Supplementary Information

Supplementary Tables

Supplementary Table S1

List of genes differentially expressed in melanoma CTC cell lines compared with high purity (80%) primary melanomas and metastases available in the TCGA database and with CCLE melanoma cell lines (refer to Figures 1D-F and Figure S5A) and comparing Mel-167 and PEM-22 CTC lines to the matched tumor samples in each case (refer to Figure S5B).

Supplementary Table S2

Table listing genes and definitions for curated pathway signatures "SREBP_TARGET", "SREBF1_TARGET", "SREBF2_TARGET" and "FERROPTOSIS" (refer to Figures 1D-F, 2C, 3C, 6A-B, 7A-B and Supplementary Figures S2, S3B, S10C-D).

Supplementary Table S3

Table sheet #1 lists detailed GSVA pathway enrichment scores for all samples from melanoma CTC lines, Mel-167-matched primary tumors, PEM-22-matched metastases, TCGA high purity primary melanomas, metastases and melanoma cell lines from CCLE; Sheet #2 showing statistics for GSVA enrichment of CTC lines vs. CCLE lines (Primary tumor-derived) or CCLE lines (Metastasis-derived); Sheet #3 showing statistics for CTC lines vs. CCLE lines cultured with three different media conditions: 1. DMEM+10% FBS; 2. EMEM+10% FBS); 3. RPMI-1640+10%FBS (refer to Figures 1D-F).

Supplementary Table S4

Table showing GSEA pathway enrichment analysis of SREBF2 binding to promoters (refer to Figure 3C).

Supplementary Table S5

List of genes differentially expressed in *TF-KD* melanoma CTCs compared with shRNA-control cells (refer to Figures 5A, 6A-B and S10).

Supplementary Table S6

Table listing genes within Cluster 1 and Cluster 2 (shown in Figure 7A) across selected pathways identified from differential GSVA enrichment analysis of melanoma CTCs. Gene-level statistics and heatmaps for pathways with significantly higher enrichment scores in Cluster 2 vs Cluster 1 CTCs are shown in individual tab (see Figure 7A).

Supplementary Table S7

Table listing patients' clinical information including age, sex, tumor stage, treatment and therapeutic response as defined using RECIST1.1. Patients with RECIST-defined as stable disease were not included in the response/progression analyses (refer to Figure 7B). The testing results of five melanoma CTC lines in mouse models were also listed.

Supplementary Table S8

Table listing TCGA SKCM expression subtypes (see Figure S10E)

Supplementary Methods

Chromatin immunoprecipitation and real time PCR (ChIP-qPCR)

Cultured CTCs were washed with ice-cold phosphate buffered saline (PBS) and homogenized while on ice using lysis buffer (0.32M sucrose, 5 mM CaCl2, 3 mM MgAc, 0.1 mM EDTA, 10 mM Tris-HCL (pH 8.0), 1 mM DTT and 0.1% Triton X-100). The lysate was crosslinked using 1% formaldehyde for 10 min at room temperature and then quenched with glycine. ChIP was performed using Magna ChIP™ Hisens Chromatin Immunoprecipitation Kit (Millipore Cat#17-104-60) following the manufacturer's instructions. Briefly, the precipitated nuclei were resuspended in SCW buffer containing protease inhibitor cocktail, followed by sonication using a Covaris S2, yielding an average fragment size of 250bp. The chromatin was then immunoprecipitated using Protein A Dynabeads (In vitrogen) with 5mg of ChIP-grade antibody overnight at 4°C. The beads were repeatedly washed with high- and low-salt buffers, and the DNA fragments were then de-crosslinked using proteinase K at 65°C for 4 hrs, followed by ethanol precipitation. Real time q-PCR quantification was performed to assess enrichment of DNA fragments of the promoter regions of interest.

Immunohistochemistry (IHC) staining by Ventana system

Immunohistochemistry staining of melanosome triple cocktail markers (HMB45 + A103 + T311, Roche Cat# 06527787001) were applied to cyrosectioned tissue slides and staining was carried out using Ventana BenchMark Ultra System (Roche Diagnostics). Briefly, Sections were cut to 4 μm (Microm HM 355 S) and dry at 80° C for 15 min. Anti-melnosome antibody was diluted in 1:25 (antibody diluent from Ventana) and fill into a Ventana antibody dispenser. The Ventana staining procedure includes pretreatment with

Cell Conditioner 2 (pH 6) for 60 min (standard), followed by incubation with 1:25 diluted antibody at 37 °C for 32 min. Upon antibody incubation perform Ventana standard signal amplification, ultraWash, counter- staining with one drop of Hematoxylin for 4 min and one drop of bluing reagent for 4 min. For chromogenic detection we have used ultraView Universal DAB Detection Kit (Ventana).

IHC slides scanning and analyses

Slides were scanned at 40X resolution using an Aperio Scanscope slide scanner (Leica Biosystems). Image quantification was performed using the VIS image analysis software (Visiopharm Inc). A tissue finding app was applied to find all tissue regions in the image. Debris, bubbles, and tissue areas not from the site of interest were excluded by hand. A Bayesian classifier was used to identify tumor regions from non-tumor tissue. The features utilized the green channel from the RGB color space, the red channel of the RGB color space, the color deconvolution for hematoxylin, and the color deconvolution for the purple DAB stain. A poly smoothing spatial filter (3 x 3 pixels) was applied to remove noise from the red channel feature, followed by squaring of pixel values and a mean spatial filter (3 x 3 pixels) to enhance contrast. Post-processing was applied to fill holes and remove small objects. The application sampled the image at 50% to speed computation time. The result was a markup image with tumor areas and non-tumor areas delineated. For each image, the analysis application quantified the tumor area relative to the total tissue resected.

TF protein ELISA assay

Mel-167 CTCs and HepG2 liver cell culture supernatant and total cell lysates were prepared 48 hours after cells were plated into fresh media. Intracellular and secreted TF

protein was measured using a solid-phase sandwich Enzyme-Linked Immunosorbent Assay (ELISA) according to manufacturer instructions (Thermo Scientific Cat#: EHTF). TF protein copy number was calculated based on a standard curve generated using recombinant TF proteins with serial dilutions as provided in the ELISA kit.

Whole-exome sequencing

For whole-exome sequencing (WES), the AllPrep DNA/RNA Mini Kit (Qiagen, Hilden, Germany) was used for dual extraction of both genomic DNA and RNA. DNA was quantified in triplicate using a standardized PicoGreen® dsDNA Quantitation Reagent (Invitrogen, Carlsbad, CA) assay. The quality control identification check was performed using fingerprint genotyping of 95 common SNPs by Fluidigm Genotyping (Fluidigm, San Francisco, CA). Library construction was performed using the KAPA Library Prep kit, with palindromic forked adapters from Integrated DNA Technologies. All library construction, hybridization and capture steps were automated on the Agilent Bravo liquid handling system. Flowcells were sequenced utilizing Sequencing-by-Synthesis chemistry for HiSeq 4000 flowcells. Each pool of whole exome libraries was sequenced on paired 76 cycle runs with two 8 cycle index reads across the number of lanes needed to meet coverage for all libraries in the pool.

Somatic mutation calling

Exome sequencing data from CTC-derived cell lines were used to identify somatic single nucleotide variations (sSNVs) and somatic small insertions and deletions (sINDELs). Output from Illumina software was processed by the Picard and GATK toolkits developed at the Broad Institute. The BAM files were generated by aligning with bwa version 0.5.9 to the NCBI Human Reference Genome Build hg19. Prior to variant calling, the impact of

oxidative damage (oxoG) to DNA during sequencing was quantified as described previously(6). The cross-sample contamination was measured with ContEst(7) based on the allele fraction of homozygous SNPs, and this measurement was used in MuTect. From the aligned BAM files, somatic alterations were identified using a set of tools developed at the Broad Institute (www.broadinstitute.org/cancer/cga). The details of our sequencing data processing have been described previously(8,9). Following our standard procedure, sSNVs were detected using MuTect (version 1.1.6)(8); sINDELs were detected using Strelka (version 1.0.11)(10). Then an allele fraction specific panel-of-normals (PoN) filter was applied to filter false positive germline variants and common artifacts from mutation calls, which compares the detected variants to a large panel of normal exomes or genomes and removes variants that were observed in the panel-of-normals. All somatic mutations, insertions and deletions were annotated using Oncotator (version 1.4.1)(11). sSNVs and sINDELS in only 99 cancer genes (cancer gene panel in targeted sequencing) were used for mutation status analysis in 5 CTC lines for comparable mutation status comparison between sSNVs and sINDELs identified from 2 CTC lines (PEM-22 and PEM-78) using Anchored Multi-plexed PCR (AMP)(12) and those identified from 3 CTC lines (Mel-167, Mel-182-1, and Mel-182-2) by WES. The oncoplot of cancer gene mutations (COSMIC Cancer Gene Census)(13) in Mel-182-1 CTC line, two Mel-167 primary tumors, one Mel-167 parental CTC cell line, and six Mel-167 isogenic CTC clones was generated by the maftools (14).

General pipeline for RNA-seq analyses

RNA-seq reads were aligned to the GRCh38 build of the human genome reference using STAR 2.6.1d (15). Duplicated reads were identified and flagged using Picard Tools v2.18.24 MarkDuplicates. Gene-level TPM values were calculated using RNA-SeQC v2.2.2

(16). The GENCODE v26 annotation was used for the STAR alignment and all other quantifications. All methods were run as part of the pipeline developed for the GTEx Consortium(17). Gene-level quantifications were obtained by collapsing the GENCODE annotation to a single transcript per gene, using the procedure described in GTEx Consortium(17), All bulk RNA-seq samples (all CTC line samples and tumor samples from this study, TCGA SKCM, CCLE Skin) and single-cell RNA-seq (scRNA-seq) samples (early culture CTC line samples, primary CTC samples) were processed with this pipeline. Batch effect was then removed by the *removeBatchEffect* function in the R package 'limma' (v.3.38.2)(18).

Primary CTC single cell RNA-seq analysis

The protein coding gene (annotation from gencode.v26.GRCh38) level TPM data from putative primary CTC samples, CTC culture samples, and white blood cell samples were normalized and transformed using *Linnorm* function (DataImputation = TRUE, minNonZeroPortion = 0.75) in the R package 'Linnorm' (v. 2.6.1)(19). With this *Linnorm* transformed TPM data of 4,114 genes, hierarchical clustering was performed using the *Linnorm.HClust* function (input="Linnorm", method_dist = "pearson", method_hclust = "ward.D") in the R package 'Linnorm' (v. 2.6.1)(19). Putative primary CTC samples not clustered together with white blood cell samples by hierarchical clustering and expressing at least one marker gene among previously reported 19 melanoma CTC markers(1) above the maximum gene expression values in white blood cell samples (WBC) were considered as sequence-validated primary CTC samples. GSVA analysis for the primary CTC samples was performed on the linnorm transformed TPM data (*Linnorm* function with "DataImputation = TRUE"). Differential gene expression analysis for primary CTC samples (C2 vs. C1) was performed on the raw count data of 8,325 genes (genes with at least 5

reads in at least 5 samples) using ZINB-WaVE (Zero-Inflated Negative Binomial-based Wanted Variation Extraction; *zinbwave* function with "K=0") and DESeq2(20,21) (Supplementary Table S6).

Differential gene expression analysis

The gene-level TPM (transcripts per million) values were normalized by the *estimateSizeFactorsForMatrix* function in the R package 'DESeq2' (v.1.22.1)(22) and were log-transformed as $\log_2$ (transcripts per million (TPM)+1). An empirical Bayes moderated *t*-test was used to compare expression of protein coding genes (annotation from gencode.v26.GRCh38) between two groups of interest. Fold change of protein coding genes (annotation from gencode.v26.GRCh38.genes.bed) was then calculated as the ratio of average gene expression values between two groups of interest. GSEA (Gene Set Enrichment Analysis) was performed on the rank metric(fold change)-sorted list of genes by the R package 'fgsea' (v.1.8.0)(23). For scRNA-seq samples, cells with less than 1,500 genes detected or an average expression (in TPM) of housekeeping genes below 5 were excluded from downstream analysis(24).

Pathway gene sets

The Molecular Signatures Database (MSigDB) gene sets v.6.1(25) and IRON_ION_HOMEOSTASIS gene signatures from the Gene Ontology Consortium(26,27) were used to represent broad and well-defined biological processes. Manually curated gene sets were included in the analyses: 1. 'SREBP_TARGET', 'SREBF1_TARGET' (genes predominantly regulated by *SREBF1*) and 'SREBF2_TARGET' (genes predominantly regulated by *SREBF2*)(28). 2. Genes involved in ferroptosis (29). More details are found in Supplementary Table S2.

Gene Set Variation Analysis (GSVA)

Single-sample gene set variance analysis was performed using the *gsva* function (method="gsva", mx.diff=TRUE) from the R package 'GSVA' (v.1.30.0)(30). GSVA implements a non-parametric method of gene set enrichment to generate an enrichment score for each gene set within a sample. GSVA enrichment scores were generated for each gene set using the log-transformed protein coding gene level TPM data.

ChIP-seq data analysis

ChIP-seq reads were aligned to the GRCh38 build of the human genome reference using the Bowtie2(31). Only uniquely mapped reads with mapping quality higher than 30 were retained and duplicated reads were removed. Reads overlapping with ENCODE blacklist regions were also removed. (https://www.encodeproject.org/files/ENCFF419RSJ/@@download/ENCFF419RSJ.bed.gz). Peaks from each MEL-167 replicate were called using the MACS2(32) with a liberal *P* value threshold of 1×10^{-3} for the Irreproducible Discovery Rate (IDR) framework(33,34). 4,441 peaks were identified in the promoter regions with a 0.05 IDR threshold. Read density profiles were generated by the DeepTools program(35) with the *bamCoverage* function (--binSize 20 --normalizeUsing BPM --smoothLength 60 --extendReads 150 --centerReads) and profiles were visualized by the Integrative Genomics Viewer (IGV). Read density profiles in the multiple genes (TSS ± 1 kb) in the gene sets of interest were visualized by *computeMatrix* and *plotHeatmap* functions in the DeepTools program. Genomic annotation of ChIP-seq peaks was performed using the R package 'ChIPseeker' (v.1.18.0)(36). For functional enrichment analysis of genes with SREBF2 binding in promoters, ranked list of genes was generated based on the signal value. Then, GSEA was

performed on the ranked gene list using the *GSEA* function in the R package 'clusterProfiler' (v.3.10.0)(37).

SREBF2 binding motif prediction

Potential *SREBP* binding sites were searched within a promoter region (CpG island) of the *TF* gene using the CIS-BP TF binding tool (Motif model : "PWMs-LogOdds", Threshold : "8"; Database Build 1.02)(2).

Tumor purity quantification from TCGA SKCM

Consensus Purity Estimate (CPE) values from TCGA SKCM tumor samples were obtained from the previous study(38). TCGA SKCM tumor samples with CPE values higher than 0.8 were considered as "high purity" tumor samples.

TCGA SKCM expression subtypes

The five TCGA SKCM expression subtypes ("MITF high", "Intermediate", "MITF low", "Keratin high", and "Immune") were identified by applying BayesNMF with a consensus hierarchical clustering approach(39,40).

Patient survival analysis

Disease-specific survival information of TCGA SKCM patients ('DSS': disease-specific survival event, 'DSS.time': disease-specific survival time) and other clinicopathologic variables were obtained from an integrated TCGA pan-cancer clinical data resource(41). Kaplan-Meier curves (with the log-rank test *P* values) were plotted using the *Surv* function in the R package 'survival' (v.2.43-1). Cox proportional hazards multivariate analysis was performed using the *coxph* function in the R package 'survival' (v.2.43-1). 'Immune'

expression subtypes were excluded from the Cox proportional hazards multivariate analysis because there were only two patients with high tumor purity (CPE>0.8) with *SREBF2* mRNA expression higher than 75th percentile (high *SREBF2* group) or lower than 25th percentile (low *SREBF2* group).

Supplementary References

1. Hong X, Sullivan RJ, Kalinich M, Kwan TT, Giobbie-Hurder A, Pan S*, et al.* Molecular signatures of circulating melanoma cells for monitoring early response to immune checkpoint therapy. Proc Natl Acad Sci U S A **2018**;115(10):2467-72 doi 10.1073/pnas.1719264115.
2. Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P*, et al.* Determination and inference of eukaryotic transcription factor sequence specificity. Cell **2014**;158(6):1431-43 doi 10.1016/j.cell.2014.08.009.
3. Miyamoto DT, Zheng Y, Wittner BS, Lee RJ, Zhu H, Broderick KT*, et al.* RNA-Seq of single prostate CTCs implicates noncanonical Wnt signaling in antiandrogen resistance. Science **2015**;349(6254):1351-6 doi 10.1126/science.aab0917.
4. Aceto N, Bardia A, Miyamoto DT, Donaldson MC, Wittner BS, Spencer JA*, et al.* Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. Cell **2014**;158(5):1110-22 doi 10.1016/j.cell.2014.07.013.
5. Yu M, Bardia A, Aceto N, Bersani F, Madden MW, Donaldson MC*, et al.* Cancer therapy. Ex vivo culture of circulating breast tumor cells for individualized testing of drug susceptibility. Science **2014**;345(6193):216-20 doi 10.1126/science.1253533.
6. Costello M, Pugh TJ, Fennell TJ, Stewart C, Lichtenstein L, Meldrim JC*, et al.* Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation. Nucleic Acids Res **2013**;41(6):e67 doi 10.1093/nar/gks1443.
7. Cibulskis K, McKenna A, Fennell T, Banks E, DePristo M, Getz G. ContEst: estimating cross-contamination of human samples in next-generation

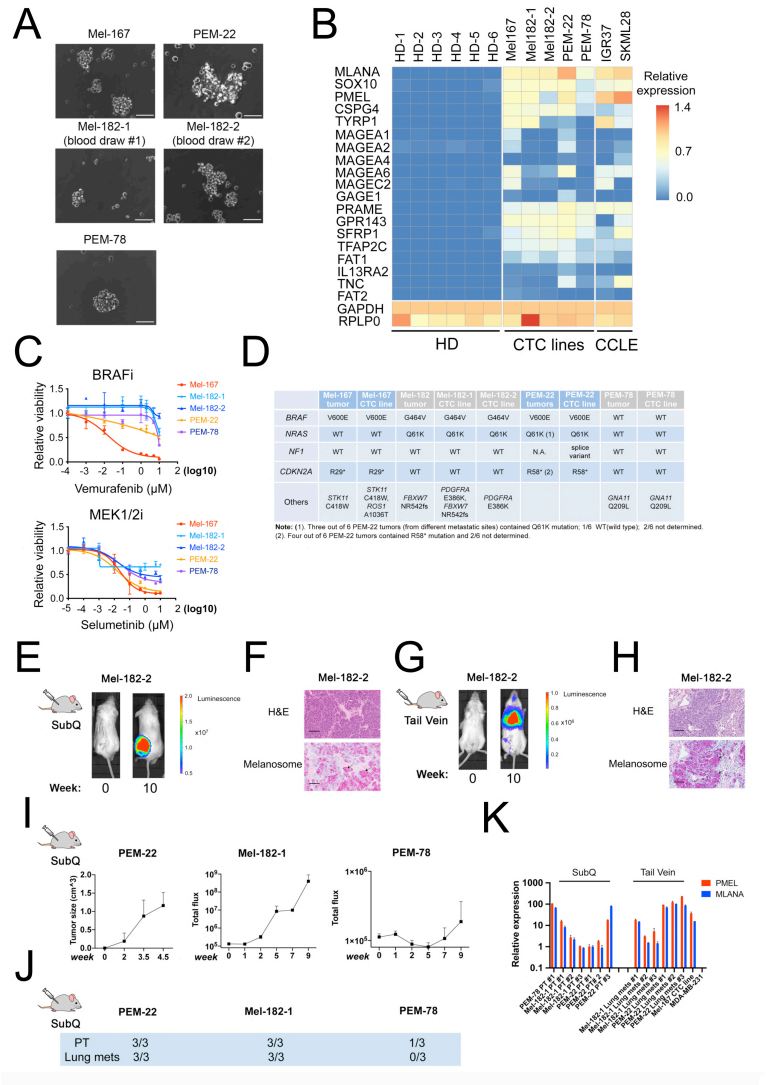
sequencing data. Bioinformatics **2011**;27(18):2601-2 doi 10.1093/bioinformatics/btr446.
8. Chapman MA, Lawrence MS, Keats JJ, Cibulskis K, Sougnez C, Schinzel AC*, et al.* Initial genome sequencing and analysis of multiple myeloma. Nature **2011**;471(7339):467-72 doi 10.1038/nature09837.
9. Berger MF, Lawrence MS, Demichelis F, Drier Y, Cibulskis K, Sivachenko AY*, et al.* The genomic complexity of primary human prostate cancer. Nature **2011**;470(7333):214-20 doi 10.1038/nature09744.
10. Saunders CT, Wong WS, Swamy S, Becq J, Murray LJ, Cheetham RK. Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. Bioinformatics **2012**;28(14):1811-7 doi 10.1093/bioinformatics/bts271.
11. Ramos AH, Lichtenstein L, Gupta M, Lawrence MS, Pugh TJ, Saksena G*, et al.* Oncotator: cancer variant annotation tool. Hum Mutat **2015**;36(4):E2423-9 doi 10.1002/humu.22771.
12. Zheng Z, Liebers M, Zhelyazkova B, Cao Y, Panditi D, Lynch KD*, et al.* Anchored multiplex PCR for targeted next-generation sequencing. Nat Med **2014**;20(12):1479-84 doi 10.1038/nm.3729.
13. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N*, et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res **2019**;47(D1):D941-D7 doi 10.1093/nar/gky1015.
14. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res **2018**;28(11):1747-56 doi 10.1101/gr.239244.118.
15. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S*, et al.* STAR: ultrafast universal RNA-seq aligner. Bioinformatics **2013**;29(1):15-21 doi 10.1093/bioinformatics/bts635.
16. DeLuca DS, Levin JZ, Sivachenko A, Fennell T, Nazaire MD, Williams C*, et al.* RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics **2012**;28(11):1530-2 doi 10.1093/bioinformatics/bts196.
17. Consortium GT, Laboratory DA, Coordinating Center -Analysis Working G, Statistical Methods groups-Analysis Working G, Enhancing Gg, Fund NIHC*, et al.* Genetic effects on gene expression across human tissues. Nature **2017**;550(7675):204-13 doi 10.1038/nature24277.
18. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W*, et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res **2015**;43(7):e47 doi 10.1093/nar/gkv007.
19. Yip SH, Wang P, Kocher JA, Sham PC, Wang J. Linnorm: improved statistical analysis for single cell RNA-seq expression data. Nucleic Acids Res **2017**;45(22):e179 doi 10.1093/nar/gkx828.
20. Risso D, Perraudeau F, Gribkova S, Dudoit S, Vert JP. A general and flexible method for signal extraction from single-cell RNA-seq data. Nat Commun **2018**;9(1):284 doi 10.1038/s41467-017-02554-5.
21. Van den Berge K, Perraudeau F, Soneson C, Love MI, Risso D, Vert JP*, et al.* Observation weights unlock bulk RNA-seq tools for zero inflation and single-cell applications. Genome Biol **2018**;19(1):24 doi 10.1186/s13059-018-1406-4.

22. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol **2014**;15(12):550 doi 10.1186/s13059-014-0550-8.
23. Sergushichev A. An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. BioRxiv **2016** doi https://doi.org/10.1101/060012.
24. Sade-Feldman M, Yizhak K, Bjorgaard SL, Ray JP, de Boer CG, Jenkins RW*, et al.* Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma. Cell **2018**;175(4):998-1013 e20 doi 10.1016/j.cell.2018.10.038.
25. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst **2015**;1(6):417-25 doi 10.1016/j.cels.2015.12.004.
26. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM*, et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet **2000**;25(1):25-9 doi 10.1038/75556.
27. The Gene Ontology C. The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res **2019**;47(D1):D330-D8 doi 10.1093/nar/gky1055.
28. Horton JD, Shah NA, Warrington JA, Anderson NN, Park SW, Brown MS*, et al.* Combined analysis of oligonucleotide microarray data from transgenic and knockout mice identifies direct SREBP target genes. Proc Natl Acad Sci U S A **2003**;100(21):12027-32 doi 10.1073/pnas.1534923100.
29. Stockwell BR, Friedmann Angeli JP, Bayir H, Bush AI, Conrad M, Dixon SJ*, et al.* Ferroptosis: A Regulated Cell Death Nexus Linking Metabolism, Redox Biology, and Disease. Cell **2017**;171(2):273-85 doi 10.1016/j.cell.2017.09.021.
30. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics **2013**;14:7 doi 10.1186/1471-2105-14-7.
31. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods **2012**;9(4):357-9 doi 10.1038/nmeth.1923.
32. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE*, et al.* Model-based analysis of ChIP-Seq (MACS). Genome Biol **2008**;9(9):R137 doi 10.1186/gb-2008-9-9-r137.
33. Li Q, Brown JB, Huang H, Bickel PJ. Measuring reproducibility of high-throughput experiments. The Annals of Applied Statistics **2011**;5(3):1752-79 doi DOI: 10.1214/11-AOAS466.
34. Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S*, et al.* ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res **2012**;22(9):1813-31 doi 10.1101/gr.136184.111.
35. Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T. deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res **2014**;42(Web Server issue):W187-91 doi 10.1093/nar/gku365.
36. Yu G, Wang LG, He QY. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics **2015**;31(14):2382-3 doi 10.1093/bioinformatics/btv145.

37. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS **2012**;16(5):284-7 doi 10.1089/omi.2011.0118.
38. Aran D, Sirota M, Butte AJ. Systematic pan-cancer analysis of tumour purity. Nat Commun **2015**;6:8971 doi 10.1038/ncomms9971.
39. Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD*, et al.* Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell **2018**;174(4):1033 doi 10.1016/j.cell.2018.07.036.
40. Kim J, Kwiatkowski D, McConkey DJ, Meeks JJ, Freeman SS, Bellmunt J*, et al.* The Cancer Genome Atlas Expression Subtypes Stratify Response to Checkpoint Inhibition in Advanced Urothelial Cancer and Identify a Subset of Patients with High Survival Probability. Eur Urol **2019**;75(6):961-4 doi 10.1016/j.eururo.2019.02.017.
41. Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD*, et al.* An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. Cell **2018**;173(2):400-16 e11 doi 10.1016/j.cell.2018.02.052.

Supplementary Figures and Legends

Figure S1



	Mel-167 tumor	Mel-167 CTC line	Mel-182 tumor	Mel-182-1 CTC line	Mel-182-2 CTC line	PEM-22 tumors	PEM-22 CTC line	PEM-78 tumor	PEM-78 CTC line
BRAF	V600E	V600E	G464V	G464V	G464V	V600E	V600E	WT	WT
NRAS	WT	WT	Q61K	Q61K	Q61K	Q61K (1)	Q61K	WT	WT
NF1	WT	WT	WT	WT	WT	N.A.	splice variant	WT	WT
CDKN2A	R29*	R29*	WT	WT	WT	R58* (2)	R58*	WT	WT
Others	*STK11* C418W	*STK11* C418W, *ROS1* A1036T	*FBXW7* NR542fs	*PDGFRA* E386K, *FBXW7* NR542fs	*PDGFRA* E386K			*GNA11* Q209L	*GNA11* Q209L

Note: (1). Three out of 6 PEM-22 tumors (from different metastatic sites) contained Q61K mutation; 1/6 WT(wild type); 2/6 not determined.
(2). Four out of 6 PEM-22 tumors contained R58* mutation and 2/6 not determined.

Figure S1. Establishment and characterization of melanoma CTC lines.

(**A**) Representative bright field images of 5 cultured CTC lines derived from four patients with metastatic melanoma (Mel-167, PEM-22, Mel182, PEM-78). Mel-182-1 and Mel-182-2 are independent CTC lines derived from two independent blood draws from a patient treated at times when he was receiving different therapeutic regimens. Scale bar, 100 μm.

(**B**) Heatmap representing a 19-gene panel of transcripts that are characteristic of melanoma lineage CTCs using a droplet digital PCR assay (1). The expression of the 19

markers in early cultures of melanoma CTCs (<1000 cells) is shown. The ATCC (CCLE) melanoma cell lines IGR37 and SK-ML-28 were used as positive controls and white blood cell cDNA from healthy donors (HD) processed through the CTC-iChip was used as negative control. Each column represents a sample and each row represents an individual melanoma marker or the housekeeping genes, GAPDH and RPLP0. Data are normalized to GAPDH and represent three independent experimental repeats.

(**C**) Sensitivity of the five melanoma CTC lines to the BRAF inhibitor vemurafenib (upper panel) and the MEK1/2 inhibitor selumetinib (lower panel). Y-axis, relative cell viability; X-axis, drug concentrations in μM, log(10) scale.

(**D**) Sequencing of melanoma CTC lines, compared with matched biopsy samples of metastatic tumors, using whole exome sequencing (Mel-167/Mel-182-1/Mel-182-2) or cancer panel sequencing using the Anchored Multi-plexed PCR method (PEM-22 and PEM-78). For PEM-22 two additional mutations (NRAS Q61K and CDKN2A R58*) were detected from RNA-seq data (for both tumor and CTC line). For Mel-182, two independent CTC lines were derived at different times during the patient's treatment, both of which carry the PDGFRA E386K mutation, which was not evident in the earlier metastatic tumor biopsy. For PEM-22, the single CTC line carries the NRAS and CDKN2A mutations, both of which were present in subsets of 6 independent metastatic biopsies.

(**E**) Representative image of GFP-luciferase-tagged Mel-182-2 CTC-derived subcutaneous tumor in NSG mice (100,000 cells per injection). Tumors were monitored over time by live imaging (IVIS) and images are shown post-injection and at week 10. N = 3 mice tested.

(**F**) H&E (top) and melanosome marker (bottom) staining (purple, Roche Cat# 790-4677) of the CTC-derived primary tumor shown in S1E (tumor cells marked by black arrows). Scale bar, 100μm.

(G) Representative image of NSG mouse inoculated by tail vein with GFP-luciferase tagged Mel182-2 CTCs (500,000 cells per injection), shown before injection and at week 10, when most tumor cells are lodged in the lungs. N = 3 mice tested.

(**H**) H&E (top) and melanosome marker (bottom) staining of lung metastases following tail vein injection shown in Figure S1G (tumor cells marked by black arrows). Scale bar, 100μm.

(**I**) Time course of tumor formation following subcutaneous inoculation of 200,000 cultured PEM-22, Mel-182-1 and PEM-78 CTCs into immunosuppressed NSG mice (n = 3 for each CTC line). Tumors were monitored over time by IVIS luciferin imaging (Mel-182-1/PEM-78) or by size (PEM-22).

(**J**) Fraction of injected NSG mice developing primary tumors (PT) following subcutaneous inoculation (200,000 cell inoculum) or developing lung metastases (lung mets) following tail vein inoculation (500,000 cell inoculum; IVIS quantitation) for three independent CTC cultures, PEM-22, Mel-182-1 and PEM-78 (n = 3 for each CTC line).

(**K**) q-PCR quantitation of melanosome marker transcripts SILV (PMEL) and MLANA in primary tumors and in lung metastases from NSG mice injected with PEM-22, Mel-182-1 or PEM-78 CTC lines. The breast cancer cell line MDA-MB-231 cDNA was used as a negative control. Actin was used as control and data were internally normalized to PEM-22 PT #1.

Figure S2

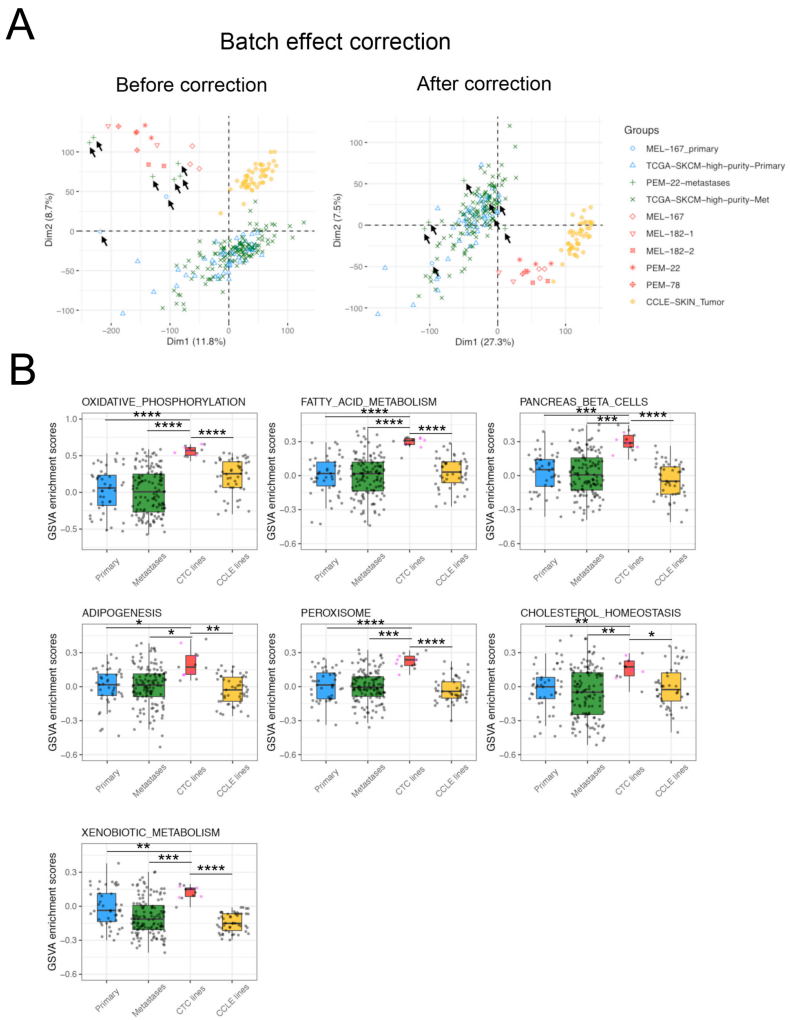


Figure S2. Pathway enrichment analyses comparing TCGA high-purity primary tumors & metastases, melanoma CTC lines, and CCLE melanoma lines.

(**A**) Global gene expression profiles of samples used in this study, TCGA, and CCLE datasets before and after batch effect correction. PCA analysis was performed to visualize the global gene expression profiles of the above samples. The x-axis is the first principal component and the y-axis is the second principal component. Samples are color-coded by sample type (blue: primary tumors, green: metastases, red: CTC lines, yellow: CCLE lines).

(**B**) Boxplots showing significantly upregulated pathways specific to CTC cell lines by comparing the differences in GSVA enrichment scores between melanoma CTC cell line samples (n=13, including 5 distinct lines colored in pink and 8 sample repeats in grey), versus high-purity primary melanoma (n=45) and high-purity metastatic melanomas (n=129) from TCGA and standard melanoma cell lines (CCLE, n=49). The mean GSVA enrichment scores were used for replicates of each CTC line. The differences in mean GSVA enrichment scores > 0.1 and FDR-adjusted *P* values < 0.05 were used as cutoffs for all significantly upregulated pathways in CTC lines. Y-axis, GSVA enrichment scores. ****$P < 0.0001$; ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$. A detailed list of GSVA enrichment scores for all pathways across all samples is found in Supplementary Table S3.

Figure S3

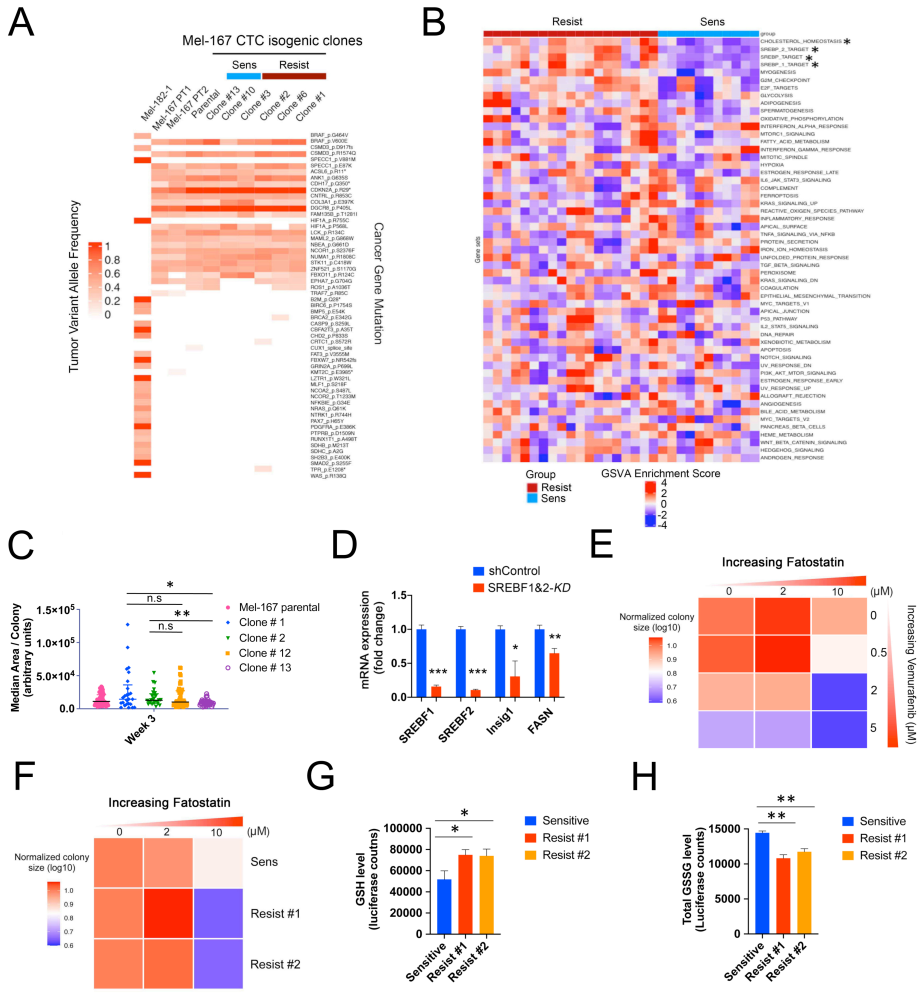


Figure S3. Molecular characterization of Mel-167 CTC subpopulations and effect of SREBP activity on BRAF sensitivity.

(**A**) Whole-exome sequencing (WES) of Mel-167 isogenic clones, with cancer gene mutations (rows) represented as oncoplot heatmap. Shown are mutations in two Mel-167 biopsied tumor specimens PT1 and PT2, in the bulk Mel-167 parental CTC line, and in 6 independent single CTC-derived clones #1, #2, #3, #6, #10 and #13. Clones #10 and #13 are sensitive to BRAFi, while clones #1, #2, #3, #6 are resistant (see Figures 2A-C). An unrelated CTC line (with distinct mutational profile), Mel-182-1, is shown as control. The following mutation types were included in the analysis: " Missense_Mutation",

"Nonsense_Mutation", " Splice_Site", "Frame_Shift_Ins" and "Frame_Shift_Del". The color scale of the heatmap is based on the calculated tumor variant allele frequency of each cancer gene. A detailed description of the analysis is found in Supplementary Methods.

(**B**) Heatmap of GSVA enrichment scores of all pathways showing a difference between BRAF inhibitor-sensitive (blue group in upper heading) and resistant isogenic CTC isogenic lines (red group in upper heading). Each row represents a pathway and each column represents a sample. Statistical significance was assessed by two-sided Welch's t-test. The differentially expressed pathways between the two groups were defined as the gene sets with FDR-adjusted P values < 0.05 and absolute mean difference in GSVA scores between two groups > 0.25 and are marked with asterisks. ***P*< 0.01; **P*< 0.05.

(**C**) Quantitation of soft agar colony size of selected resistant (#1, #2) and sensitive (#12, #13) isogenic CTC clones after week 3 of growth in soft agar. Y axis, calculated area (arbitrary units) of individual colonies in each isogenic clone as well as Mel-167 parental line. Statistical Significance was assessed using Mann-Whitney test.

(**D**) Real time q-PCR quantitation of *SREBF1/2* mRNAs and their downstream target mRNAs, *Insig1* and *FASN,* in a representative vemurafenib-resistant Mel-167 line (Resist #1), demonstrating effective knockdown following transfection with siRNAs against *SREBF1* and *SREBF2 (SREBF1&2-KD)*. Y-axis: relative fold change of siRNA treated CTCs as compared to control siRNA. *Actin was* used for internal normalization. Data was obtained from three independent biological repeats. Statistical significance was assessed by two-sided Welch's *t*-test. $P = 0.0007$ for *SREBF1*; $P = 0.0003$ for *SREBF2*; $P = 0.0288$ for *Insig1*; $P = 0.0029$ for *FASN*. ****P* < 0.001; ***P*< 0.01; **P*< 0.05.

(**E**) Heatmap representing the size of soft agar colonies following treatment of Mel-167 CTCs with increasing concentrations of the BRAFi Vemurafenib and SREBP inhibitor

Fatostatin, , showing cooperative cell toxicity. The drug effect on colony number is shown in Figure 2F.

(**F**) Heatmap representing the size of soft agar colonies by the parental (sensitive) Mel-167 CTs, compared with two Vemurafenib-resitant derivative clones (#1, #2), following treatment with increasing concentrations of the SREBP inhibitor Fatostatin. Statistical significance was assessed by two-sided T test with Welch's correction. P = 0.0035, comparing resistant clone #1 to sensitive line and P = 0.0016 comparing resistant clone #2 to sensitive line. The Vemurafenib-resistant CTCs with increased endogenous SREBF1 and SREBF2 expression show increased sensitivity to Fatostatin. The drug effect on colony number is shown in Figure 2H.

(**G**) Total GSH level measured in two Mel-167-CTC clonal lines with acquired resistance to Vemurafenib compared with the control sensitive cells (refer to Figure 2I). Statistical significance was assessed by two-sided T test with Welch's correction. P = 0.0202, comparing Resist #1 to Sensitive and P = 0.0226 comparing Resist #2 to Sensitive line. * $P < 0.05$.

(**H**) Total GSSG level measured in Vemurafenib-resistant compared with the control sensitive cells (refer to Figure 2I). Statistical significance was assessed by two-sided T test with Welch's correction. P = 0.0016, comparing Resist #1 to Sensitive and P = 0.0020 comparing Resist #2 to Sensitive line. ** $P < 0.01$.

Figure S4

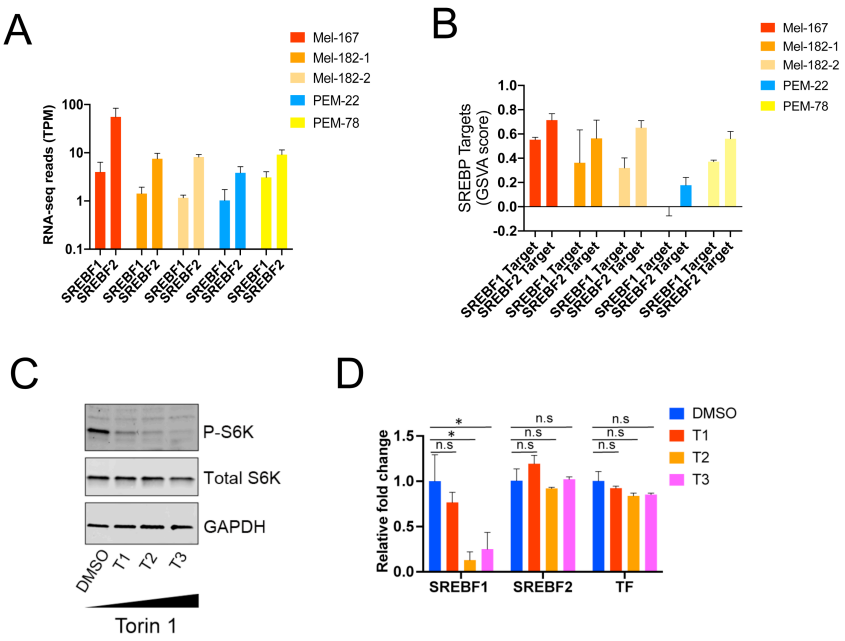


Figure S4. Differential effects of SREBF1 and SREBF2 in melanoma CTCs.

(**A**) Bar graph showing *SREBF1* and *SREBF2* mRNA expression levels in all five CTC lines Y-axis: RNA-seq reads in log10 scale (TPM, transcripts per million). No. of RNA-seq samples: Mel-167, n = 3; Mel-182-1, n = 2; Mel-182-2, n = 2; PEM-22, n = 3; PEM-78, n = 3.

(**B**) Bar graph showing GSVA scores of SREBF1- and SREBF2-targets in all five CTC lines as shown in (A). The definition of SREBF1- and SREBF2-targets can be found in Supplementary Table S2.

(**C**) Effective suppression of mTOR signaling by increasing Torin1 concentrations concentrations (T1 = 0.0156 µM; T2 = 0.0313 µM; T3= 0.0625 µM), shown by Western blot analysis of phosphorylated S6K and total S6K protein in Mel-167 CTCs. GAPDH was used as loading control.

(**D**) q-PCR quantification of *SREBF1*, *SREBF2* and *TF* mRNA in Mel-167 CTCs treated with the mTOR inhibitor Torin 1 at increasing concentrations (T1 = 0.0156 μM; T2 = 0.0313 μM; T3= 0.0625 μM), demonstrating effective downregulation of *SREBF1* but not *SREBF2*. The SREBP target gene *Transferrin (TF)* is not affected by mTOR inhibition, consistent with its regulation through *SREBF2*, rather than *SREBF1*. Statistical significance was assessed by two-sided Welch's t-test. * $P < 0.05$; n.s, not significant.

Figure S5

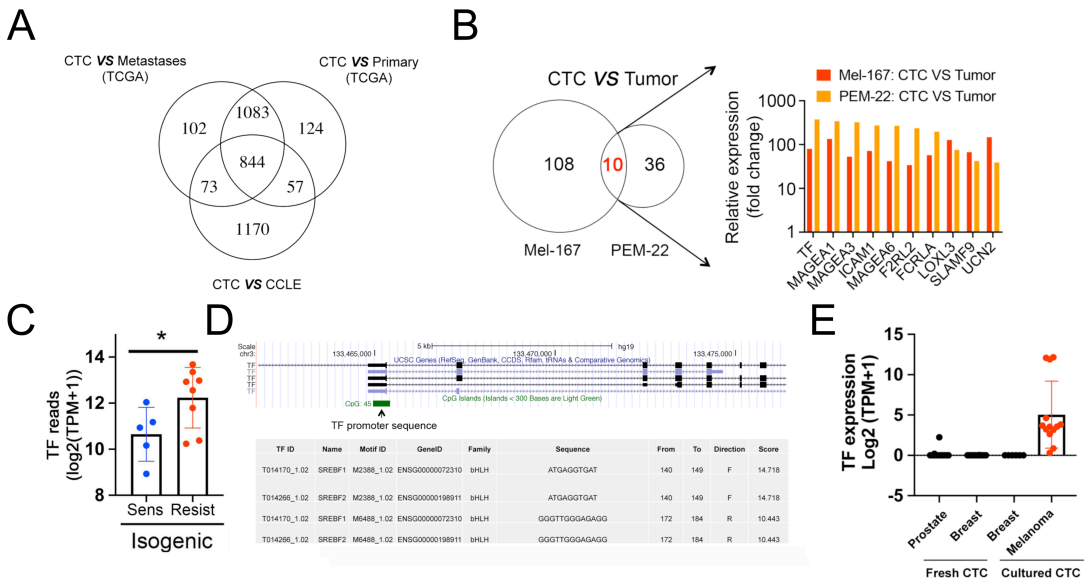


Figure S5. Direct transcriptional regulation of *TF* by SREBF2.

(**A**) Venn diagram of genes that are commonly up-regulated (fold change > 2 in CTCs) in three pairwise comparisons: (1) all 5 CTC lines *vs.* TCGA-derived high-purity primary tumor; (2) all 5 CTC lines *vs.* TCGA-derived high-purity metastases; (3) all 5 CTC lines *vs.* CCLE melanoma lines. This resulted in a total of 844 genes, which is also shown in Figure 3D (No. of genes within the blue circle: 586 + 258 = 844). The top 10 genes (based on

mean fold change) that are commonly up-regulated in these three pairwise comparisons are listed in Figure 3D. A detailed list of differentially expressed genes can be found in Supplementary Table S1.

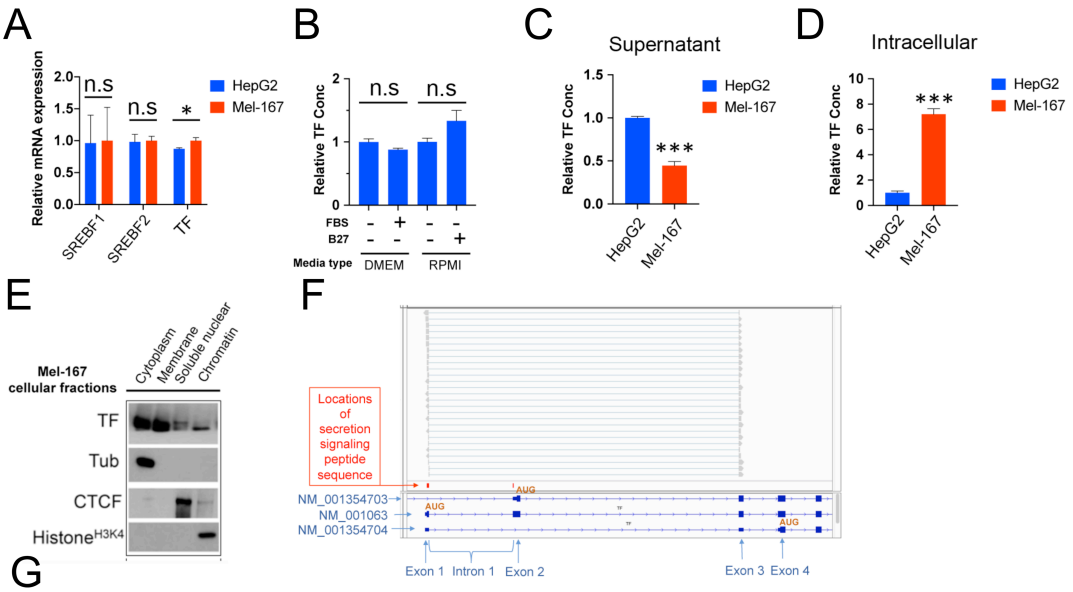
(**B**) Melanoma CTC-enriched transcripts. Left: Venn diagram showing the overlap between two pair-wise comparisons of cultured CTCs and their available matched tumor samples. For patient Mel-167, the cultured CTCs are compared with the primary tumor specimen which was available; for patient PEM-22, the cultured CTCs were compared with biopsies of six independent metastatic lesions (primary tumor was not available, but multiple on-treatment biopsies of superficial lesions were obtained during clinical course); Right: Graphical representation of the top 10 differentially expressed genes, common to both pair-wise comparison (X axis), with fold changes (Y axis). Mean expression levels for the top 10 upregulated genes in CTC lines (reads in TPM) are : *TF* = 114; *MAGEA1* = 86; *MAGEA3* = 170; *ICAM1* = 1428; *MAGEA6* = 159; *F2RL2* = 46; *FCRLA* = 519; *LOXL3* = 181; *SLAMF9* = 19; *UCN2* = 34.

(**C**) Increased expression of *TF* mRNA in untreated Mel-167 isogenic lines with intrinsic resistance to BRAFi, compared with those that are intrinsically sensitive to the drug. Bar graph showing *TF* RNA-seq reads (log2(TPM+1)) of 8 resist (red, #1 to #8) versus 5 sensitive isogenic lines (blue, #9 to #13) as shown in Figure 2A. Statistical significance was assessed by two-sided Welch's *t*-test. * $P < 0.05$.

(**D**) The *TF* promoter region (CpG island) spanning 471 bp as shown in the UCSC genome browser ("YourSeq" track in the top panel) used to identify putative *SREBP* binding sites using the CIS-BP (Catalog of Inferred Sequence Binding Preferences) TF binding tool (http://cisbp.ccbr.utoronto.ca/TFTools.php)(2). The binding motif, position of binding sites and binding scores are listed (bottom panel).

(**E**) Boxplot showing the absence of *TF* mRNA expression (log_2(TPM+1)) in freshly isolated prostate CTCs (3), freshly isolated breast CTCs (4), and breast CTC cell lines (5) compared with melanoma CTC lines.

Figure S6



	Total # of reads that span intron 1	# of reads that align to exons 1 & 3 but not to exon 2	% of reads that align to exons 1 & 3 but not to exon 2	95% CI lower bound	95% CI upper bound
HepG2	52,635	471	0.9%	0.8%	1.0%
Mel-167	1,975	27	1.4%	0.9%	2.0%

Figure S6. Characterization of secreted and non-secreted *TF* variants in Mel-167 CTCs.

(**A**) q-PCR quantitation of *SREBF1, SREBF2* and *TF* mRNA expression in liver cancer HepG2 and melanoma Mel-167 CTCs. Actin was used as an internal control and data normalized to Mel-167 CTCs. Statistical significance was assessed by two-sided Welch's *t*-test. * $P < 0.05$; n.s, not significant.

(**B**) Relative quantification of TF protein (ELISA assay) in culture media for HepG2 (DMEM) in the presence or absence of Fetal Bovine Serum (FBS; 10%), and in culture

medium for Mel-167 CTCs (RMPI) in the presence or absence of B27 (2%). Under both culture conditions, the very low amounts of TF present in either FBS or B27 supplements do not significantly alter total TF levels. n.s, not signficant. Statistical significance was assessed by two-sided Welch's *t*-test. n.s, not significant.

(**C-D**) Relative quantification of TF protein by ELISA of culture supernatant from HepG2 or Mel-167 CTCs over a period of 48 hours (Figure panel C) and intracellular TF protein levels assayed by direct cell lysis (Figure panel D). Data shown as normalized to HepG2 cells The absolute quantification of TF in the respective cell types is describedas follows: as noted above in Figure S6B, we first tested the contribution of the culture medium: the B27 "stem-cell medium" used to culture CTC lines has minimal levels of TF, and the total amount of TF in supernatant at 48 hrs was not significantly different when cells were maintained in the presence or absence of B27. Similarly the FCS used to culture HepG2 cells has minimal levels of TF. We therefore tested cells under full growth-supporting culture conditions. The total amount of TF protein secreted by Mel-167 CTCs is 45% of that produced by HepG2 cells (P = 0.0008, two-sided T test with Welch's correction), whereas the total intracellular TF concentration in Mel 167 CTCs is 5.2 fold higher (P=0.0005, two-sided T test with Welch's correction). Intracellular and secreted TF levels for any given cell line are not strictly comparable, since the former reflects steady state baseline expression, while the second reflects increasing accumulation over 48 hrs of incubation. Nonetheless, for Mel 167 CTCs, the total amount of intracellular TF was calculated as 19.22 ng (per milliliter lysate from 1 million cells), compared with 3.10 ng in HepG2 cells; the total amount of TF in culture supernatant of Mel-167 CTCs was 1.52 ng (per 1 milliliter media from 1 million cells over 48 hr collection), compared to 3.41ng in the supernatant of HepG2. Statistical significance was assessed by two-sided Welch's *t*-test. n.s, not significant. *** $P < 0.001$.

(**E**) Western blot analysis of TF subcellular localiation in Mel-167 CTCs. The four independent fractions (cytoplasm, membrane, soluble nuclear and chromatin) were purified using the Subcellular Protein Fractionatiaon kit (ThermoFisher Scientific Cat #78840). Antibodies against α-Tubulin (a marker for cytoplasmic protein), CTCF (a marker for soluble nuclear fraction) and Histone H3K4 (a marker for chromatin fraction) were used as controls in the analysis.

(**F**) Schematic representation showing the genomic structure of the first 4 exons of the *TF* gene, including the location of the secretion signaling peptide sequence (red bars) spanning the 3' end of exon 1 and exon 2. The location of the primary and of two alternative translational starts (AUG) are noted, with the AUG in exon 4 producing an in-frame iron-binding protein lacking the secretion signaling peptide sequence. The IGV display shows 27 mRNA reads from Mel-167 RNA-seq that align to exons 1 and 3 but not exon 2. This implies that those reads cannot arise from transcripts with a full secretion signaling peptide sequence because the transcripts from which they arise lack exon 2. The splicing of exon 1 to exon 3 results in a transcript with an open reading frame starting in exon 4 (NM-001354704), which is in-frame with the canonical splice variant NM_001063, but lacks the secretion signal. In addition, NM-001354703 encodes a transcript lacking exon 1, and hence also lacking the signal sequence.

(**G**) Percent of reads spanning TF intron 1 that cannot arise from transcripts with a full secretion signaling peptide sequence because the transcripts from which they arise lack exon 2 in either HepG2 or Mel-167 RNA-seq.

Figure S7

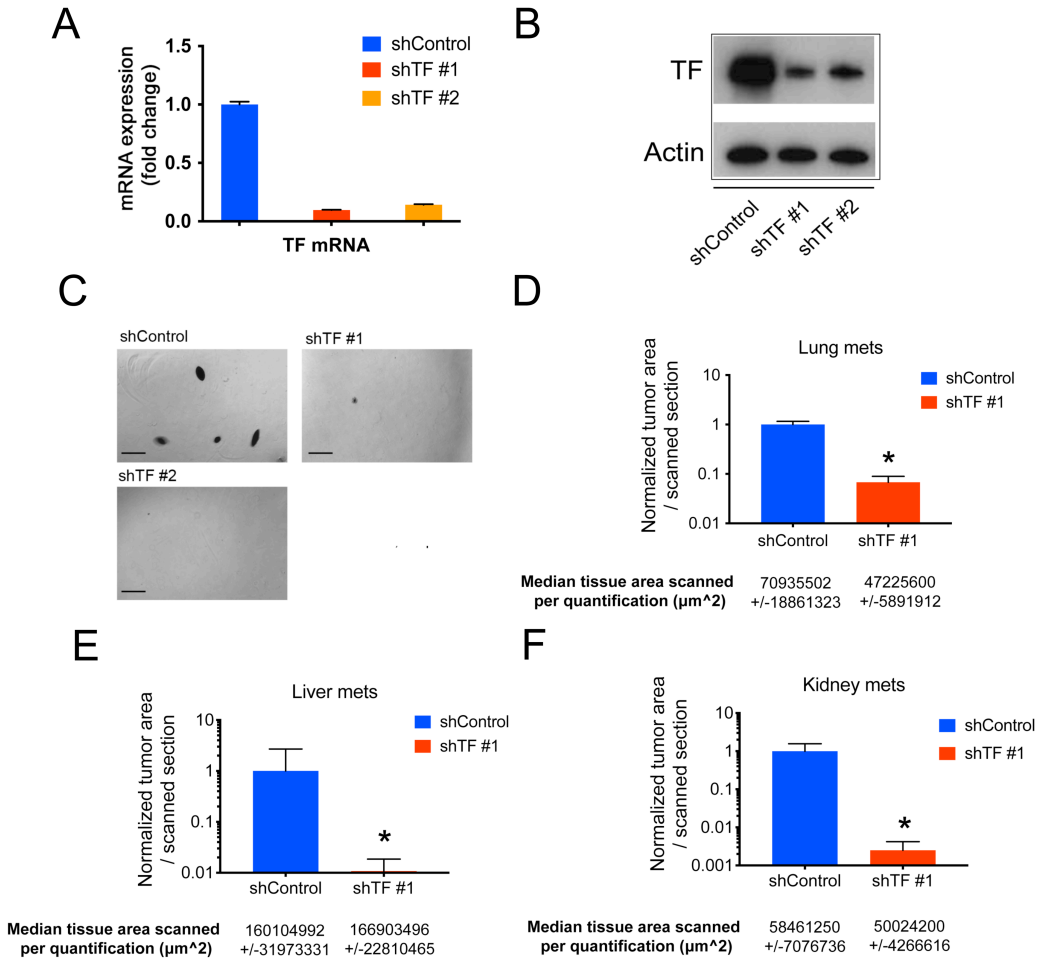


Figure S7. Functional consequences of TF depletion

(A) Real time q-PCR analysis of *TF* mRNA in *Mel-167* CTCs, following transfection with two independent shRNA constructs targeting *TF,* compared to shControl. Y-axis: relative fold change of sh*TF*, with *Actin* internal normalization. Data was obtained from two independent repeats.

(B) Western blot analysis of TF protein expression in Mel-167 CTCs, following knockdown using either of two shRNA constructs (Actin loading control).

(C) Representative images of soft agar colonies from Mel-167 CTCs transfected with either of two independent shRNAs against *TF* (shControl*).* Colony quantification is shown in Figure 4C. Scale bar, 500 μm.

(**D-F**) Bar graph quantification of metastatic burden in mice following tail-vein injection of Mel-167 CTCs expressing shTF#1 versus shControl, as calculated by scanning and automated image analysis of lung (D), liver (E) and kidney (F) tissues. Quantification of total tumor cell area within each scanned IHC section (stained with melanosome marker, Roche Cat#790-4677; VIS software, Visiopharm Inc, see detailed analysis in Supplementary Information). Y axis: fraction of melanoma cells per total tissue area within each scanned tissue section. The median tissue area subjected to automated scanning with standard deviations (in μm^2) is indicated below each graph. Number of slides per group= 4. Two-sided Welch's *t*-test was used to assess the statistical significance. * $P < 0.05$.

Figure S8

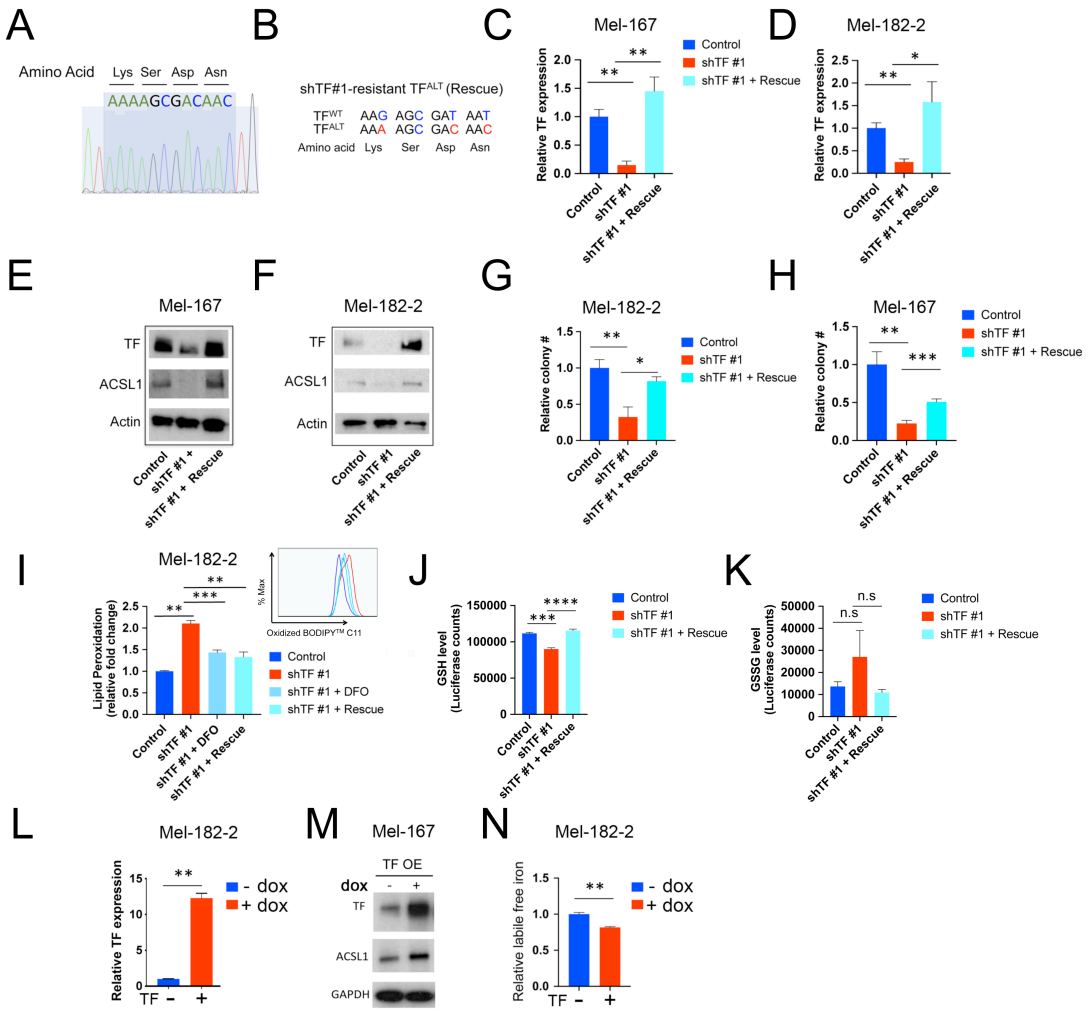


Figure S8. Molecular consequences of *TF* deletion in Mel-167 and Mel-182-2 CTCs.

(**A-B**) Sequence (**A**) and nucleotide tracing (**B**) of the synthetic *TF* ^{ALT} construct, in which three 3rd-position nucleotides in the coding region of *TF* cDNA (red) have been altered within the targeting sequence of shTF#1, thereby rendering the cDNA resistant to shRNA knockdown. Infection of *TF* ^{ALT} together with shTF#1 in Mel-167 and Mel-182-2 CTC lines enables rescue of the TF-KD phenotype, confirming specificity.

(**C**) q-PCR quantitation of *TF* mRNA expression in Mel-167 CTCs infected with control vector, shTF #1, or shTF#1 plus the TF^{ALT} rescue construct (sequential antibiotic

selections). Statistical significance was assessed by two-sided Welch's *t*-test. n.s, not significant. ** $P < 0.01$.

(**D**) q-PCR quantitation of *TF* mRNA expression in Mel-182-2 CTCs infected with control vector, shTF #1, or shTF#1 plus the TF^{ALT} rescue construct. Statistical significance was assessed by two-sided Welch's *t*-test. n.s, not significant. ** $P < 0.01$; * $P < 0.05$.

(**E**) Western blot analysis of TF protein expression in Mel-167 CTCs infected with control vector, shTF #1, or shTF#1 plus the TF^{ALT} construct, demonstrating effective rescue of comparable TF expression. ACSL1 was used as a readout of SREBP activity, again showing rescue of the knockdown, and Actin was used as loading control.

(**F**) Western blot analysis of TF protein expression in Mel-182-2 CTCs infected with control vector, shTF #1, or shTF#1 plus the TF^{ALT} rescue construct. ACSL1 was used as a readout of SREBP activity and Actin was used as loading control.

(**G**) Soft agar colony numbers of Mel-182-2 infected with shControl, shTF#1 alone, or shTF#1 with TF^{ALT} rescue. Y-axis, relative colony number normalized to control. Statistical significance was assessed by two-sided Welch's t-test. ** $P < 0.01$; * $P < 0.05$.

(**H**) Soft agar colony number quantification of Mel-167 CTCs infected with shControl, shTF#1 alone, or shTF#1 with TF^{ALT} rescue. Y-axis, relative colony number normalized to control. Statistical significance was assessed by two-sided Welch's t-test. *** $P < 0.001$; ** $P < 0.01$.

(**I**) Quantitation of lipid peroxidation levels by flow cytometry (BODIPY™ 581/591 C11 molecular sensor) in Mel-182-2 CTCs, following infection with shControl, shTF#1, shTF#1 plus 50 μM iron chelator deferoxamine (DFO, incubated for 12 hours before flow cytometry assay), or shTF#1 with TF^{ALT} rescue, with the histogram shown on the top right. The fraction of cells positive for lipid peroxidation was calculated and data are

normalized to shControl. Statistical significance was assessed by two-sided Welch's t-test. ** $P < 0.01$; *** $P < 0.001$. Data for Mel-167 cells are shown in Figure 5C.

(**J**) Total GSH level measured in Mel-167 CTCs infected with control vector, shTF #1, or shTF#1 plus the TF^{ALT} rescue construct (refer to Figure 5F). Statistical significance was assessed by two-sided T test with Welch's correction. $P = 0.0001$, comparing Control to shTF #1 and $P < 0.0001$ comparing shTF#1 to shTF#1 + Rescue. *** $P < 0.001$; **** $P < 0.0001$.

(**K**) Total GSSG level quantification in Mel-167 CTCs infected with control vector, shTF #1, or shTF#1 plus the TF^{ALT} rescue construct (refer to Figure 5F). Statistical significance was assessed by two-sided T test with Welch's correction. n.s, not significant.

(**L**) q-PCR quantitation of *TF* mRNA expression in Mel-182-2 CTCs infected with a TF cDNA that is doxycline-inducible (-dox VS +dox, 200ng/ml, induced for 48 hours). Statistical significance was assessed by two-sided Welch's *t*-test. n.s, not significant. ** $P < 0.01$.

(**M**) Western blot quantification of dox-inducible TF overexpression (200ng/ml dox induced for 48 hours) in Mel-167 CTCs cultured using 2%B7. ACSL1 was used as a marker for SREBP activity and GAPDH was used as loading control.

(**N**) Quantitation of intracellular labile free iron by flow cytometry using a fluorescent reporter (Goryo chemical Cat# GC903-01) in Mel-182-2 CTC lines. The normalized geometric means of fluorescence intensity were calculated for CTCs expressing a TF cDNA that is doxycline-inducible (-dox VS +dox, 200ng/ml, induced for 48 hours). Statistical significance was assessed by two-sided Welch's t-test. ** $P < 0.01$. Data for Mel-167 CTCs is shown in Figure 5I.

Figure S9

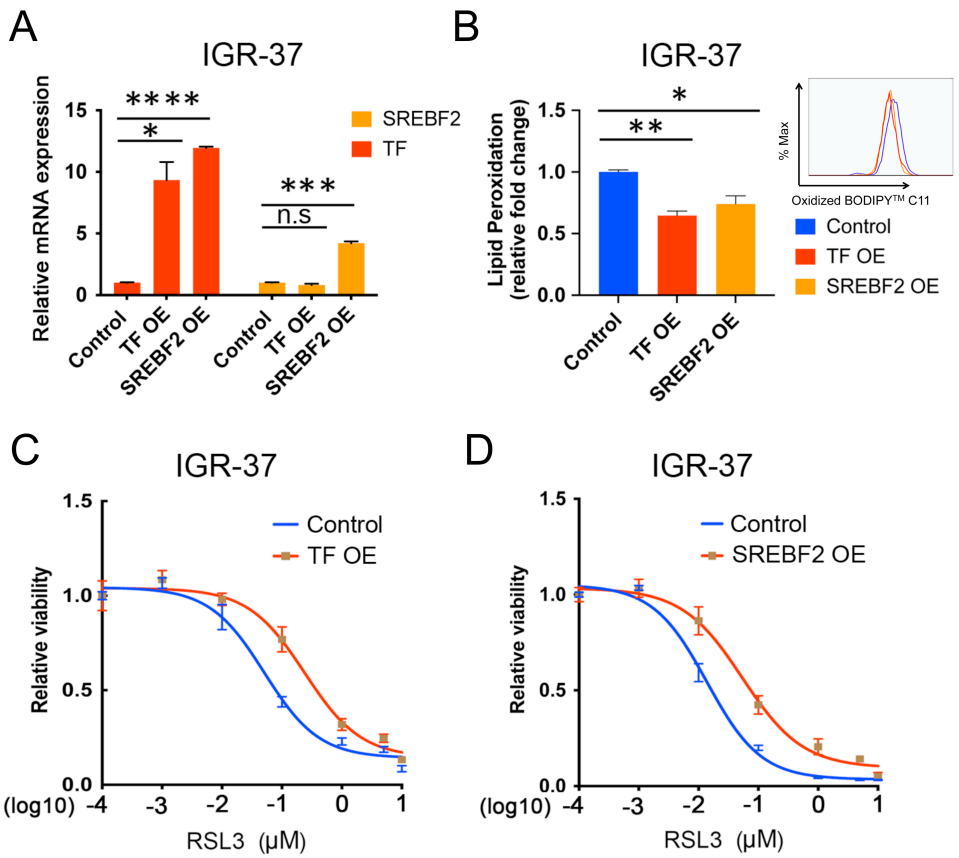


Figure S9. Overexpression of TF or SREBF2 in IGR-37 melanoma cell line.

(**A**) q-PCR quantitation of *SREBF2* and *TF* mRNA expression in IGR-37 melanoma cells infected with control vector, *TF* cDNA or the mature form of *SREBF2* cDNA, or. Statistical significance was assessed by two-sided Welch's *t*-test. n.s, not significant. **** $P < 0.0001$; *** $P < 0.001$; * $P < 0.05$.

(**B**) Lipid peroxidation was quantified by flow cytometry (using the oxidation molecular sensor BODIPY™ 581/591 C11) following ectopic overexpression (OE) of *TF* or *SREBF2* in the CCLE melanoma cell line IGR-37, which does not express endogenous *TF,* with the histogram shown on the top right. The fraction of cells positive for lipid peroxidation was calculated and the data normalized to control vector. Y axis represents fold change. Statistical significance was assessed by two-sided Welch's t-test: $P = 0.0011$ for TF OE

compared to Control; P = 0.0162 for SREBF2 OE compared to Control. * P < 0.05; ** P < 0.01.

(**C**) Resistance to the ferroptosis inducer RSL3 conferred by overexpression of *TF* in IGR-37 melanoma cells. Y axis represents relative cell viability. X axis represents drug concentrations (μM) in log10 scale. Statistical significance is assessed by two-sided T test with unequal variance and P values are generated when comparing cell viabilities between control and TF OE groups treated with RSL3 at different concentrations: P = 0.0061 (0.1 μM); P = 0.0203 (1 μM); P = 0.0219 (5 μM); P = 0.0185 (10 μM).

(**D**) Resistance to the ferroptosis inducer RSL3 conferred by overexpression of *SREBF2* in IGR-37 melanoma cells. Y axis represents relative cell viability. X axis represents drug concentrations (μM) in log10 scale. Statistical significance is assessed by two-sided T test with unequal variance. P values are calculated by comparing cell viabilities between the two groups treated with RSL3 at different concentrations: P = 0.0088 (0.01 mM); P = 0.0106 (0.1 μM); P = 0.0208 (1 μM); P = 1.40×10^{-6} (5 μM).

Figure S10

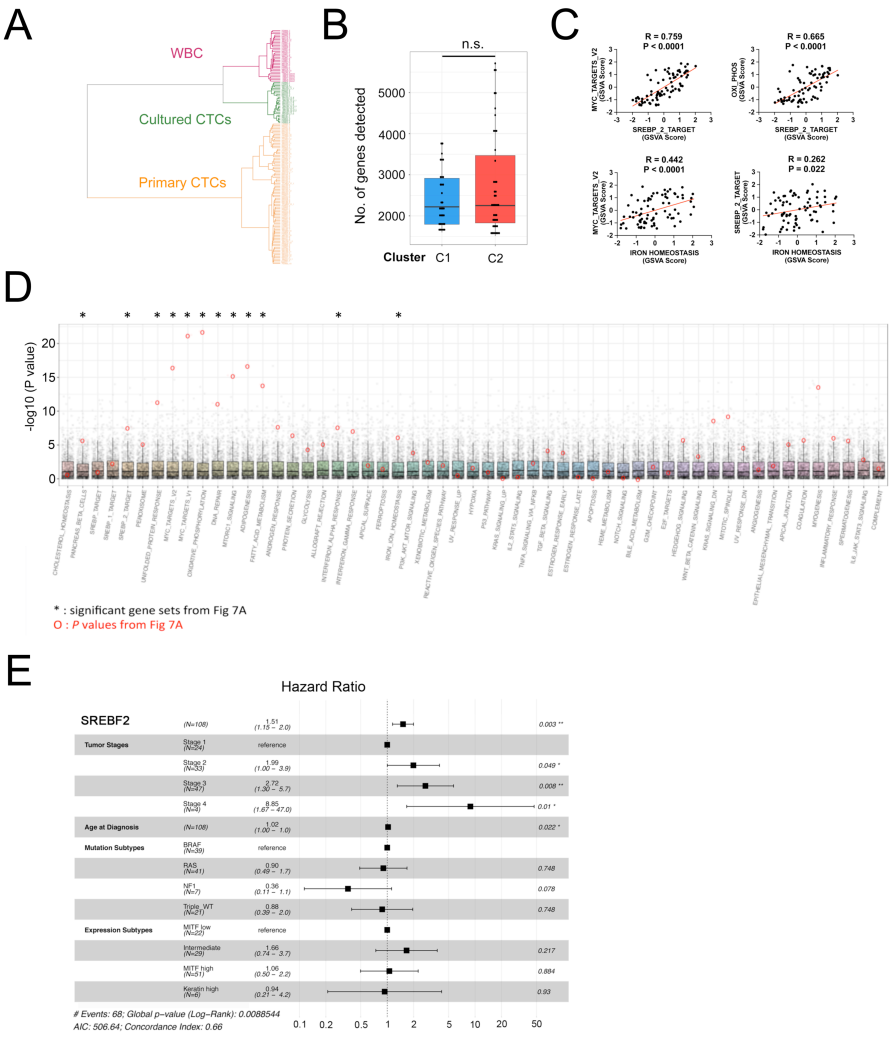


Figure S10. Clustering analysis of lipogenesis and iron homeostasis signatures in primary melanoma CTCs.

(**A**) Validation of single cell CTC RNA-seq from individually selected candidate CTCs following microfluidic enrichment, using hierarchical clustering strategy. 132 freshly isolated candidate single CTCs (primary) were purified through the CTC-iChip and initially differentiated from contaminating leukocytes based on their based on their large cell size and being negative for CD45 surface marker expression, as determined by live cell staining(3). Their identity as melanoma CTCs was then confirmed by RNA-seq analysis, and clustering with respect to 20 single CTCs (from patients Mel-167, Mel-182-

1, Mel-182-2 and PEM-22) that had been incubated in culture medium <8 weeks (to ensure viability but prior to proliferation as cell lines), versus 6 leukocytes (WBC, negative control) from healthy donors. The clustering resulted in 8 candidate CTCs that grouped with cultured CTCs (green) and 95 candidate CTCs clustering together (orange). Another 29 candidate CTCs grouped with WBCs (red) and were discarded (likely resulting from contamination during selection). The 103 primary CTCs were then subjected to further validation based on expression of at least one established melanoma CTC marker (1). This resulted in a total of 76 sequence-confirmed primary CTCs, (RNA-Seq clustering shown in Figure 7A).

(**B**) Box plot analysis showing the number of genes detected for each single CTC sample in cluster 1 (C1) and cluster 2 (C2) (shown in Figure 7A). There are no significant differences between C1 and C2 clusters. Wilcoxon rank-sum test was used to assess the statistical significance, $P = 0.6636$.

(**C**) Correlation between GSVA pathway enrichment scores of two pathways of interest across 76 validated primary CTCs showing "SREBF2_TARGET" signature *vs* "MYC_TARGET_V2" (upper left); "SREBF2_TARGET" *vs* "Oxi_Phos, OXIDATIVE_PHOSPHORYLATION" (upper right); "IRON HOMEOSTASIS" *vs* "MYC_TARGET_V2" (lower left); "IRON HOMEOSTASIS" *vs* "SREBF2_TARGET" (lower right). Pearson correlation coefficient and associated *P* values are shown in each plot. Curated pathway gene lists are found in Supplementary Table S2. GSVA scores of each pathway and expression levels of all genes within each selected pathway are listed in Supplementary Table S6.

(**D**) Boxplot of P values for the difference in GSVA enrichment scores between C1 and C2 (in Figure 7A) for randomly permuted gene sets. A boxplot of P values for the difference in GSVA enrichment scores between cluster 1 (C1) and cluster 2 (C2) for the gene sets in

which gene set size did not change but genes in each gene set were randomly chosen from 4,114 genes (4,144 genes from the Linnorm transformed TPM data; refer to the 'Primary CTC scRNA-seq analysis' section from supplementary methods). Each grey dot represents the P values from one of the 1,000 randomly generate gene sets and red dot represents the P values from the Figure 7A. P values were calculated from two-sided Welch's t-test.

(**E**) Reduced *SREBF2* mRNA expression is associated with improved metastasis-free survival. Cox proportional hazards multivariate analysis of high-purity (Consensus Purity Estimate>0.8) primary melanoma and metastases samples from the TCGA dataset showing significant association of *SREBF2* mRNA expression with disease-specific survival, after correcting for "Tumor Stages", "Age at Diagnosis", "Mutation Subtypes", and "Expression Subtypes" (see Supplementary Table S8).