Supplementary materials and methods

Patients, tissue specimens and peripheral blood

This study was approved by the Ethics Committee of Nanfang Hospital, Southern Medical University. One hundred and five NPC patients were recruited in this study. All NPC patients gave their informed consent to participate in the study. All patients were initially diagnosed by pathology and received standardized treatment between 2005 and 2009 at Nanfang Hospital, Southern Medical University, Guangzhou, China. The clinical characteristics of the patients are listed in supplementary Table 3. All tissue specimens were collected at the time of diagnosis before any treatment and were immediately frozen in liquid nitrogen.

The follow-up data were summarized at the end of December, 2018 at Nanfang Hospital, Southern Medical University, Guangzhou, China. All patients were evaluated every 2-3 months during the first year, every 6 months during the second to fifth year and every year thereafter. All follow-up examinations were performed by physicians who were unaware of this study. Recurrence and death were the primary endpoints. The time to recurrence was calculated from the date of completion of the treatment to the date of diagnosis of tumour recurrence. The overall survival was calculated from the date of completion of the treatment to the date of death or last follow-up.

The peripheral blood used in this study was donated by the healthy staff of this research group. All volunteers gave informed consent to participate this study.

Cells

Human embryonic kidney 293T cells and mouse hepatoma Hepa1-6 cells were purchased from Cell Bank of Chinese Academy of Sciences. Human NPC C666-1 (EBV+) cells and hepatocellular carcinoma HCCLM3 cells were purchased from Shanghai Biological Technology Co., Ltd (Shanghai, China). The immortalized human nasopharyngeal epithelial cell line NP69 and four human NPC cell lines, including 5-8F, CNE1, CNE2 and 6-10B, were obtained from the Clinical Research Center of Nanfang Hospital, Southern Medical University. Human regulatory T (Treg) cells expressing $CD4^{+}CD25^{+}CD127^{low/-}$ biomarkers were isolated and purified by FACS from the peripheral blood mononuclear lymphocytes (PBMC) of healthy volunteers. All cells were cultured under standard conditions.

Reagents and plasmids

The miR-200a mimics, miR-200a inhibitor, negative control miRNA mimics and negative control miRNA inhibitor were purchased from RiboBio Co., Ltd. (Guangzhou, China) and used according to the manufacturer's instruction. All plasmids used in this study were constructed by Genecopoeia Co., Ltd. (Guangzhou, China) except the dual luciferase reporter vectors carrying wild type or mutated sequences of the CXCL12 3'UTR, which were constructed by RiboBio. All plasmid sequences were verified by sequencing. SB431542 and AMD3100 were purchased from Selleck.

Real time PCR

Total RNA was extracted using RNAzol® RT RNA isolation reagent (Genecopoeia), and reverse transcription and fluorescence quantitative detection were performed using an All-in-One qPCR kit (Genecopoeia) according to the manufacturer's instruction. miRNA reverse transcription and fluorescence quantitative detection were performed using an All-in-One miRNA qPCR kit (Genecopoeia) according to the manufacturer's instruction. GAPDH and U6 genes were used as mRNA and miRNA internal controls,

respectively. All tests were performed in triplicate. Please refer to Supplementary Table 1 for the details of the primer sequences.

Immunohistochemistry

Immunohistochemistry was performed to detect the expression of the target genes in tissues using an All-in-One IHC kit according to the manufacturer's instruction (Beijing zhongshan Jinqiao Biotechnology Co., Ltd). Please refer to Supplementary Table 2 for detailed description of antibodies.

Analysis was performed by two independent pathologists who were blinded to the clinical outcome. The EBNA1 protein is encoded by the EB virus gene; hence, it is not encoded by the human genome. Therefore, an all-or-nothing criterion was used to determine whether the tissue sample is EBNA1-positive or –negative, and a total score system was used to determine the level of EBNA1 expression. The expression level of EBV-EBNA1 was scored as a proportion of immunopositive staining area multiplied by staining intensity. The proportion of immunopositive staining area was based on the percentage of positive tumour cells: 0 (0%), 1 (1%-25%), 2 (26%-50%), 3 (51%-75%) or 4(76%-100%). Staining intensity was scored as 0 (negative), 1 (weak), 2 (moderate) or 3 (intense). Therefore, each case was ultimately considered negative if the total score was 0 and positive if the final score was ≥1. The total score 0 was defined as negative expression of EBNA1 protein and marked with symbol '－'. The total score from 1 to 3 was defined as low expression of EBNA1 protein and marked with symbol '+'. The total score ≥4 were defined as high expression of EBNA1 protein and marked with symbol'+ +'; score ≥4 was marked with symbol '+ + +'.

The TGFβ1, SMAD3, p-SMAD3, c-JUN, p-c-JUN and CXCL12 proteins are encoded by the human genome. Therefore, a total score system was used to determine the levels of expression of human proteins. The expression levels of human proteins were scored as proportion of immunopositive staining area multiplied by staining intensity. The proportion of immunopositive staining area was based on the percentage of positive tumour cells: 0 (0%), 1 (1%-25%), 2 (26%-50%), 3 (51%-75%) or 4 (76%-100%). Staining intensity was scored as 0 (negative), 1 (weak), 2 (moderate) or 3 (intense). Therefore, each case was ultimately considered negative expression if the total score was 0 (－), low expression if the total score was from 1 to 2 (±) or 3 (+) and high expression if the total score was 4 (+ +) or ＞4 (+ + +).

The infiltrating lymphocytes around the tumour cells were identified as Treg cells by positive Foxp3 staining. To grade the $Foxp3^+$ Treg cells, photomicrographs with 100 positive cells per 400× high-power field were acquired and used as a reference for grading. Three representative fields were selected at 400× magnification. The results were counted manually in a single 400× microscopic field for each area of every specimen. Representative areas were considered negative if no $Foxp3^+$ Treg cells infiltrated, low density infiltration if less than or equal to 100 positive cells (≤100) per field were detected and high density infiltration if more than 100 cells (＞100) per field were detected.

***In vitro* migration assay of Treg cells**

Peripheral blood mononuclear cells (PBMCs) from healthy volunteers were resuspended in PBS and stained with CD4, CD8, CD25 and CD127 monoclonal antibodies for 15-30 min. CD4 positive, CD25 positive, CD127 negative or low expression and CD8 negative cells ($CD4^+$ $CD25^+$ $CD127^{-/low}$ $CD8^-$) were sorted by flow cytometry and identified as Treg cells. Treg cells (1×10^5) were resuspended in 100 μl serum-free medium and were added to the upper chamber of 5-μm pore size Transwell inserts (Corning). The Transwell chambers were then moved to 24-well plates containing 600 μl

cell-conditioned medium and incubated at 37°C for 4-5 h. Cells were fixed in 4% paraformaldehyde and stained with 0.1% crystal violet. Migrated cells attached to the lower surface of the chamber membrane were quantified using 3 random fields. Three independent experiments were performed for each assay.

Animal study

All animal studies complied with the protocols approved by Southern Medical University Animal Care and Use Committee.

(1) Generation of humanized immune reconstitution mouse model of NPC to verify the associations of EBV-EBNA1, miR-200a and accumulation of Treg cells.

Human NPC cells (5×10^5, negative control group: 5-8F-NC; EBNA1 overexpression group: 5-8F-EBNA1; miR-200a overexpression group: 5-8F-miR-200a; EBNA1 and miR-200a overexpression group: 5-8F-EBNA1-miR-200a) were injected subcutaneously into the flanks of nude mice (4-week-old female mice, n=5/group), and tumour growth was measured twice weekly using calipers (tumour volume was calculated using the equation $v = (width^2\times length)/2$). Sixteen days after tumour transplantation, depletion of NK cells and subsets of monocyte/macrophages was achieved by intraperitoneal injection of 50 μg of an anti-asialo-GM1 antibody (eBioscience, Cat. 16-6507-39). Before the end of experiment, repeated injections were necessary if the interval was more than a week (weekly injections of anti-asialo-GM1 antibody if necessary). At the same time, doxycycline (4 mg/kg•d) was intragastrically administered daily to prevent infection. After 3 days of NK cell and monocyte/macrophage depletion, humanize immune reconstitution was performed by tail intravenous injection of human PBMC, and a total of 1×10^7 cells were administered in two injections at 8 h interval (5×10^6 cells/injection×2 injections). After 50-60 h of humanized immune reconstitution, xenografts were harvested and IHC detection and FACS analysis were performed.

(2) Generation of humanized immune reconstitution mouse model of NPC to verify the associations of the TGFβ1-SMAD3 axis and accumulation of Treg cells.

Human NPC cells (5×10^5; 5-8F-EBNA1) were injected subcutaneously into the flanks of nude mice (4-week-old female mice, n=5/group), and tumour growth was measured twice weekly using calipers (tumour volume was calculated using the equation $v = (width^2\times length)/2$). Five days after tumour transplantation, SB431542 or PBS (negative control) were injected in the tumours every other day. Sixteen days after tumour transplantation, depletion of NK cells and subsets of monocyte/macrophages was achieved by intraperitoneal injection of 50 μg of an anti-asialo-GM1 antibody (eBioscience, Cat. 16-6507-39). Before the end of experiment, repeat injections were necessary if the interval was more than a week (weekly injections of anti-asialo-GM1 antibody if necessary). At the same time, doxycycline(4 mg/kg•d) was intragastrically administered daily to prevent infection. After 3 days of NK cell and monocyte/macrophage depletion, humanize immune reconstitution was performed by tail intravenous injection of human PBMC, and a total of 1×10^7 cells were administered in two injections at 8 h intervals (5×10^6 cells/injection×2 injections). After 50-60 h of humanized immune reconstitution,

xenografts were harvested and FACS analysis was performed.

(3) Generation of humanized immune reconstitution mouse model of NPC to verify the associations of the CXCL12-CXCR4 axis and accumulation of Treg cells.

Human NPC cells (1×10^6; 5-8F-EBNA1) were injected subcutaneously into the flanks of nude mice (4-week-old female mice, n=5/group), and tumour growth was measured twice weekly using calipers (tumour volume was calculated using the equation $v = (width^2\times length)/2$). Five days after tumour transplantation, recombinant human CXCL12 protein (1 μg/kg; every other day) or AMD3100 (5 mg/kg daily) were administered by injection in the tumours. Sixteen days after tumour transplantation, depletion of NK cells and subsets of monocyte/macrophages was achieved by intraperitoneal injection of 50 μg of an anti-asialo-GM1 antibody (eBioscience, Cat. 16-6507-39). Before the end of experiment, repeat injections were necessary if the interval was more than a week (weekly injections of anti-asialo-GM1 antibody if necessary). At the same time, doxycycline(4 mg/kg•d) was intragastrically administered daily to prevent infection. After 3 days of NK cell and monocyte/macrophage depletion, humanize immune reconstitution was performed by tail intravenous injection of human PBMC, and a total of 1×10^7 cells were administered in two injections at 8 h intervals (5×10^6 cells/injection×2 injections). After 50-60 h of humanized immune reconstitution, xenografts were harvested and FACS analysis was performed.

(4) Generation of humanized immune reconstitution mouse model of human hepatocellular carcinoma (HCC) to verify the associations of the CXCL12-CXCR4 axis and accumulation of Treg cells.

Human HCC cells (1×10^6; HCCLM3) were injected into the abdominal cavity of nude mice (4-week-old female mice, n=5/group) , and mice weight was measured twice weekly. Five days after tumour transplantation, recombinant human CXCL12 protein (1 μg/kg; every other day) or AMD3100 (5 mg/kg daily) were administered by intraperitoneal injection. Sixteen days after tumour transplantation, depletion of NK cells and subsets of monocyte/macrophages was achieved by intraperitoneal injection of 50 μg of an anti-asialo-GM1 antibody (eBioscience, Cat. 16-6507-39). Before the end of experiment, repeat injections were necessary if the interval was more than a week (weekly injections of anti-asialo-GM1 antibody if necessary). At the same time, doxycycline(4 mg/kg•d) was intragastrically administered daily to prevent infection. After 3 days of NK cell and monocyte/macrophage depletion, humanize immune reconstitution was performed by tail intravenous injection of human PBMC, and a total of 1×10^7 cells were administered in two injections at 8 h intervals (5×10^6 cells/injection×2 injections). After 50-60 h of humanized immune reconstitution, xenografts were harvested and FACS analysis was performed.

(5) Generation of immunocompetent syngeneic mouse model of mouse hepatocellular carcinoma (HCC) to verify the associations of the CXCL12-CXCR4 axis and accumulation of Treg cells.

Mouse HCC cells (1×10^6; Hepa1-6) were injected into the abdominal cavity of C57BL/6 mice

(4-week-old female mice, n=5/group) and mice weight was measured twice weekly. Five days after tumour transplantation, recombinant human CXCL12 protein (1 μg/kg; every other day) or AMD3100(5 mg/kg daily) were administered by intraperitoneal injection. Twenty-one day after tumour transplantation, xenografts were harvested and FACS analysis was performed.

Statistical analysis

All statistical analyses were performed using the Graphpad Prism (version 5.0) statistical software package except the correlation analysis of clinical samples, which was performed using the SPSS (version 13.0) statistical software package. The Bivarate correlate model and Pearson correlation coefficient were utilized to evaluate the relationship between EBNA1 expression and clinicopathological characteristics in NPC patients. The Linear regression model and Pearson correlation coefficient were utilized to evaluate the relationship between Treg cells infiltration level (the dependent variable) and clinicopathological characteristics (the independent variables) in NPC patients. Survival analysis was performed using the Kaplan-Meier method. The Student's t-test was used for comparisons of two independent groups. The One-way Analysis of Variance (ANOVA) was used for comparisons of multiple groups. The Repeated Measures of General Linear Model was used for comparisons of tumour growth. A p value of less than 0.05 was considered statistically significant (*$p<0.05$; **$p<0.01$; ***$p<0.001$). Data are presented as mean±SD.