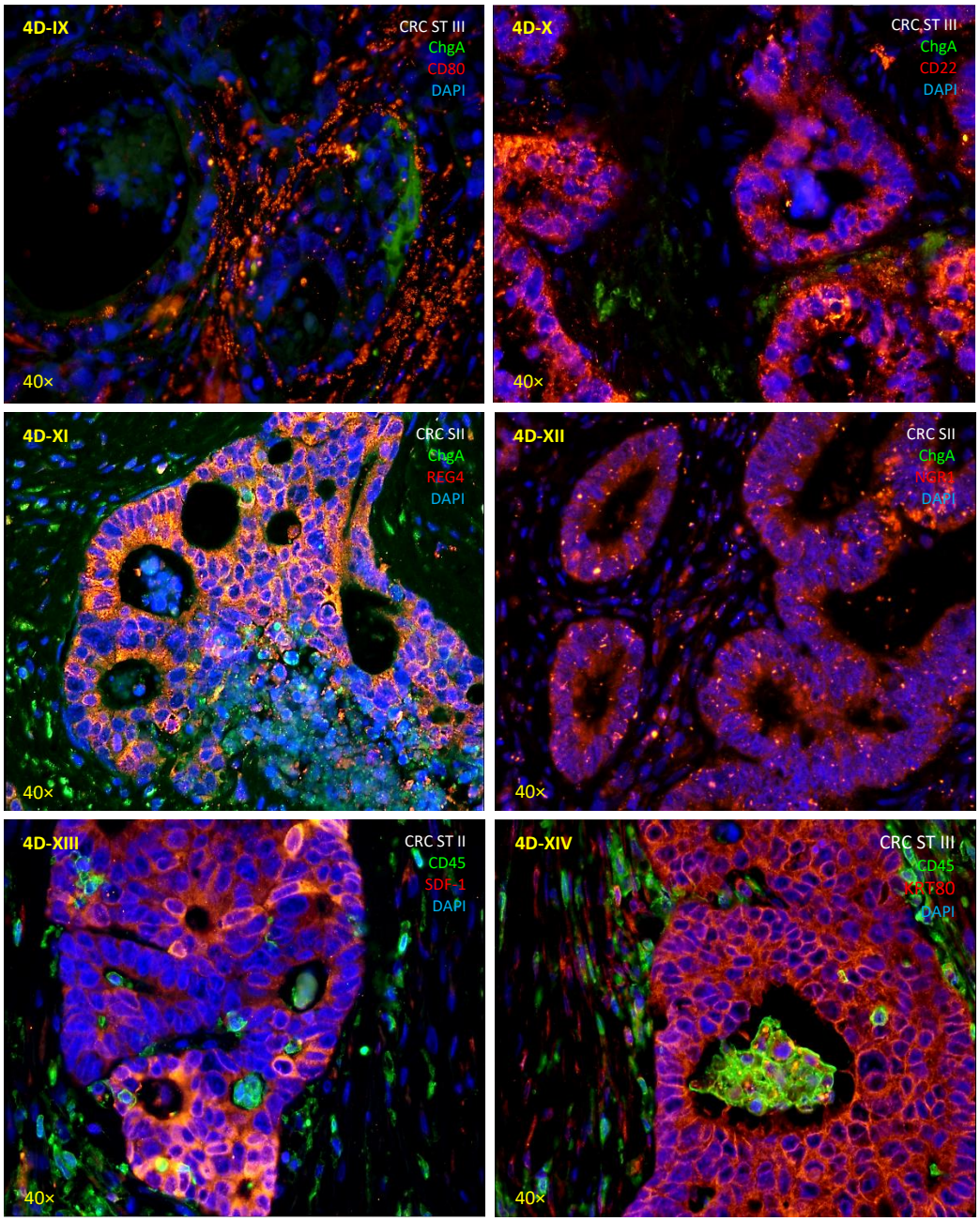


**Panel C**: Residual expression of stem cell markers in CRCs

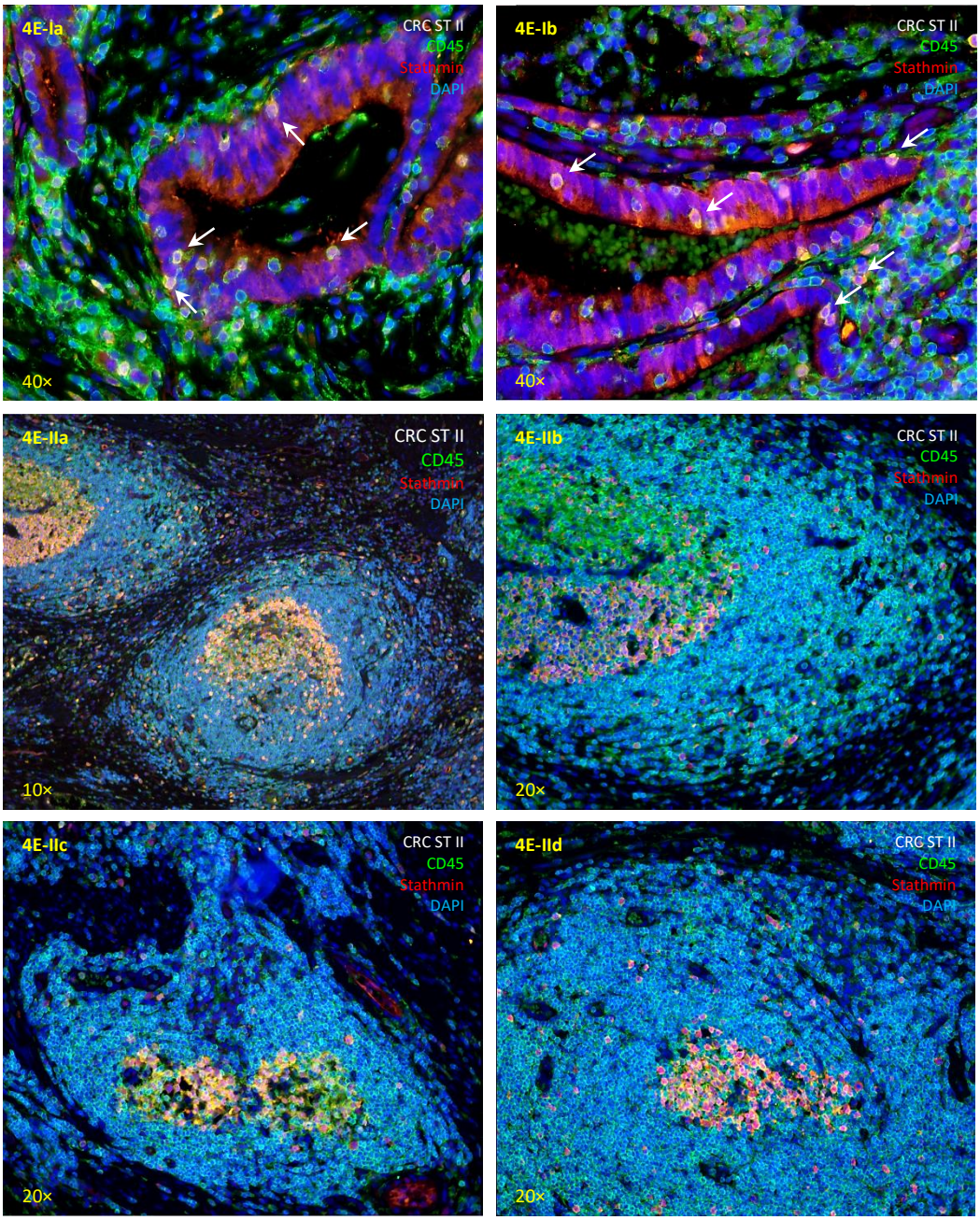


**Panel D**: Residual expression of immune-related markers in CRCs

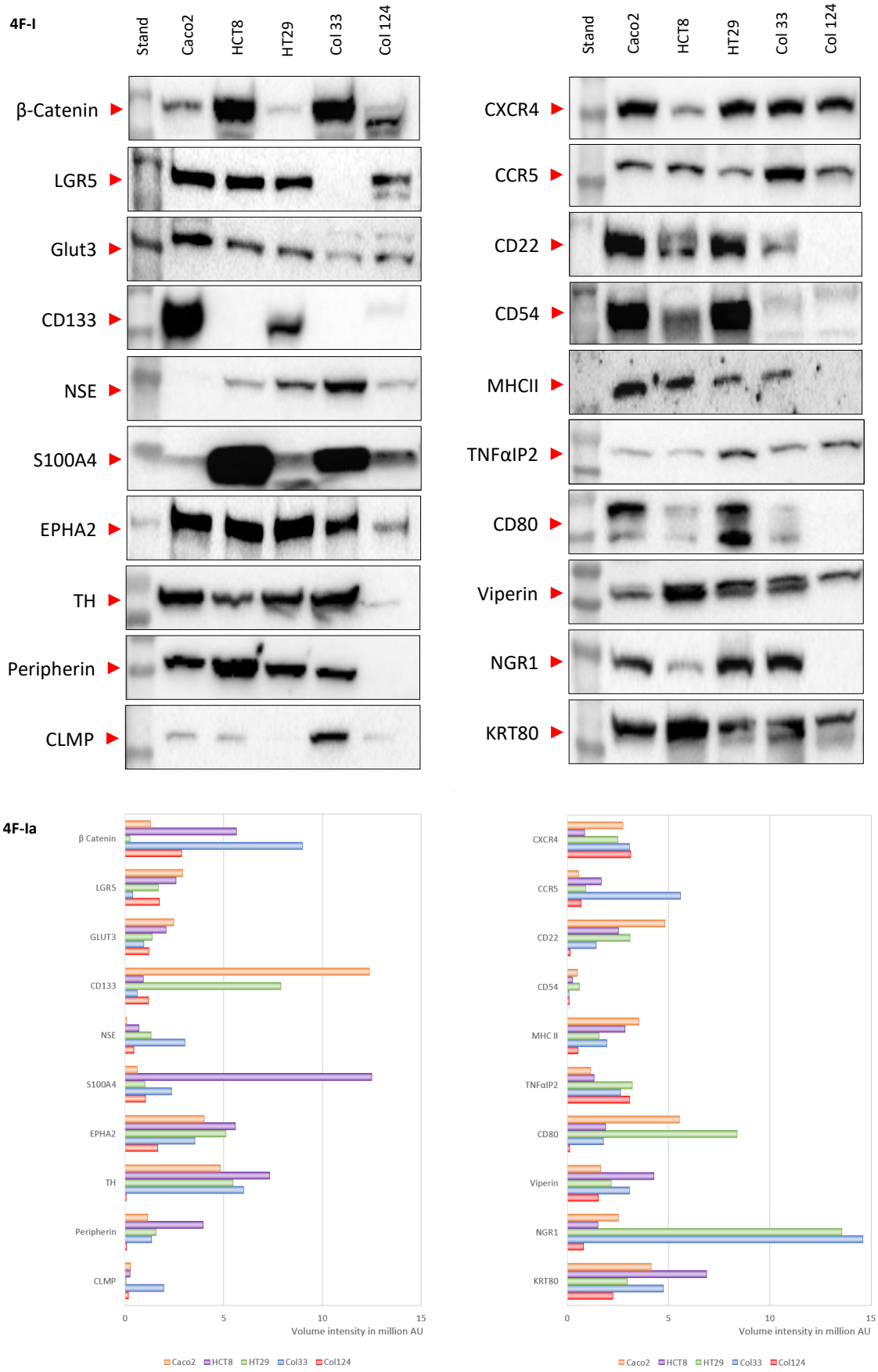



4D-IX
CRC ST III
ChgA
CD80
DAPI
40×
4D-X
CRC ST III
ChgA
CD22
DAPI
40×
4D-XI
CRC SII
ChgA
REG4
DAPI
40×
4D-XII
CRC SII
ChgA
NGB1
DAPI
40×
4D-XIII
CRC ST II
CD45
SDF-1
DAPI
40×
4D-XIV
CRC ST III
CD45
KRT80
DAPI
40×

**Panel E**: Follicle resident stathmin-positive cells migrating into the epithelial layer of CRCs.



**Panel F**: The neuroimmune signature of commercial and primary colorectal carcinoma cell lines



**Panel G**: Expression of lineage and state-specific markers in CRC cell lines

**4G-I**

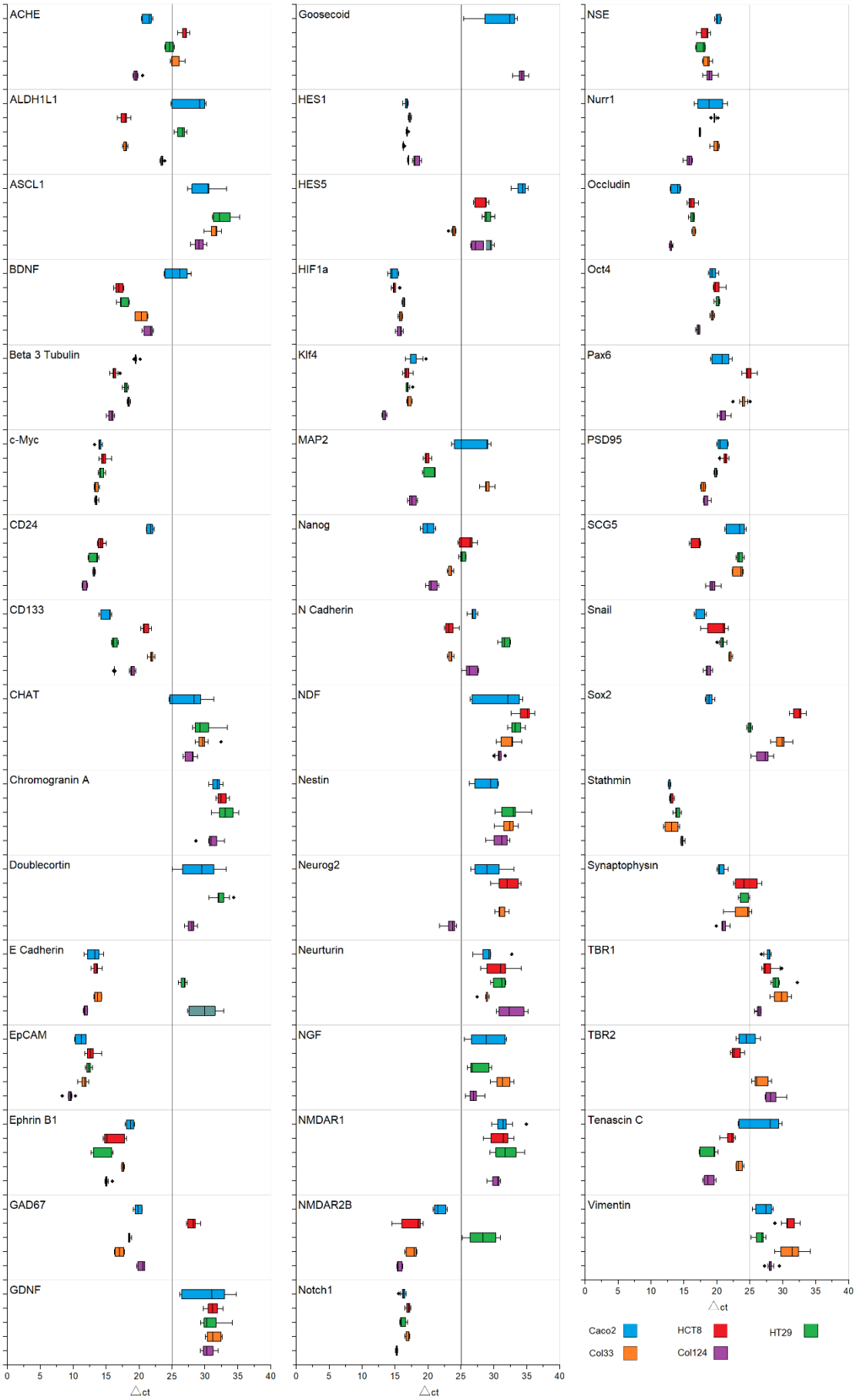
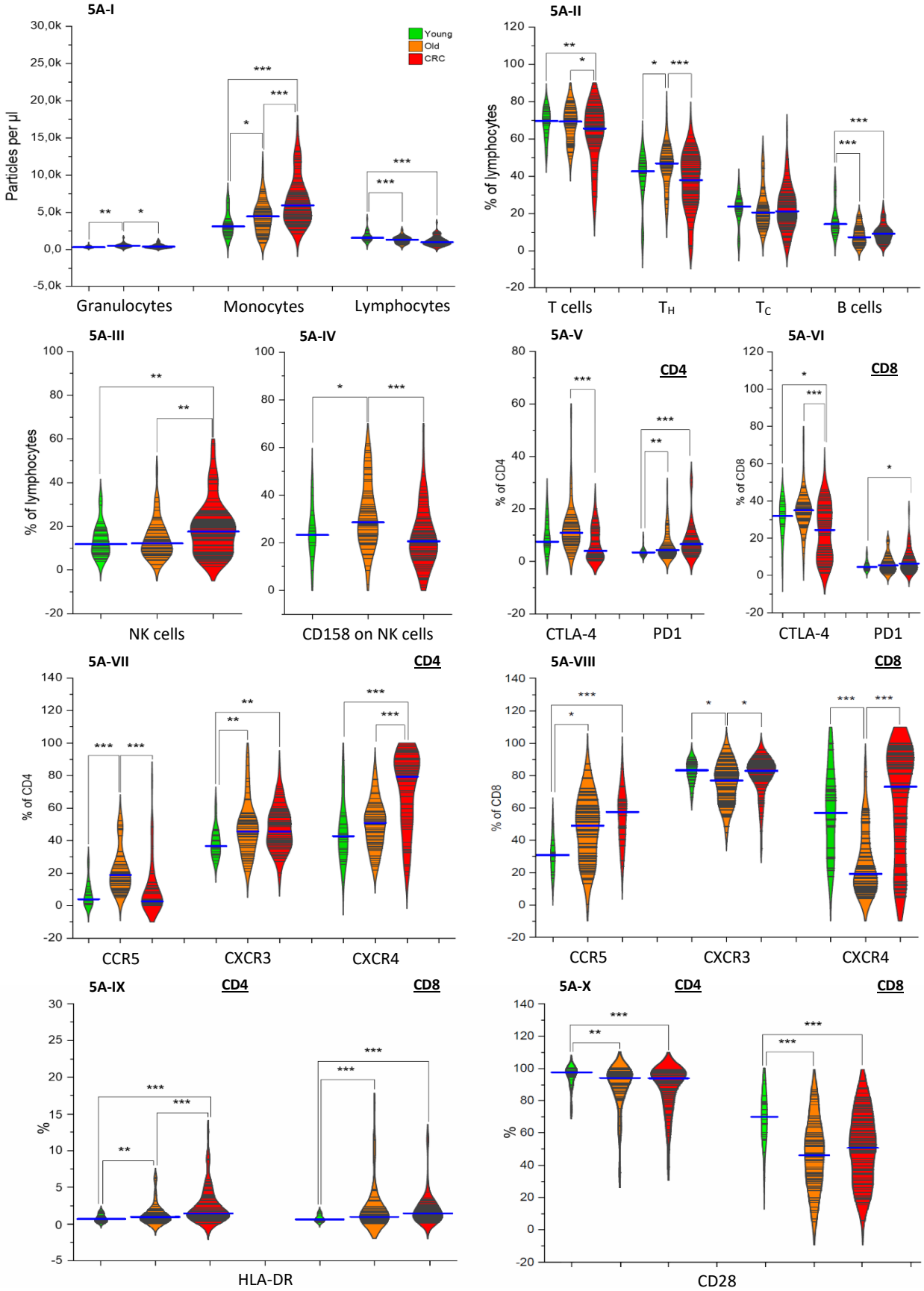
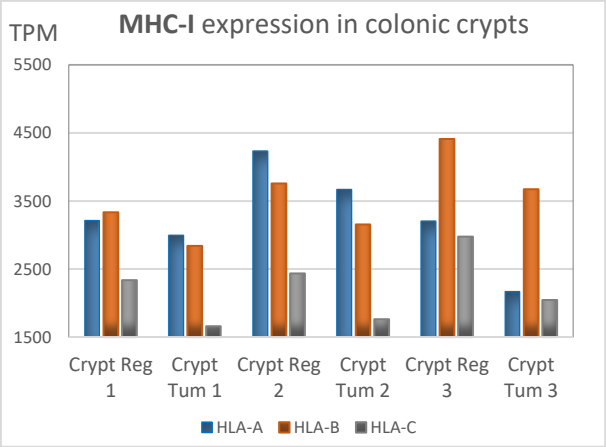


**Figure 4**: Colorectal carcinoma tumors exhibit hallmarks congruent with NICA cells. **Panel A**: Expression of ChgA in CRC early development. In the early stages of CRC, an increased number of ChgA-expressing cells are detected in the crypts, especially near the subjacent lymphatic tissue. ChgA expression is observed either focally or extensively in the crypts, from the basal portions upwards, as shown in **Figures 4A-I-VIII**. **Panel B**: Residual neuronal-related markers in CRCs. At various stages of CRC development, neuronal markers such as TH (**Figure 4B-I**), peripherin (**Figure 4B-II**), somatostatin (**Figure 4B-III**), NSE (**Figure 4B-IV**), PYY (**Figure 4B-V**), and the neuron-related structural protein stathmin (**Figure 4B-VI**) were observed. Colorectal carcinomas generally express DCLK1 (**Figures 4B-VII** & **VIII**), a marker originally identified as neuronal and also expressed in tuft cells of the gastrointestinal tract. **Panel C**: Residual expression of stem cell markers in CRCs. Colorectal carcinomas retain wide expression of stem cell markers including LGR5 (**Figure 4C-I**), CD133 (**Figure 4C-II**), SLC2A3 (**Figure 4C-III**), ASCL2 (**Figure 4C-IV**), β-catenin (**Figure 4C-V**) and Stathmin, which is not localized to the tumor stroma (**Figure 4C-VI**) but found as having infiltrated cells from the *Lamina propria* (**Figure 4C-VI**, arrows). **Panel D**: Residual expression of immune-related markers in CRCs. Several immune-related markers are found in CRCs, including MHCII at very low levels (**Figure 4D-I**), the innate immune-related proteins Gp2 and ZG16 (**Figures 4D-I** and **4D-II** respectively), REG4 (**4D-III**), PigR (**4D-IV**), TNFaIP2 (**4D-V**), CD54 (**4D-VI**), Viperin (**4D-VII**), DEAF6A (**4D-VIII**), CD80 (**4D-XI**), CD22 (**Figure 4D-X**), REG4 (**Figure 4D-XI**), NGR1 (**Figure 4D-XII**), SDF-1 (**Figure 4D-XIII**), and the structural NICA-related KRT80 (**Figure 4D-XIV**). **Panel E**: Follicular and parafollicular stathmin-positive cells migrating into the epithelial layer of CRCs. stathmin-positive cells are primarily found in the germinal center of gut-associated lymphatic follicles, from where they migrate to the colon epithelium and fuse (**Figures 4E-Ia-b** & **4E-IIa-d**, white arrows). **Panel F**: The neuroimmune signature of CRC cell lines and primary colorectal carcinoma cells. Colorectal carcinoma cells exhibit a dual neuronal and immune signature, along with a prominent expression of relevant stemness markers, which is not observed in all standard CRC cell lines or primary cell cultures. Most cell lines like Caco2, HCT8, and HT29 express LGR5 and Glut3, as well as the structural protein β-catenin (**4F-I**). The neuronal signature is recognizable in almost all cell types by the expression of NSE, EPHA2, TH, and Peripherin. The immune-related signature is confirmed by the expression of CXCR5 and CCR5 in all CRC cells. CD22, CD80, NGR1, and MHCII are expressed in all cells except the primary cell line Col124. TNFaIP2 and viperin are expressed in all CRC cells, as the structural protein KRT80 is observed in NICAS cells. **Figure 4F-Ib** reflects the densitometric measurements related to β-actin or GAPDH. **Panel G**: **Expression of lineage and state-specific markers in CRC cell lines**. We used qPCR to evaluate the expression of neuronal (*MAP2, NSE, Synaptophysin, ChAT, GDNF*), neurodevelopmental (*ASCL1, SOX2, NESTIN, NEUROG2*), stemness (*OCT4, NANOG, CD133*), epithelial (*EpCAM, E-cadherin*), and mesenchymal (*vimentin, Snail, tenascin C*) markers across five CRC cell lines (Caco2, HCT8, HT29, Col33, Col124) (**Figure 4G-I**). Expression patterns are heterogeneous between cell lines, with some exhibiting high expression of neuronal and neuroendocrine genes and others displaying classical epithelial and less mesenchymal profiles. IHC pictures are representative of n=10 patients. WB and qPCR results are representative of at least 4 biological replicates.

**Panel A**: Age- and cancer-associated changes in peripheral immune cell distribution



**Panel B**: Downregulation of MHC-I is concomitant with an MHC-II upregulation in CRC isolated crypts



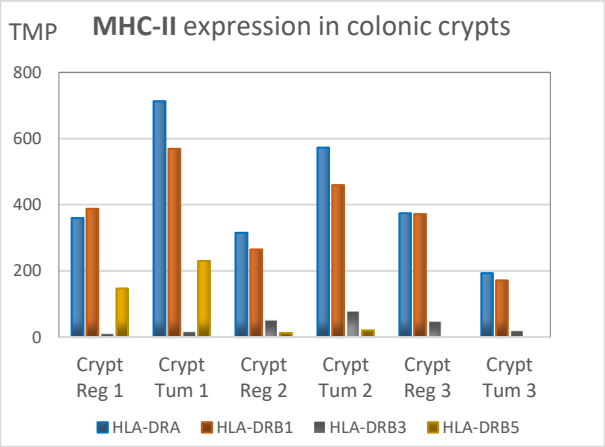
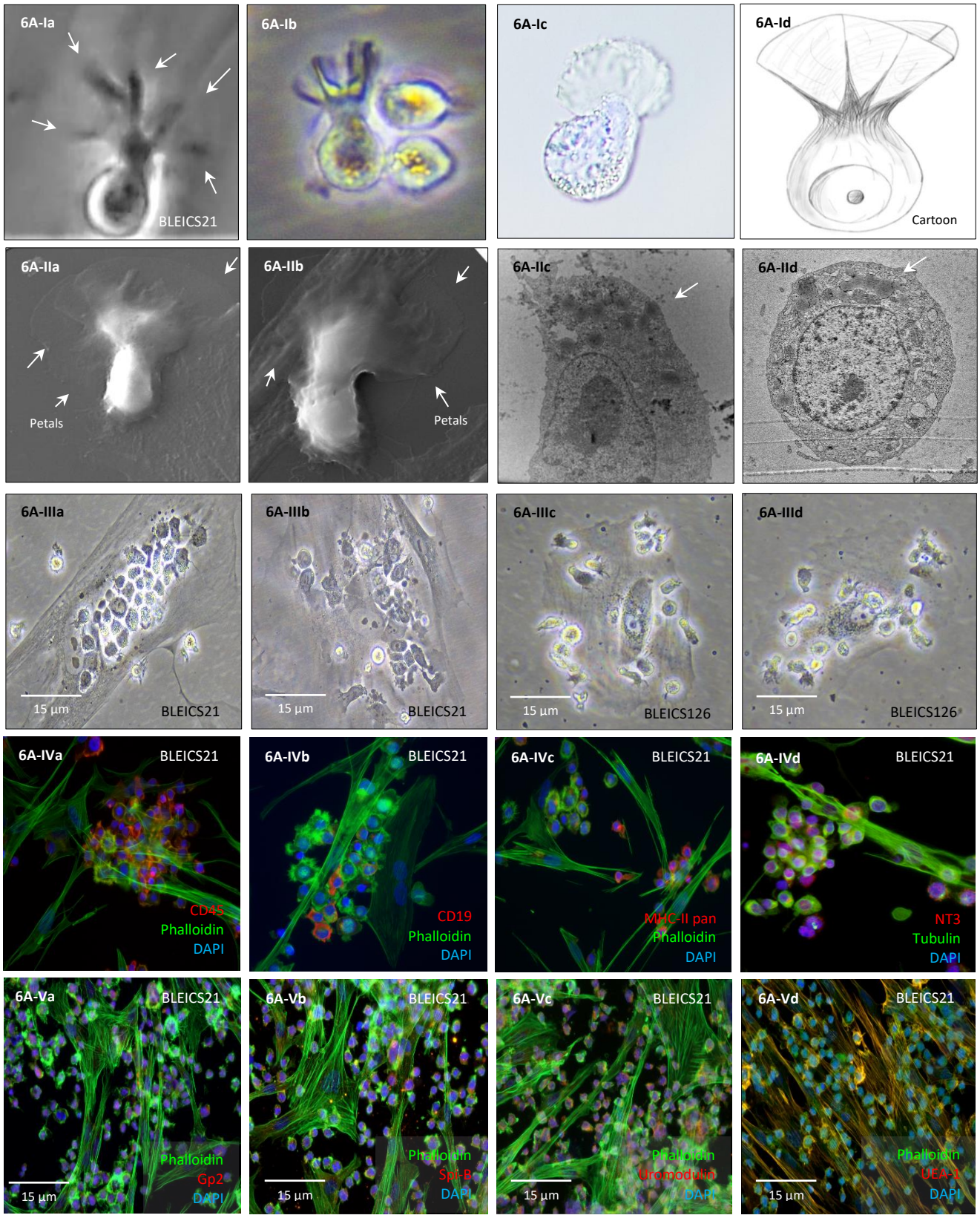
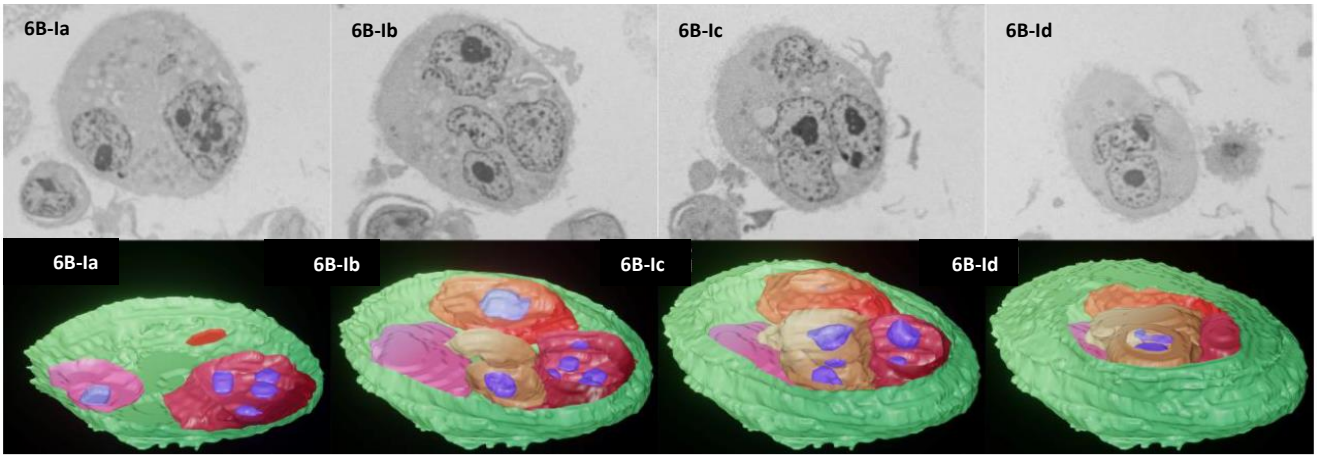



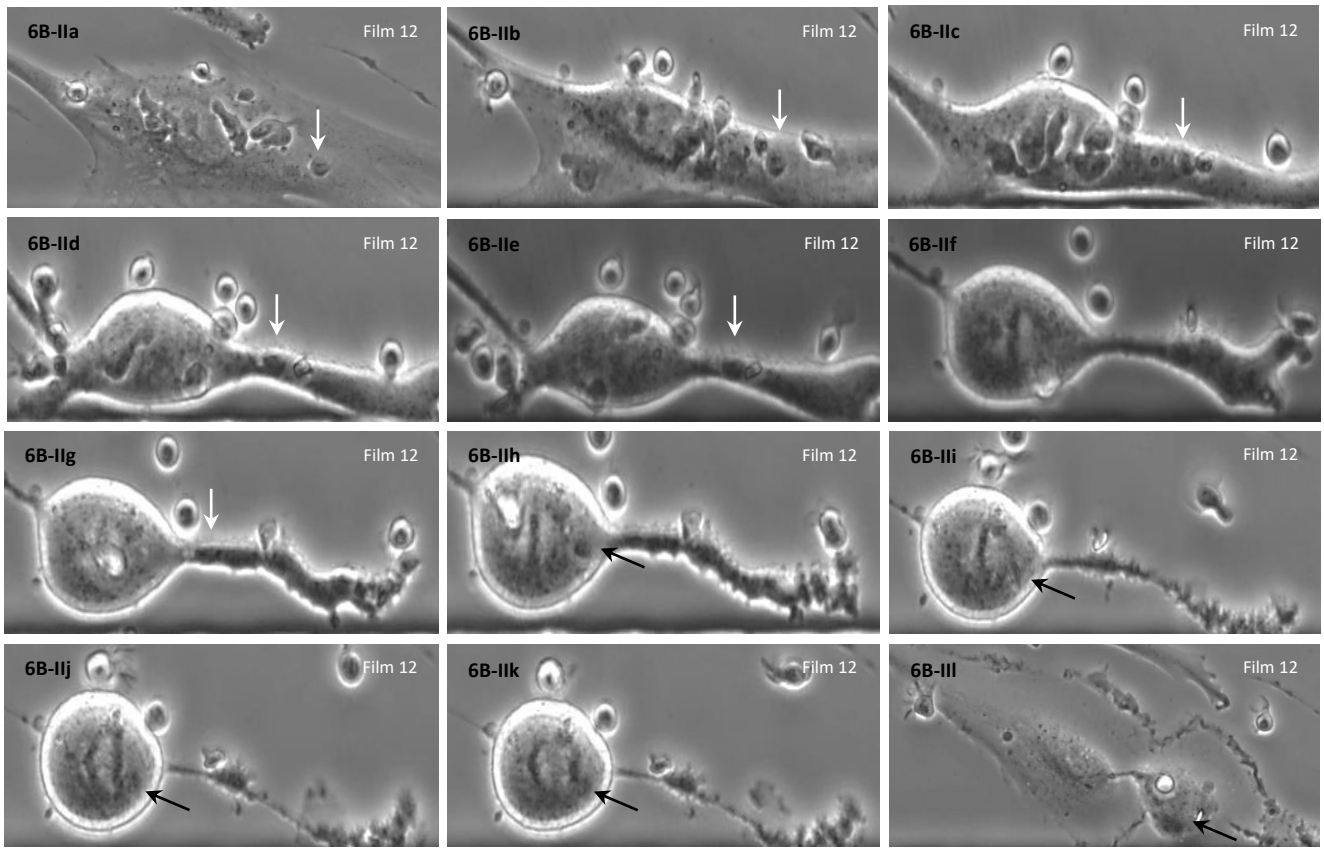

**Figure 5**. **Immune landscape alterations in colorectal cancer patients**. The immune status of patients with CRC shows peripheral impairment in cell number and distribution. **Panel A**: Peripheral immune cell distribution in patients with colorectal carcinoma. FACS analysis revealed a reduction in the populations of peripheral immune cells and in the expression of cytokines and immune checkpoint molecules. 3 distinct populations grouped by age and condition were analyzed: a young group (<25 years old), an older group without previous cancer history (median age 83), and a CRC group (median age 72). The older group had significant monocytosis compared to the younger group. This monocytosis was highly significant in CRC patients compared to the healthy same-age group. The older group had fewer lymphocytes compared to the younger group (**Figure 5A-I**). As to lymphocyte distribution among the three groups, the CRC group had significantly fewer T helper cells than the healthy same-age group. There was no detectable change in the T cytotoxic population. Older individuals tend to have significantly fewer B cells, which was also observed in the CRC group (**Figure 5A-II**). As expected, NK cells were significantly elevated in the CRC group compared to the healthy same-age group (**Figure 5A-III**). However, expression of the activation molecule CD158 was lower expressed in the NK cells of CRC patients (**Figure 5A-IV**). Expression of the immune checkpoint CTLA-4 was lower $CD4^+$ and $CD8^+$ cells in the CRC group, whereas PD1 did differ significantly changes between groups (6A-V and 6A-VI, respectively). The chemokine receptors on $CD4^+$ cells show a significant upregulation of CCR5 and CXCR4 in CRC patients (**Figure 5A-VII**). In $CD8^+$ populations, only CXCR4 was significantly upregulated in CRC patients (**Figure 5A-VIII**). HLA-DR (**Figure 5A-IX**) was found significantly upregulated in $CD4^+$ cells in CRC patients compared with the same-aged group and generally upregulated in $CD4^+$ and $CD8^+$ populations compared with the young group. CD28 was downregulated in $CD4^+$ and $CD8^+$ populations only in CRC patients compared with the young group (**Figure 5A-X**). **Panel B**: Shift from MHC class I to class II gene expression in the tumor crypts. Transcriptomic analysis of tumor crypts from CRC patients reveals a decrease in the basal expression of several MHC class I-related genes, alongside an increase in MHC class II components, particularly HLA-DRA and HLA-DRB1. Flow cytometry analysis were performed with freshly collected EDTA blood samples. The young group was comprised of 30 participants, the older group included 60 participants, and the CRC group 102 patients.

**Panel A:** Morphological and phenotypic characterization of BLEICS cells associated with colorectal carcinomas



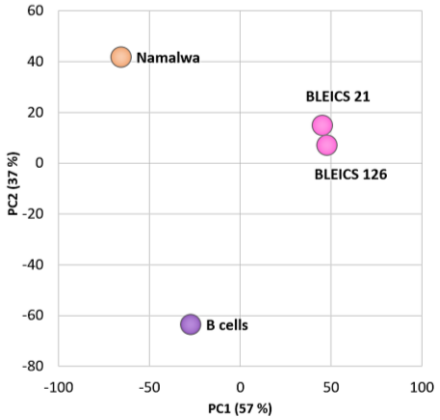
**Panel B:** Intracellular localization and dynamics of fusion between BLEICS cells and supporting fibroblasts.



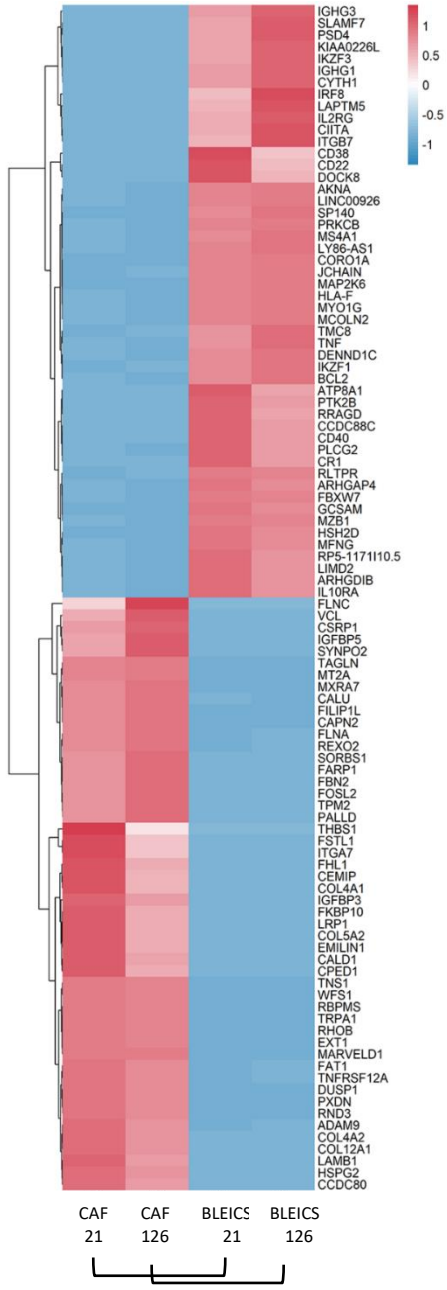



**Panel C**: BLEICS cells revealed a transcriptional background with mixed B-DC cell phenotype experienced with multiple pathogen infections

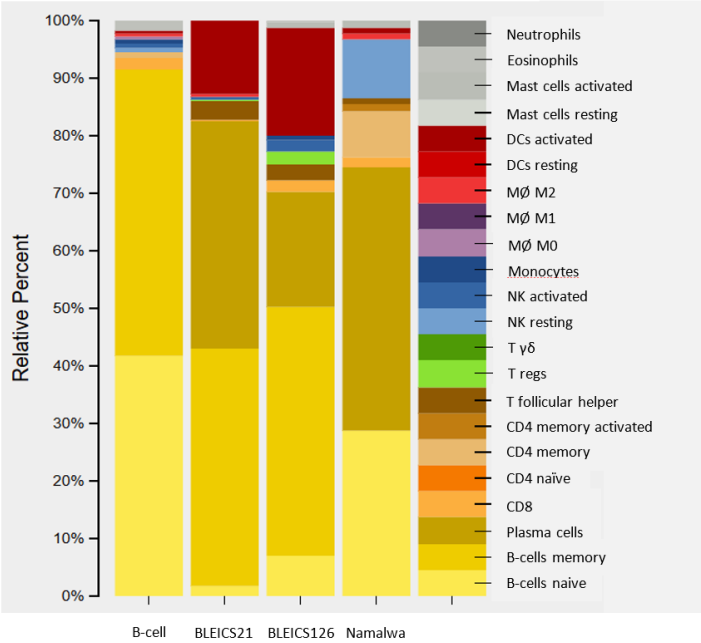
**6C-Ia**



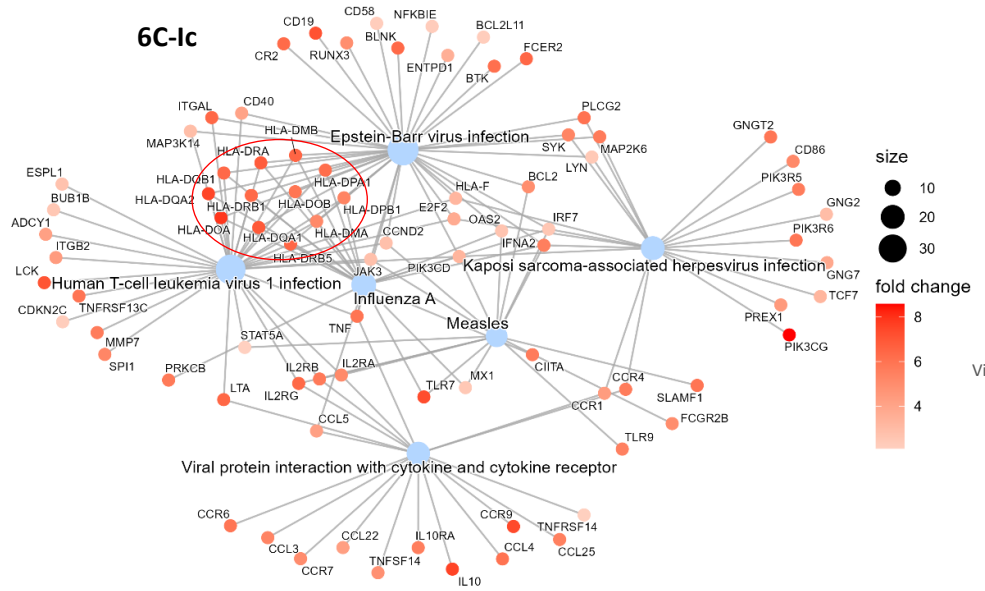

**6C-Id**



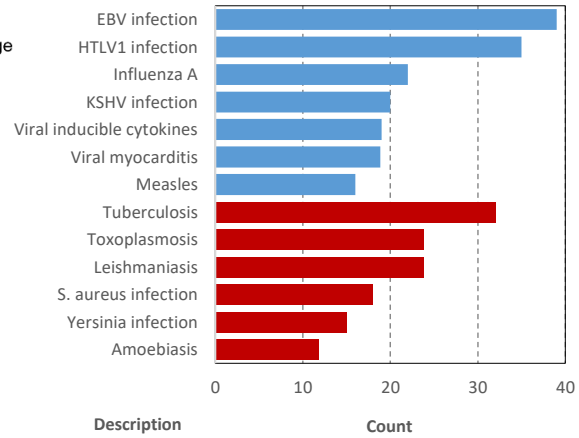

**6C-Ib**




**6C-Ic**

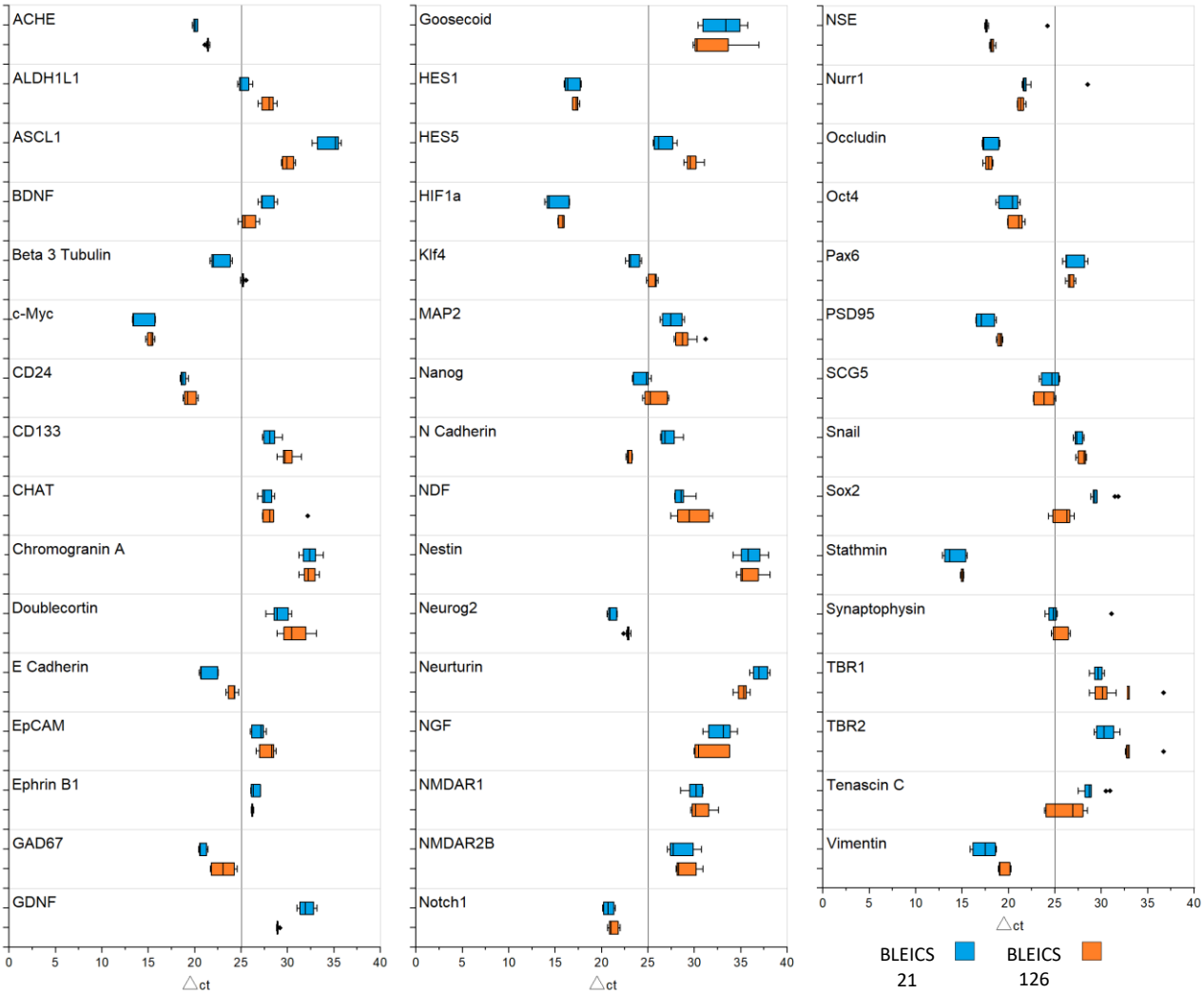



**6C-Ie**

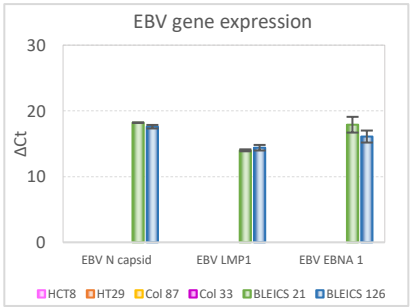


**Panel D**: BLEICS cells show a peculiar expression pattern of neuronal genes and all relevant structural EBV genes

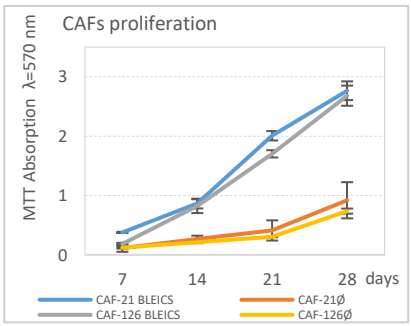
**6D-I**



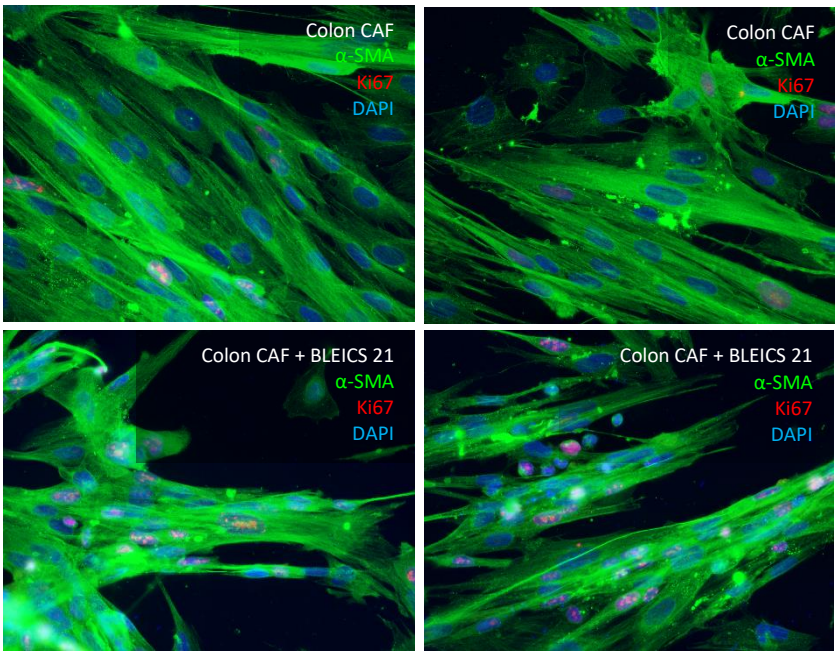

**6D-II**



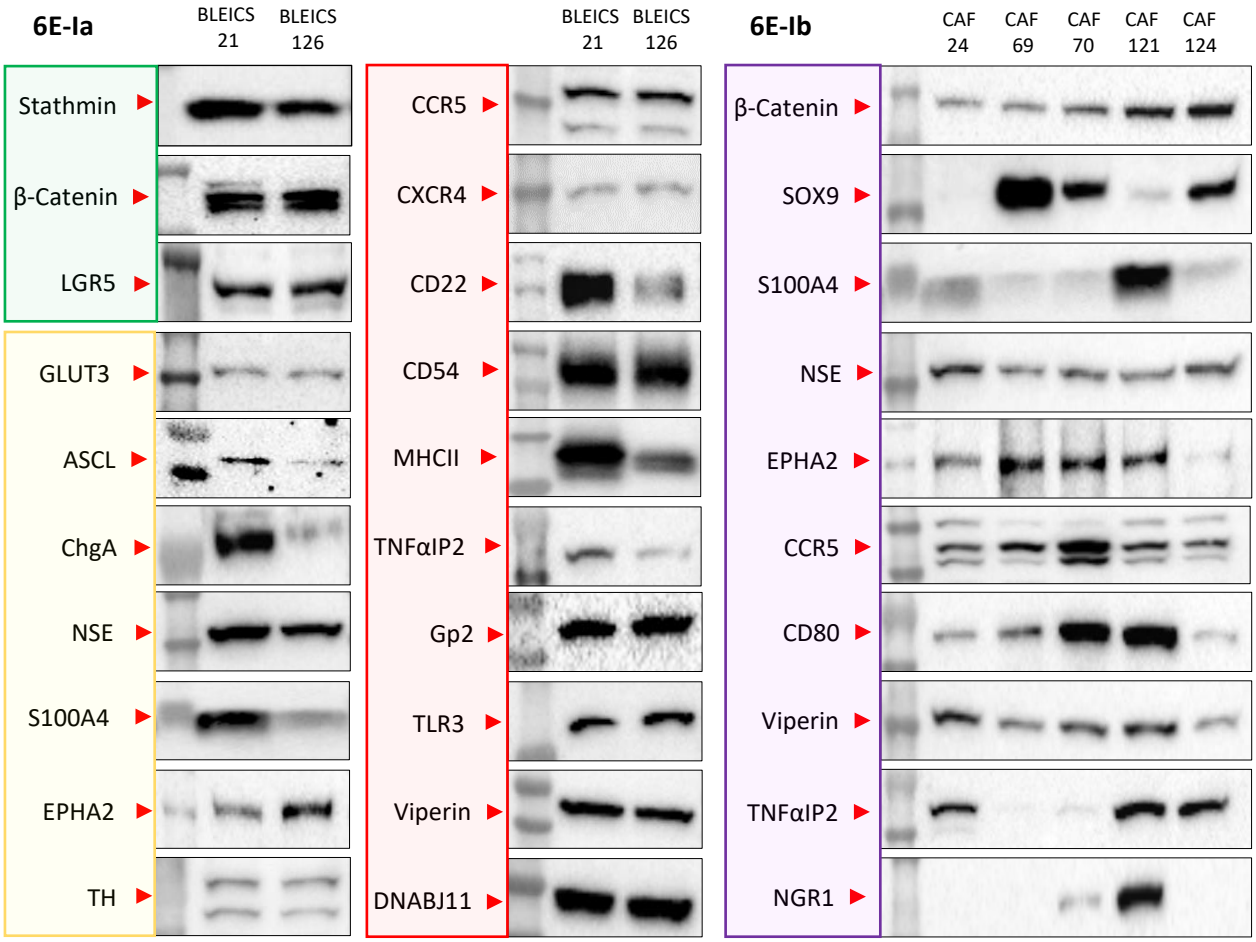

**6D-III**

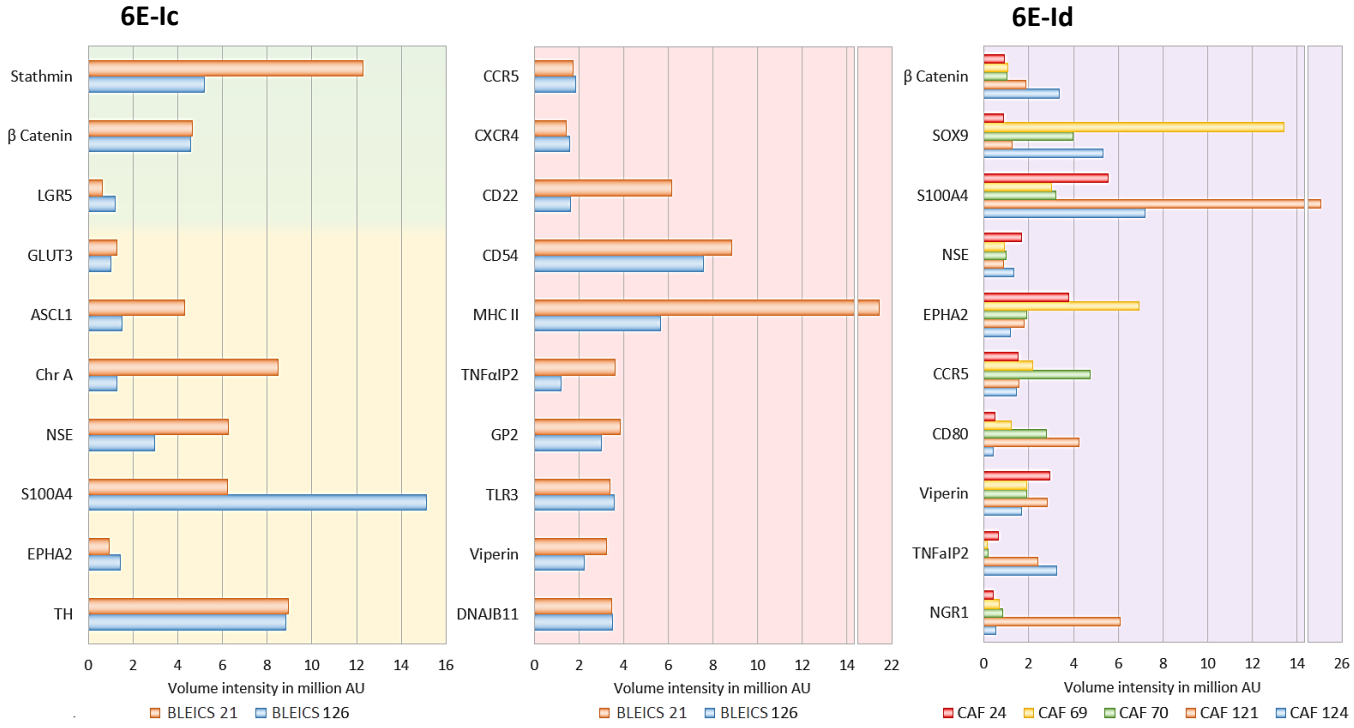



**6D-IV**



**Panel E**: Protein expression of neuronal, stem cell and immune-related markers in BLEICS cells and colon CAFs



**Panel F**: Regulation of human endogenous retroviruses (HERVs) in BLEICS cells vs. regular B-lymphocytes

**6F-I**

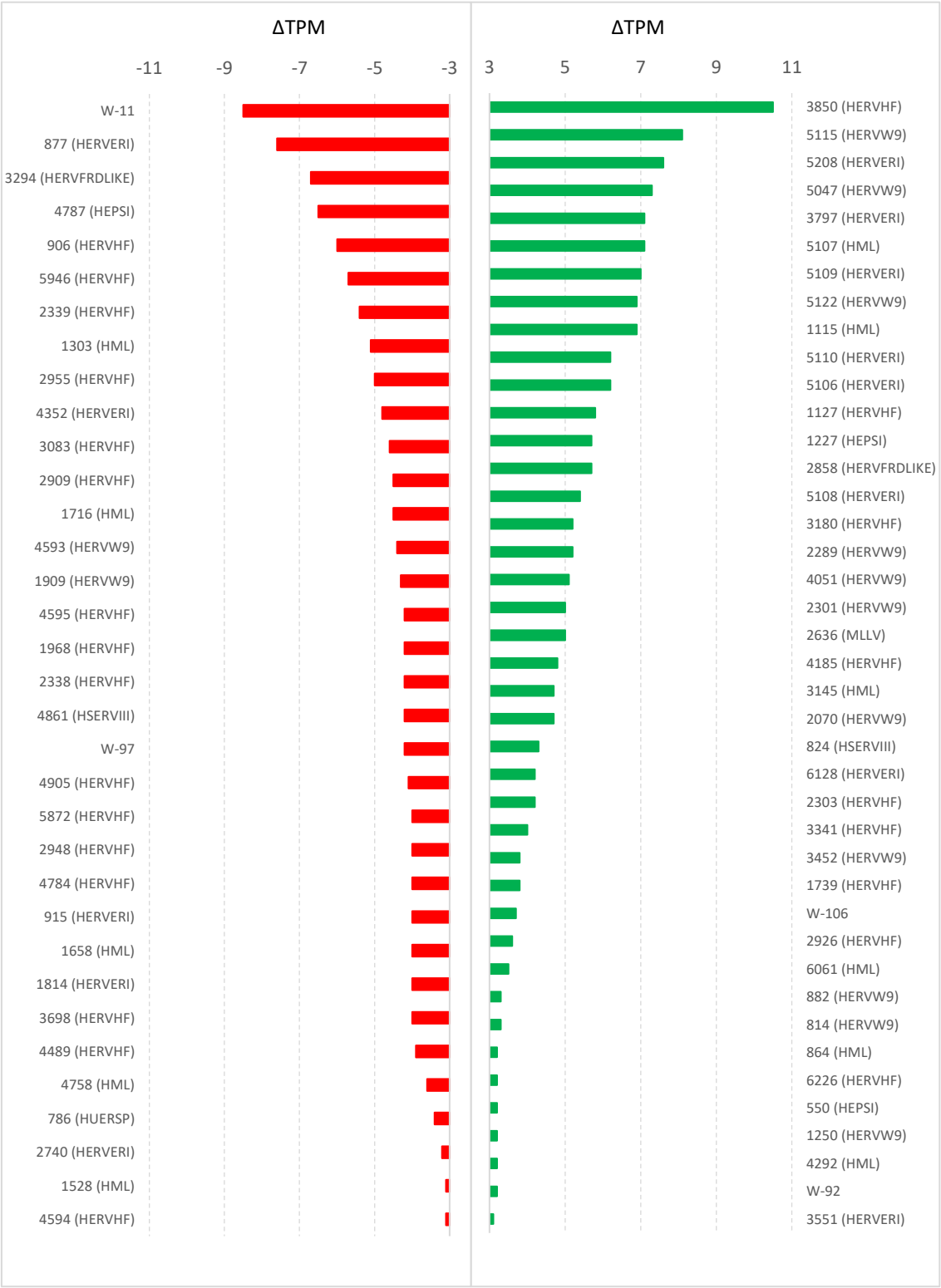
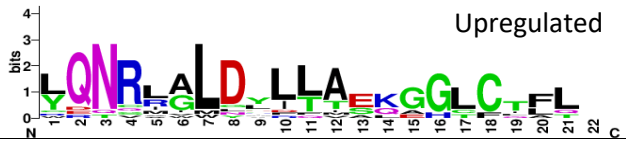


**Table 2**: Differentially expressed HERVs with ISD in BLEICS cells

| HERV | -LOG$_{10}$ (P VALUE) | ΔTPM | SUBGENES | ISD | BESTREFRV | SUPER GROUP |
|---|---|---|---|---|---|---|
| 1227 | 1.5 | 5.7 | 5LTR MA CA NC Prot RT IN TM PPT 3LTR | **LdNqiALDhLLA** | | HEPSI |
| 2636 | 2.3 | 5.0 | 5LTR PBS MA CA NC Prot RT IN TM 3LTR | **LQNRRGLDLLFiSQGGLCtALE** | Cat ECE1 endogenous retroviral element DNA, partial 57% -3795 | MLLV |
| 2858 | 1.7 | 5.7 | 5LTR PBS MA CA NC Prot RT IN TM 3LTR | **LQNRmALDiiTtAQGGTCtLtK** | | HERVFRDLIKE |
| 5106 | 1.8 | 6.2 | 5LTR PBS NC Prot RT IN SU TM PPT 3LTR | **YQNRLALDYLLA** | Homo sapiens retinoic acid-inducible endogenous re 85% -6483 | HERVERI |
| 5108 | 2.1 | 5.4 | 5LTR PBS CA DU IN SU TM PPT 3LTR | **YQNRLALDYLLA** | Human endogenous retroviral DNA (4-1), complete re 68% 1017 | HERVERI |
| 5115 | 2.2 | 8.1 | 5LTR PBS CA NC Prot RT IN SU TM PPT 3LTR | **LQNczGLYLLttEKaGLCtFLG** | Human endogenous retrovirus pHE.1 (ERV9). 90% -3615 | HERVW9 |
| 5122 | 1.5 | 6.9 | 5LTR PBS MA CA NC Prot IN SU TM PPT 3LTR | **LQNRRGLDcLTAEKGGLCtFLG** | Human endogenous retrovirus pHE.1 (ERV9). 74% -3430 | HERVW9 |
| 5208 | 1.4 | 7.6 | 5LTR PBS MA CA NC Prot RT IN SU TM PPT 3LTR | **YQNRLvLDzLLA** | Homo sapiens retinoic acid-inducible endogenous re 67% -6129 | HERVERI |
| 6128 | 1.7 | 4.2 | 5LTR PBS MA CA NC Prot RT IN SU TM 3LTR | **cQNRLALDYLLA** | Human endogenous retroviral DNA (4-1), complete re 80% 299 | HERVERI |
| 877 | ∞ | -7.6 | 5LTR PBS MA CA NC Prot RT IN SU TM PPT 3LTR | **YQNRLALyYLLA** | Human endogenous retroviral DNA (4-1), complete re 79% 261 | HERVERI |
| 906 | ∞ | -6.0 | 5LTR PBS MA CA NC Prot RT IN SU TM PPT 3LTR | **LQNRqsLDLLTAEKGGLCIFLN** | Homo sapiens RGH2 gene, retrovirus-like element. 83% -22 | HERVHF |
| 1303 | ∞ | -5.1 | 5LTR PBS MA CA NC DU Prot RT IN SU TM PPT 3LTR | **RQTVIWMGDRLMSLEylFQLQC** | Human endogenous retrovirus HERV-K(II) DNA, complete 84% 981 | HML |
| 1716 | ∞ | -4.5 | 5LTR PBS CA NC DU Prot RT IN TM PPT 3LTR | **LRziVMWLGDRV** | Human endogenous retrovirus HERV-K10. 62% 1031 | HML |
| 1814 | ∞ | -4.0 | 5LTR MA CA IN SU TM PPT 3LTR | **YQNRLALDYLLA** | Human endogenous retroviral DNA (4-1), complete re 64% 581 | HERVERI |
| 1968 | ∞ | -4.2 | 5LTR PBS NC Prot IN SU TM PPT 3LTR | **LQNRzGLDLLTAEKGGLCIFLN** | Human endogenous retrovirus, complete genome. 92% 792 | HERVHF |
| 3698 | ∞ | -4.0 | IN SU TM PPT 3LTR | **LQNhRGLDLLTAEKGGLCIFLN** | Homo sapiens RGH2 gene, retrovirus-like element. 77% 4627 | HERVHF |
| 4352 | 1.4 | -4.8 | 5LTR MA CA NC Prot IN SU TM PPT 3LTR | **YzNRLALDYHLA** | Homo sapiens retinoic acid-inducible endogenous re 96% -4626 | HERVERI |
| 4593 | ∞ | -4.4 | 5LTR PBS MA CA NC Prot RT SU TM 3LTR | **LQNRqGLDLLTAEKGGfCtFLG** | Human endogenous retrovirus pHE.1 (ERV9). 89% -2775 | HERVW9 |



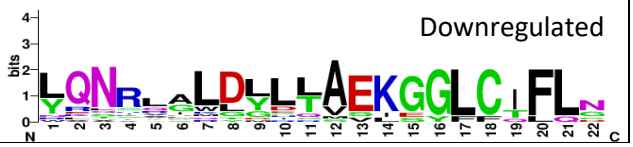
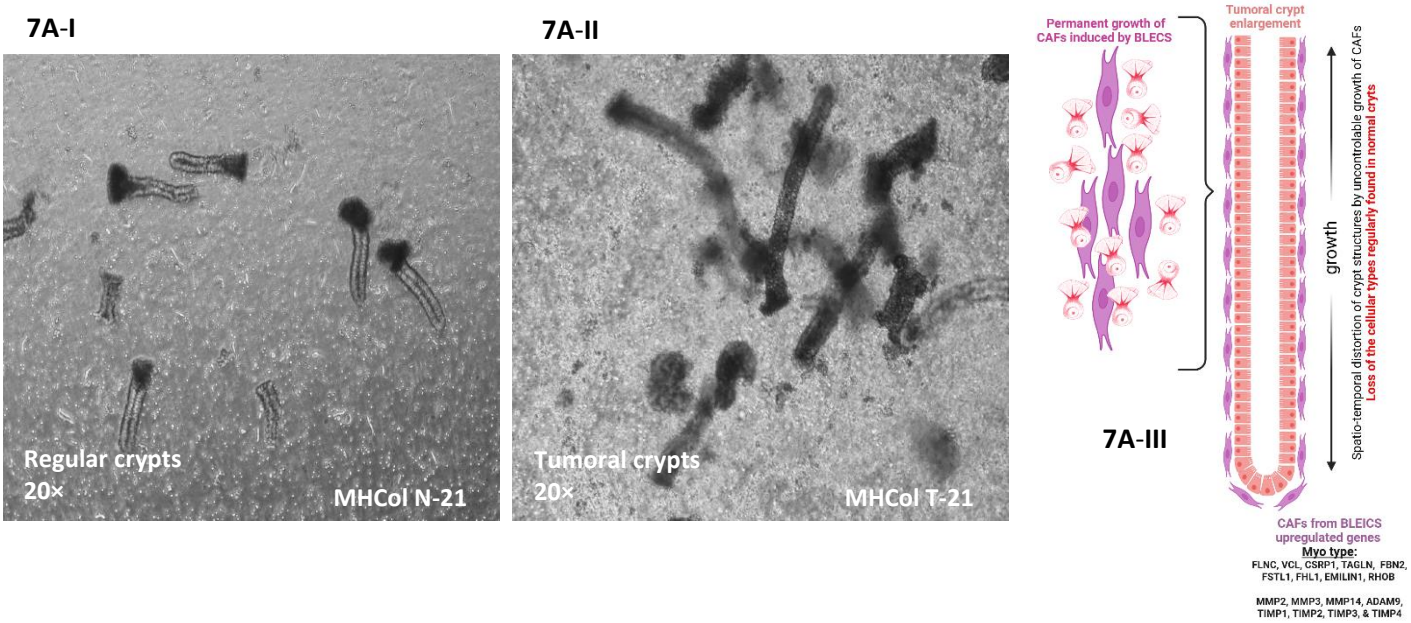



**Figure 6**: **B-lymphocyte, EVB-infected**, **Calamari-shaped** (**BLEICS**) **cell type associated with colorectal carcinomas.** **Panel A**: Morphological and phenotypical characterization of a B-lymphocyte population associated with colorectal carcinomas. Light microscopy revealed pleomorphic lymphoid characteristics of the population associated with the tumoral crypts. The cells show a body of approximately 5 µm with polarized petals (**Figures 6A-Ia-c**), as depicted in **Figures 6A-Ia-d**. These petals in the apical portion of the cells are well defined in SEM (**Figures 6A-IIa/IIb**). TEM ultrastructural analysis of these B-lymphocytes, revealed a concentration of electron-dense structures in the apical portion before the petals, congruent with fat droplets (**Figures 6A-IIc/IId**). We observed an intense interaction of this B-lymphocyte subpopulation with CAFs (**Figures 6A-IIIa-d** and the SI: Videography 13). Cytometric analysis of the cells revealed a B-cell phenotype, with expression of CD45 and CD19 (**Figures 6A-IV** & **6A-IVb,** SI: BLEICS Phenotype by FACS), as well as MHCII. BLEICS cells express the typically neuron-associated growth factor NT3 (**Figure 6A-IVd**) as well as some markers typically expressed by M cells, such as Gp2 (**Figure 6A-Va**), the transcription factor SpiB (**Figure 6A-Vb**), uromodulin (**Figure 6A-Vc**). They are labeled by UEA-1 lectin (**Figure 6A-Vd**). **Panel B**: Intracellular localization and dynamics of fusion of BLEICS cells with supporting fibroblasts. Time-lapse observations exposed a tight relationship between tumoral calamari-shaped B-lymphocytes and CAFs, revealing possible cell-cell fusions generating polykaryocytes, also observed using TEM (**Figures 6B-Ia-d** top picture). Sequential TEM sections were performed for a reconstruction of these polykaryocytes (**Figures 6B-Ia-d** bottom picture). The admixture of cells is observed in the videographic sequence **Figures 6B-IIa-l** (SI Videography 12). **Panel C**: BLEICS cells have a transcriptional background with a mixed B-DC cell phenotype and multiple pathogen infections. NGS results for the 100 most highly expressed genes show, that BLEICS cells are phylogenetically different from normal B-lymphocytes isolated from 10 different individuals, as well as from a lymphoma B cell line (Namalwa), as judged by the PCA set of variables (**Figure 6C-Ia**). Deconvolution analysis of B-lymphocytes compared with the BLEICS cells, using the CIBERSORTx tool revealed that BLEICS cells largely retain a genotype congruent with memory and plasma cells, but also a significant genotype characteristic of myeloid cells, specifically of activated DCs (**Figure 6C-Ib**). BLEICS cells revealed a signature congruent with previous infections, such as EBV, HTLV-I, influenza A, Kaposi's sarcoma-associated herpesvirus, measles, some pathogenic bacterial infections such as Mycobacterium, S. aureus, and some parasitic infections like T. gondii and Leishmania (**Figure 6C-Ic** & **6C-Ie**). Fibroblasts revealed a plethora of genes commonly found in the biology of muscle development and differentiation such as EMILIN1, RHOB, FHL1 and FSTL1 among others as reflected in the heatmap (**6C-Id**). **Panel D**: BLEICS cells show a peculiar expression pattern of neuronal genes and all relevant structural EBV genes. Having detected expression of the NT3 in calamari-shaped B-lymphocytes, we examined these cells for a panel of neuronal and stemness markers. We noted that these B-lymphocytes showed detectable expression of neuronal markers like ACHE, BDNF, GAD67, N-cadherin, Neurog2, NSE, PSD95, SCG5, and synaptophysin. Moreover, the cells expressed some stem cell markers such as HES1, HIF1a, Klf4, Nanog, Notch1, Nurr1, and vimentin, indicating EMT. Stathmin, a structural protein, was highly expressed by these cells (**Figure 6D-I**). qPCR confirmed the presence of Epstein–Barr virus (EBV) transcripts—EBV N capsid, LMP1, and EBNA1—in BLEICS cells (**Figure 6D-II**), confirming prior RNA-Seq findings (**3C-Ic**). BLEICS cells in fact stimulate the growth of their cognate fibroblasts as seen in **Figure 6D-III**. BLEICS 21 cells promote CAF proliferation, as demonstrated by ICC showing a marked increase in Ki67 expression, particularly within α-SMA-positive cells under co-culture conditions (**Figure 6D-IV**). CAFs were cultured for three days either alone (top) or in the presence of BLEICS 21 cells (bottom), followed by ICC staining to assess proliferative activity. **Panel E**: Protein expression of neuronal, stem cell, and immune-related markers in BLEICS cells and colon CAFs. To validate the neuroimmune phenotype of calamari-shaped B-lymphocytes, we looked for the expression of some markers associated with both (neuronal and immune) phenotypes, using immune techniques (**Figure 6E-Ia**). Western blots revealed the expression of stemness-related markers (green), neuronal-related markers (yellow), and an immune phenotype (red) in calamari-

shaped B-lymphocytes (**Figure 6E-Ia** and the respective densitometric analysis in **Figure 6E-Ib**, which are the result of the semi-quantitative measures against housekeeping proteins like β-actin and GADPH). Some of these markers are also found in CAFs obtained from primary CRC cultures, revealing the expression of β-catenin, Sox9, and S100A4, among others (**Figures 6E-Ib** & **6E-Id**). **Panel F**: Regulation of human endogenous retroviruses (HERVs) in BLEICS. Differential expression analysis revealed several significantly modulated human endogenous retroviruses (HERVs) in tumor-derived crypts. The most differentially expressed loci are shown in the bar plots (**Figure 6F-I**) of transcripts per million (TPM) differences between conditions. **Table 2** lists all significantly upregulated and downregulated HERVs with an immunosuppressive domain (ISD), as described by Vargiu et al.[40], along with annotations including subgenomic structure, ISD sequence, best-matching reference elements, and corresponding HERV supergroup classifications. To further illustrate conserved sequence features, sequence logos were generated for the ISD motifs of upregulated (left) and downregulated (right) HERVs using WebLogo (https://weblogo.berkeley.edu/logo.cgi) and are shown below the **Table 2**. Results are representative of n = 4 experiments. BLEICS: B-Lymphocyte, EVB-Infected, Calamari-Shaped.

**Panel A**: Spatiotemporal deformation of the colonic crypts during malignization




**Panel B**: Proposed model of cellular interactions and crypt remodeling during colorectal tumorigenesis

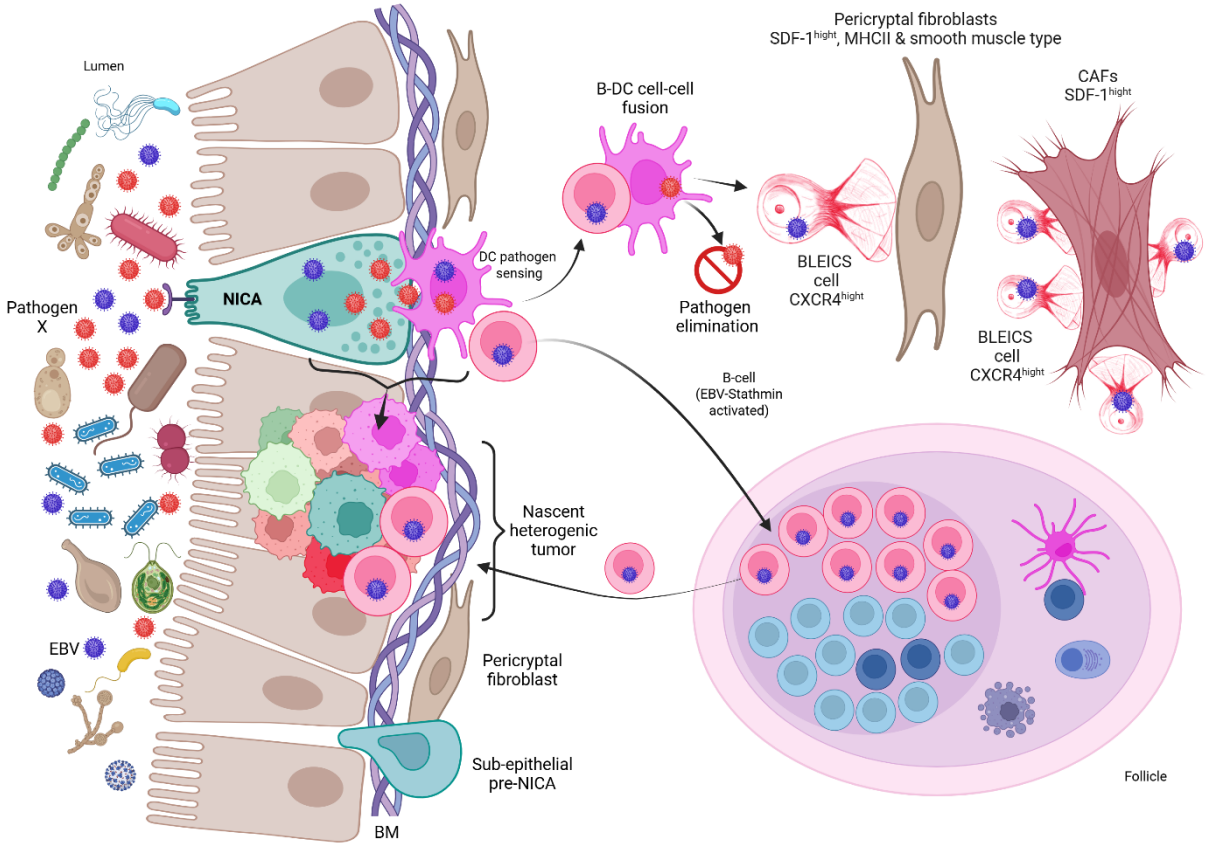


**Figure 7**: **Main events during malignization**: **Panel A**: Spatiotemporal deformation of colonic crypts during malignization. A comparison between regular and tumor crypts from the same individual reveals that tumoral crypts are enlarged up to threefold (**7A I-II**). As shown (**Figure 3D-III**), BLEICS cells are critical for the sustained growth of cancer-associated fibroblasts (CAFs). Their removal significantly slows CAF growth *in vitro*, eventually leading to irreversible senescence. This cell-cell interaction may contribute to the aberrant enlargement of tumoral crypts, as depicted in **Figure 7A-III**. CAFs were found to express higher levels of genes associated with smooth muscle biology, suggesting that the majority of these CAFs may be myofibroblast or pericryptal fibroblasts, which play a pivotal role in cancer progression. Additionally, these CAFs, like the tumoral crypts, overexpress matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs), specifically MMP3, MMP14, and TIMP1 (SI: Metalloproteinases). Thus, CAF growth forced by BLEICS cells may explain the stromal remodeling observed in tumors and the resultant spatiotemporal deformation of the crypt architecture. **Panel B**: Schematic illustration of a possible path to carcinogenesis in colonic crypts. NICA cells could originate from subepithelial compartments (pre-NICA), rather than from the colonic crypts stem cells, integrating into crypt structures to access lumen pathogens after molecular signaling. Functioning similarly to M cells, NICA cells express various pathogen-recognition molecules, potentially serving as a gateway for transit infections. Pathogens like EBV attract dendritic cells, which become activated, and recruit B-lymphocytes in response. Upon recruitment, B-lymphocytes become infected with EBV, triggering the expression of stathmin. Their interaction with activated dendritic cells leads to the formation of BLEICS ($CXCR4^{high}$) cells. Stimulated by fibroblast-secreted SDF-1, BLEICS cells engage pericryptal fibroblasts, prompting their differentiation into CAFs and ultimately drive crypt elongation. BLEICS cells infiltrate emerging tumor tissues. Additionally, fusion between infected NICA cells and dendritic cells may cause genetic instability, potentially leading to carcinogenesis and diverse tumor cell populations. Tumor cells display markers unique to NICA cells and possess MHC-II, likely acquired from dendritic cells rather than B cells, as colorectal carcinoma cells lack immunoglobulin expression. Cartoons were generated using BioRender.com.