Merkel cell carcinoma in southern China: clinicopathologic and transcriptomic analysis with insights into an intraepidermal-predominant lesion

Shuting Shen[1]*, Yuhang Pan[1]*, Bin Li[1]*, Jiafeng Li[2], Xin Zeng[1], Yutao Zhong[1], Chunkui Shao[1], Yang Xie[3]‡, Na Cheng[1]‡

[1] Department of Pathology, The Third Affiliated Hospital of Sun Yat-sen University, No. 600 Tianhe Road, Guangzhou 510630, China

[2] Department of Pathology, Dermatology Hospital, Southern Medical University, No. 2 Lujing Road, Guangzhou 510091, China

[3] Department of Dermatology, The Third Affiliated Hospital of Sun Yat-sen University, No. 600 Tianhe Road, Guangzhou 510630, China

* These authors contributed equally to this work.

‡ Corresponding authors

Corresponding author

Yang Xie

Department of Dermatology, The Third Affiliated Hospital of Sun Yat-sen University, No. 600 Tianhe Road, Guangzhou 510630, China

E-mail: xiey25@mail.sysu.edu.cn

Na Cheng

Department of Pathology, The Third Affiliated Hospital, Sun Yat-sen University, No. 600 Tianhe Road, Guangzhou 510630, China

Tel: +86-20-85253437

Fax: +86-20-85253130

E-mail: chengna2@mail.sysu.edu.cn

Supplementary Methods

Transcription Factor Activity Analysis

To further characterize regulatory programs in MCC, transcription factor activity was inferred using the decoupleR package with a univariate linear model. Regulatory interactions were obtained from the CollecTRI database. Differentially inferred transcription factor activities between tumor and adjacent normal tissues were used for exploratory analysis and data visualization.

Immune Microenvironment Analysis

To explore immune-related transcriptional features, enrichment scores for 28 immune cell populations were estimated using single-sample gene set enrichment analysis (ssGSEA) implemented in the GSVA package. Immune-related functional signatures and differential expression of immune checkpoint-related genes were analyzed using ggstatsplot. These analyses were considered exploratory and were used to provide additional context regarding the tumor immune microenvironment.

Exploratory Mutation Analysis Based on RNA-seq Data

To explore potential coding variants in MCC, an exploratory mutation analysis was performed using RNA-seq data. Reads were aligned to the human reference genome (GRCh38) using STAR in two-pass mode. Alignment files were processed using GATK (v4.6.2), including duplicate marking with MarkDuplicates and exon junction read processing with SplitNCigarReads, followed by variant calling with MuTect2.

Variants were filtered using FilterMutectCalls and annotated with ANNOVAR. Additional filtering criteria were applied to retain high-confidence variants, including minimum read support of at least 3, sequencing depth of at least 8, tumor log odds (TLOD) greater than 6.3, localization in coding regions, population allele frequency less than 0.001 in the 1000 Genomes database, and inclusion in the IntOGen cancer driver gene database.

Because variant detection was performed from RNA-seq data without orthogonal DNA-based validation, such as whole-exome or targeted sequencing, these findings were considered exploratory and were used primarily to provide an overview of the putative coding variant landscape. Variants were visualized using the maftools package.

Table S1. Antibodies used for immunohistochemical analysis

Antibody	Clone	Manufacturer	Dilution
Pan-cytokeratin(CK)	AE1/AE3	MXB Biotechnologies	1:100
Cytokeratin 7 (CK7)	MX053	MXB Biotechnologies	1:100
Cytokeratin 20 (CK20)	MX059	MXB Biotechnologies	1:100
Epithelial Membrane Antigen (EMA)	E29	MXB Biotechnologies	1:100
Carcinoembryonic Antigen (CEA)	MX068	MXB Biotechnologies	1:200
Gross Cystic Disease Fluid Protein 15 (GCDFP-15)	MX120	MXB Biotechnologies	RTU
$p16^{INK4a}$ (p16)	E6H4	Roche Diagnostics	RTU
P63	MX013	MXB Biotechnologies	1:100
Ber-EP4 (EA)	Ber-EP4	MXB Biotechnologies	RTU
Human Melanoma Black-45 (HMB-45)	HMB45	Vision BioSystems	1:100
Melanoma Antigen (Melan A)	A103	Vision BioSystems	1:50
Chromogranin A (CgA)	MX018	MXB Biotechnologies	1:100
S100	Polyclonal	Vision BioSystems	1:100
Neural cell adhesion molecule 1 (CD56)	CD564	Vision BioSystems	1:100
Synaptophysin (Syn)	27G12	Vision BioSystems	1:100
p40	BC28	ZSGB-BIO	1:100
p53	DO-7	Vision BioSystems	1:100
Retinoblastoma 1 (RB1)	1F8	MXB Biotechnologies	1:100
Thyroid transcription factor-1 (TTF-1)	SPT24	QuanHui Interntional	1:200
Ki-67	MX006	MXB Biotechnologies	1:100
Merkel cell polyomavirus (MCPyV)	CM2B4	Millipore	1:100

Abbreviations: RTU, ready to use.

Supplementary Results

Transcription factor activity analysis

Exploratory transcription factor activity analysis identified altered regulatory programs in MCC compared with normal skin. Increased inferred activity was observed in several cell cycle-associated regulators, whereas reduced activity was noted in selected regulators related to stress and inflammatory signaling (Figure S1).

Immune-related transcriptomic features

Exploratory immune-related analyses showed heterogeneous immune functional signatures across MCC samples. ssGSEA suggested variable enrichment of immune cell populations and immune-related pathways between tumor and normal tissues, and selected immune checkpoint-related genes also showed differential expression (Figure S2).

Exploratory mutation analysis

Exploratory RNA-seq-based variant analysis identified multiple coding variants across MCC samples, with nonsynonymous SNVs representing the predominant mutation class (Figures S3 and S4). Recurrently altered genes included several involved in chromatin remodeling and related regulatory pathways. Because these findings were inferred from RNA-seq data without orthogonal DNA-based validation, they should be interpreted cautiously.

Additional heatmap analysis of the intraepidermal-predominant case

A heatmap including normal skin, intraepidermal tumor, and invasive tumor samples from case 1 showed that the intraepidermal component shared a transcriptomic profile with MCC while remaining intermediate between normal skin and invasive tumor (Figure S5).

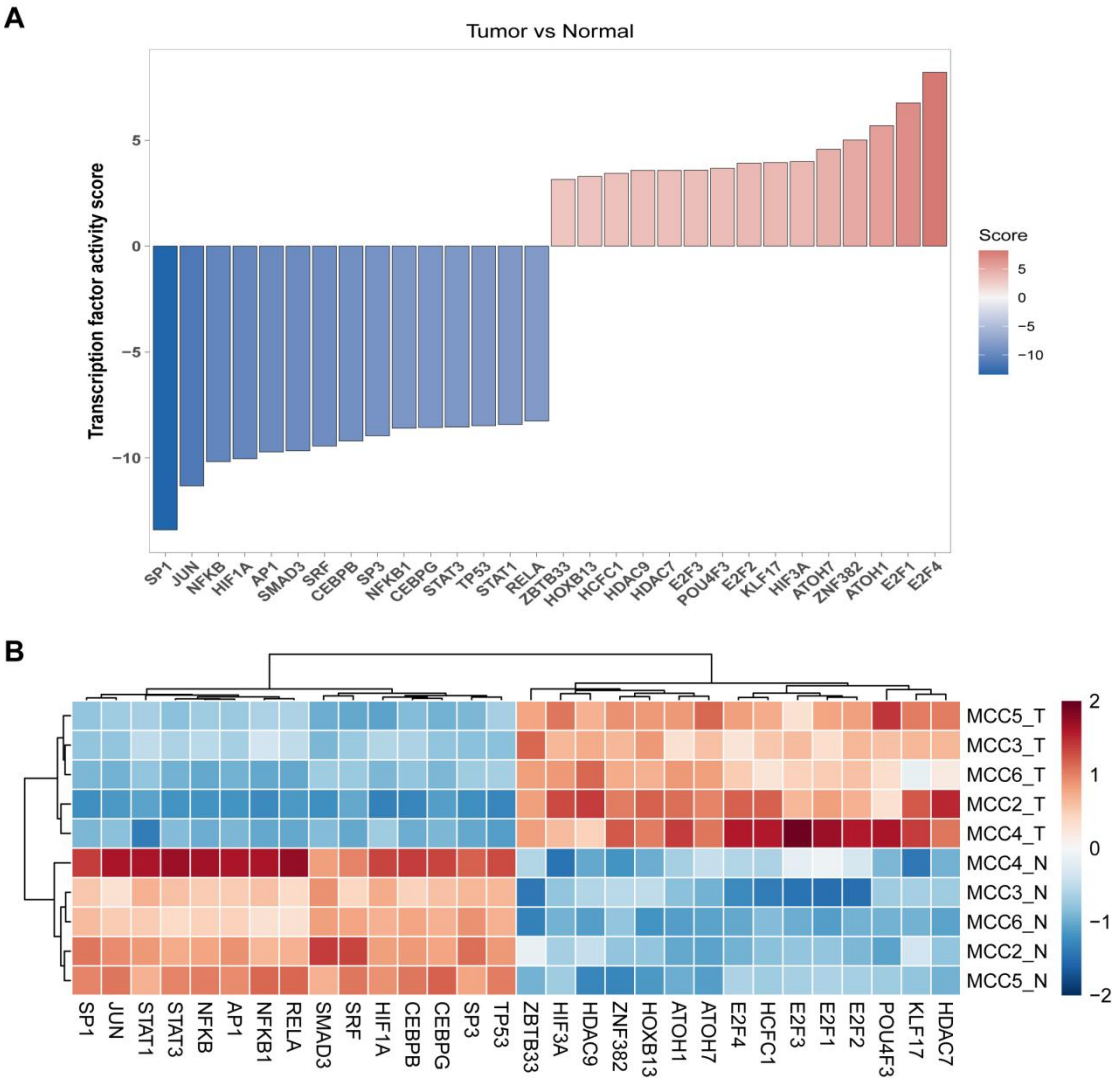

Figure S1. Transcription factor activity analysis in MCC.

(A) Bar plot showing inferred transcription factor (TF) activity scores in MCC tumors compared with normal skin, estimated using the decoupleR framework with regulatory interactions curated from the CollecTRI database. (B) Heatmap displaying TF activity patterns across individual tumor and normal samples.

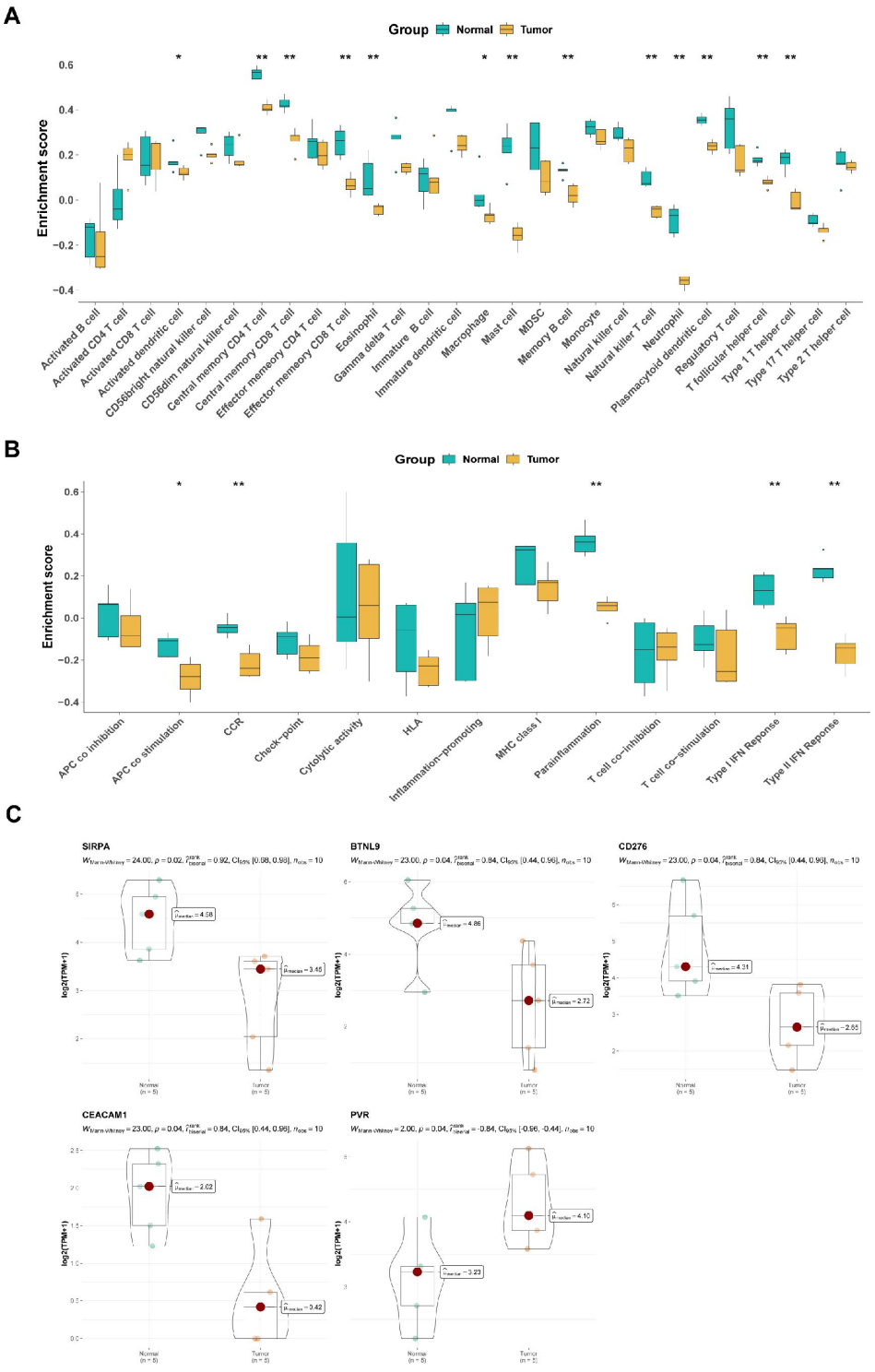

Figure S2. Immune microenvironment characteristics of MCC inferred from transcriptomic data.

(A) Comparison of immune-related functional signatures between MCC tumors and normal skin based on ssGSEA enrichment scores. (B) Relative abundance of immune cell populations inferred by ssGSEA. (C) Expression differences of representative immune checkpoint-related genes between tumor and normal samples.

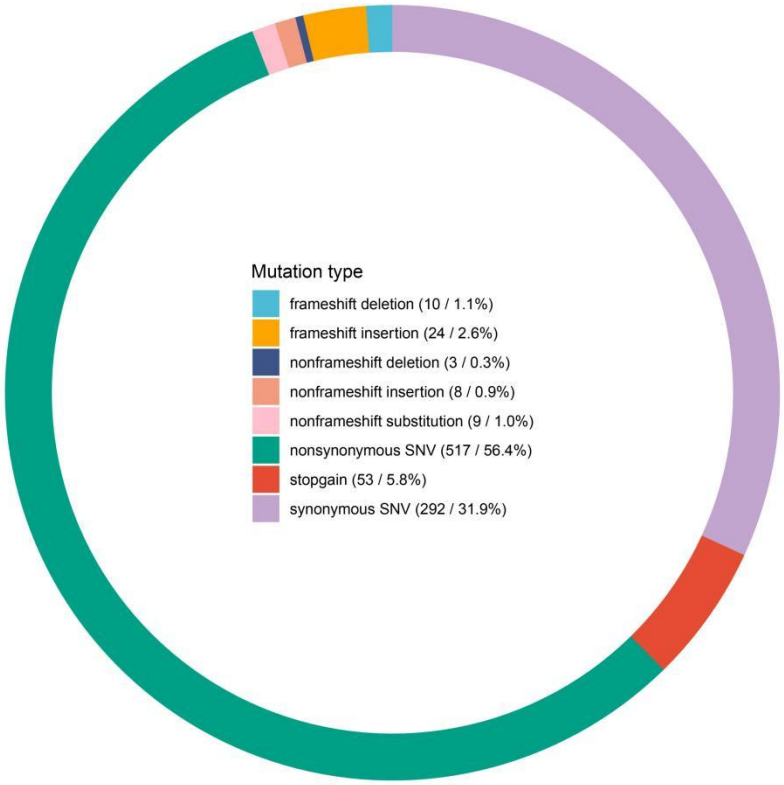

Figure S3. Distribution of mutation types identified from RNA-seq variant calling in MCC samples.

Donut chart illustrating the relative proportions of mutation classes, including nonsynonymous SNVs, synonymous SNVs, stop-gain mutations, frameshift insertions/deletions, and nonframeshift insertions/deletions. Total variants analyzed: 916.

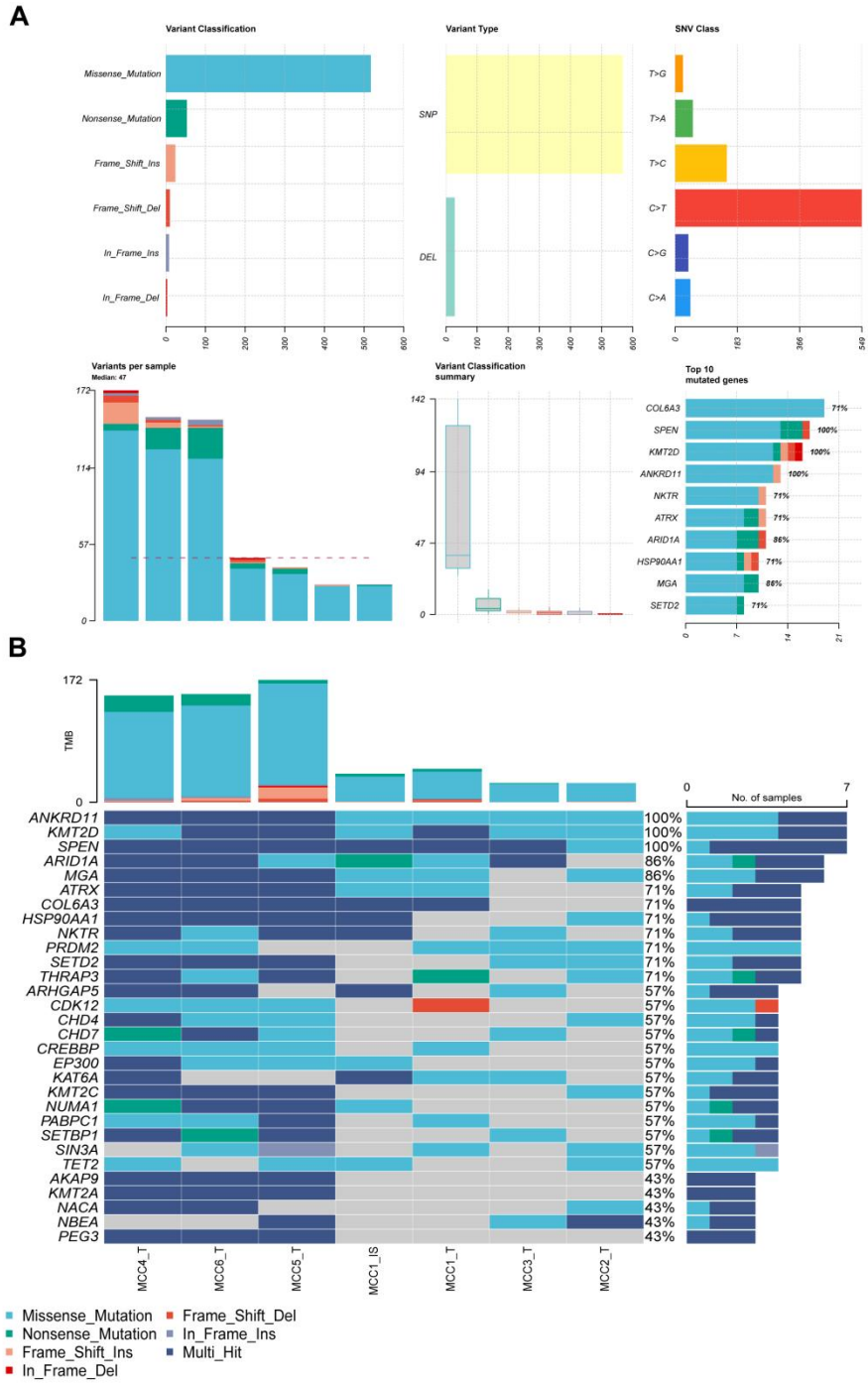

Figure S4 Mutational landscape of MCC samples inferred from RNA-seq variant calling.

(A) Overview of mutation characteristics across MCC samples, including variant classification, variant type distribution, SNV classes, variants per sample, and the most frequently mutated genes. (B) Oncoplot showing the distribution of somatic variants across MCC samples.

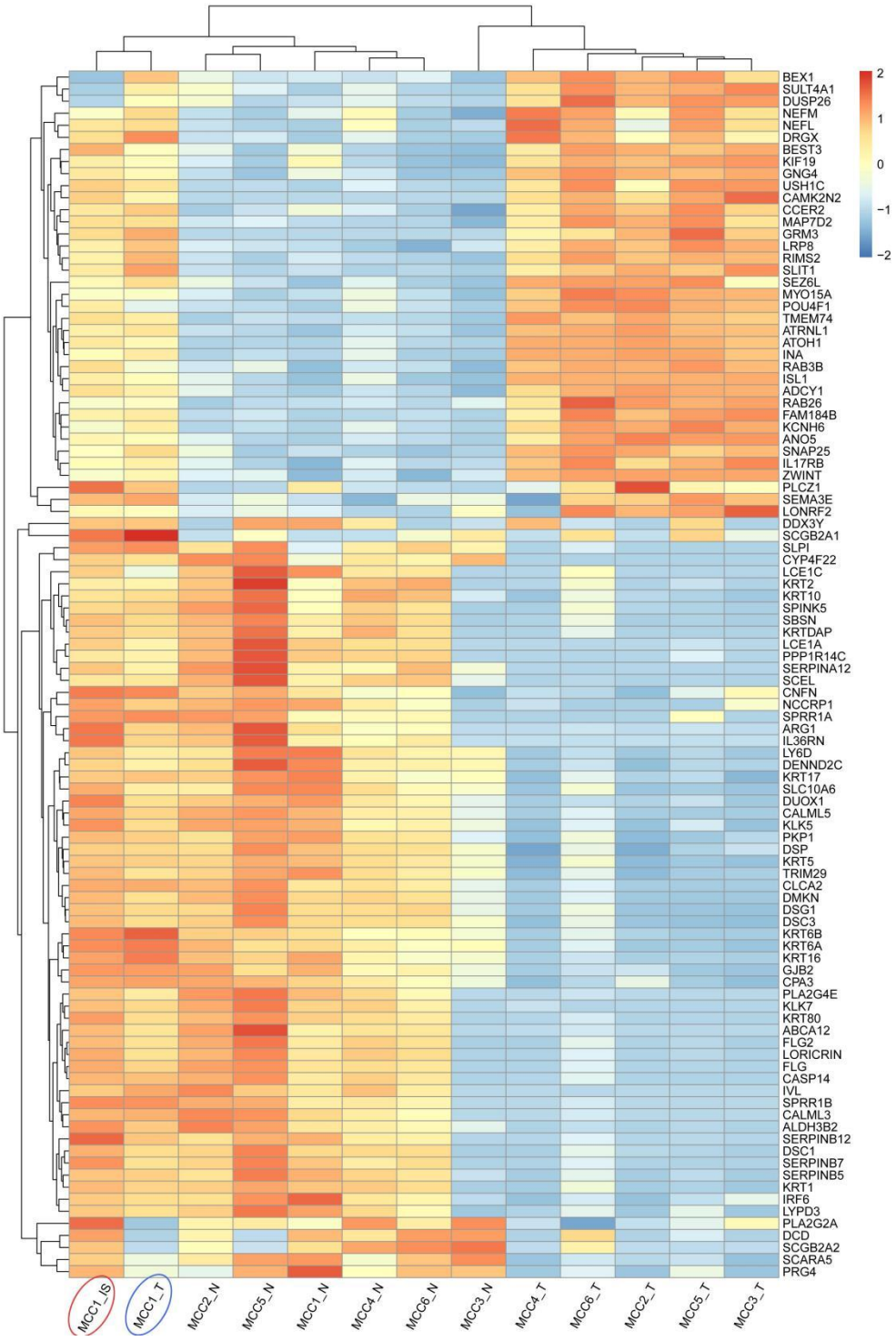

Figure S5. Heatmap of representative differentially expressed genes in MCC.
Hierarchical clustering heatmap showing representative differentially expressed genes across normal skin, intraepidermal tumor, and invasive MCC samples. Expression values are shown as scaled z-scores. Samples corresponding to the intraepidermal (MCC1-IS) and invasive (MCC1-T) components from case 1 are indicated by red and blue circles, respectively.