ANKLE1 cleaves mitochondrial DNA and contributes to cancer risk by driving the Warburg effect and apoptosis resistance

Piotr Przanowski[a] ✉, **Róża K. Przanowska**[a,b], **and Michael J. Guertin**[c,d] ✉

[a]Department of Biochemistry and Molecular Genetics, University of Virginia School of Medicine, Charlottesville, Virginia, United States of America
[b]Department of Biomedical Engineering, University of Virginia School of Medicine, Charlottesville, Virginia, United States of America
[c]Center for Cell Analysis and Modeling, University of Connecticut, Farmington, Connecticut, United States of America
[d]Department of Genetics and Genome Sciences, University of Connecticut, Farmington, Connecticut, United States of America

Alleles within the chr19p13.1 locus are associated with increased risk of both ovarian and breast cancer and increased expression of the *ANKLE1* gene. ANKLE1 is molecularly characterized as an endonuclease that efficiently cuts branched DNA and shuttles between the nucleus and cytoplasm. However, the role of ANKLE1 in mammalian development and homeostasis remains unknown. In normal development ANKLE1 expression is limited to the erythroblast lineage and we found that ANKLE1's role is to cleave the mitochondrial genome during erythropoiesis. We show that ectopic expression of ANKLE1 in breast epithelial-derived cells leads to genome instability and mitochondrial DNA (mtDNA) cleavage. mtDNA degradation then leads to mitophagy and causes a shift from oxidative phosphorylation to glycolysis (Warburg effect). Moreover, mtDNA degradation activates STAT1 and expression of epithelial-mesenchymal transition (EMT) genes. Reduction in mitochondrial content contributes to apoptosis resistance, which may allow precancerous cells to avoid apoptotic checkpoints and proliferate. These findings provide evidence that ANKLE1 is the causal cancer susceptibility gene in the chr19p13.1 locus and describe mechanisms by which higher ANKLE1 expression promotes cancer risk.

ANKLE1 | chr19p13.1 | Triple-Negative Breast Cancer | mtDNA degradation | mitophagy | apoptosis

Correspondence: ***piotr.przanowski@gmail.com; guertin@uchc.edu***

Introduction

Alleles within the cancer susceptibility locus chr19p13.1 were first found to modify the risk of breast cancer in BRCA1 mutation carriers, triple negative breast cancer (TNBC), and ovarian cancer (Antoniou et al. 2010; Bolton et al. 2010). The authors suggested that *BABAM1* was acting as the causal gene to modify breast and ovarian cancer risk based upon its physical interaction with the BRCA1 protein complex and its expression in ovarian cancer. Integration of GWAS data with expression quantitative trait loci (eQTL) analysis implicated either *ABHD8* and/or Ankyrin repeat and LEM-domain containing protein 1 (*ANKLE1*) as candidate causal genes (Lawrenson et al. 2016). The authors concluded that *ABHD8* was the most plausible causal gene in the locus based upon chromatin conformation assays, deletion analysis, expression data, and cell migration experiments (Lawrenson et al. 2016). We also integrated GWAS and eQTL analysis, along with evolutionary conservation data, ChIP-seq, and chromatin accessibility data to identify the likely causal variant and causal gene, with a focus on variants that disrupt transcription factor binding (Liu et al. 2017). We proposed that a variant within a CCCTC-binding factor (CTCF) binding site reduces the affinity for CTCF and causes *ANKLE1* expression to increase (Liu et al. 2017). Our work suggested that *ANKLE1* was the most likely the casual gene within the chr19p13.1 breast and ovarian cancer susceptibility locus.

Few studies have explored the molecular, cellular, and physiological functions of ANKLE1. ANKLE1 was first studied in 2012 as an uncharacterized endonuclease that requires its LEM domain and GIY-YIG motifs for DNA cleavage *in vivo* (Brachner et al. 2012). Interestingly, ANKLE1 has high specificity for cleaving branched DNA (Song et al. 2020). Specificity for branched DNA is consistent with an observation in *C.elegans* that its homolog resolves chromatin bridges during late mitosis (Hong et al. 2018a) and is involved in the the regulation of meiotic recombination repair and chromosome segregation (Hong et al. 2018b). However ANKLE1 is dispensable for resolving chromatin bridges, meiotic recombination, and DNA repair in mice (Braun et al. 2016). *ANKLE1* is primarily expressed in hematopoietic tissues of vertebrates, but *ANKLE1*-deficient mice are viable without any detectable phenotype in hematopoiesis (Braun et al. 2016).

Herein we confirm that ANKLE1 expression is normally limited to the erythroblast lineage and we determine that the developmental role of ANKLE1 is to cleave the mitochondrial genome during erythropoiesis. Although this appears to be the molecular and cellular function of ANKLE1, the biological relevance of this function with regards to organismal development remains unclear. We also explore the undesired role that ectopic expression of ANKLE1 plays in conferring breast cancer risk. We find that ectopic expression of ANKLE1 in epithelial breast cells leads to genome instability, the Warburg effect, and resistance to apoptosis.

Results

ANKLE1 is the causal gene for breast and ovarian cancer risk in the chr19p13.1 region.

Expression quantitative trait loci (eQTL) data have revolutionized how geneticists identify candidate causal genes from genome-wide association study (GWAS) loci. We integrated the most recent meta-analysis of breast cancer GWAS (Zhang et al. 2020) and Genotype-Tissue Expression (GTEx) project data; we found that *ANKLE1* eQTL variants colocalize with the cancer susceptibility GWAS variants (Figure 1A). In contrast, we found minimal colocalization with other genes in the region (Figure S1A). Statistical colocalization analysis (Wallace 2020) demonstrates that *AN-*

KLE1 expression and genetic risk of breast cancer share a single casual variant in the locus with a probability of 0.75 (Figure 1B). This finding is consistent with recent integrative genomic analyses (Ferreira et al. 2019; Hoffman et al. 2017; Liu et al. 2017). Alleles that are associated with increased risk of breast and ovarian cancer are associated with high expression of *ANKLE1* (Figure 1C and Figure S1B). *ANKLE1* is more highly expressed in triple negative breast cancers (TNBC) compared to adjacent normal tissue and other breast cancer subtypes (Figure 1D).

Breast cancer is classified into at least four distinct subtypes. The chr19p13.1 locus was originally identified as a breast cancer risk region in BRCA1 mutation carriers and TNBC patients. We compared the effect sizes for the most significant polymorphisms in this region and confirmed that the variants specifically contribute to risk of developing TNBC (Figure 1E). We also compared eQTL significance to disease risk effect size for all genes in the region and these values are most highly correlated with the *ANKLE1* eQTL in TNBC (Figure S1C).

Expression of ANKLE1 in normal tissue is limited to erythroblasts.

Although *ANKLE1* gene expression is associated with breast and ovarian cancer, no Mendelian inherited human diseases are caused by *ANKLE1* mutations and *ANKLE1* knockout mice have no discernible phenotypes (Braun et al. 2016). In an effort to understand the normal biological role of ANKLE1, we examined ANKLE1 protein expression in different tissues in the Human Protein Atlas (Thul et al. 2017). ANKLE1 protein is only detected in a small subpopulation of cells within the bone marrow tissue (Figure S2A and Figure 2A). Expression analysis of RNA isolated from different bone marrow cell subpopulations indicates that ANKLE1 expression is limited to erythroblasts (Figure 2B) (Konuma et al. 2011). *ANKLE1*'s expression increases during erythroblast differentiation (Figure S2B), which indicates a possible role in red blood cell development. We hypothesized that ANKLE1 promotes enucleation and DNA fragmentation in erythrocytes based on its expression and the presence of its endonuclease domain. We leveraged an *in vitro* model of erythroblast differentiation of human leukemia K562 cells to test this hypothesis (Alves et al. 2011; Witt et al. 2000). Consistent with the expression pattern observed in human erythroblast differentiation (Figure S2B) (Ludwig et al. 2019), *ANKLE1* is transcriptionally activated in differentiating K562 cells (Figure S2C). Contrary to our expectations, DNA fragmentation is indistinguishable throughout differentiation between clonal *ANKLE1* knockout (KO) and clonal wild-type (WT) K562 cells (Figure 2C). This result is consistent with earlier experiments showing that ANKLE1 is dispensable for normal erythropoiesis in *ANKLE1* KO mice (Braun et al. 2016). Lastly, differentiating *ANKLE1* KO clones enucleate at a slightly faster rate

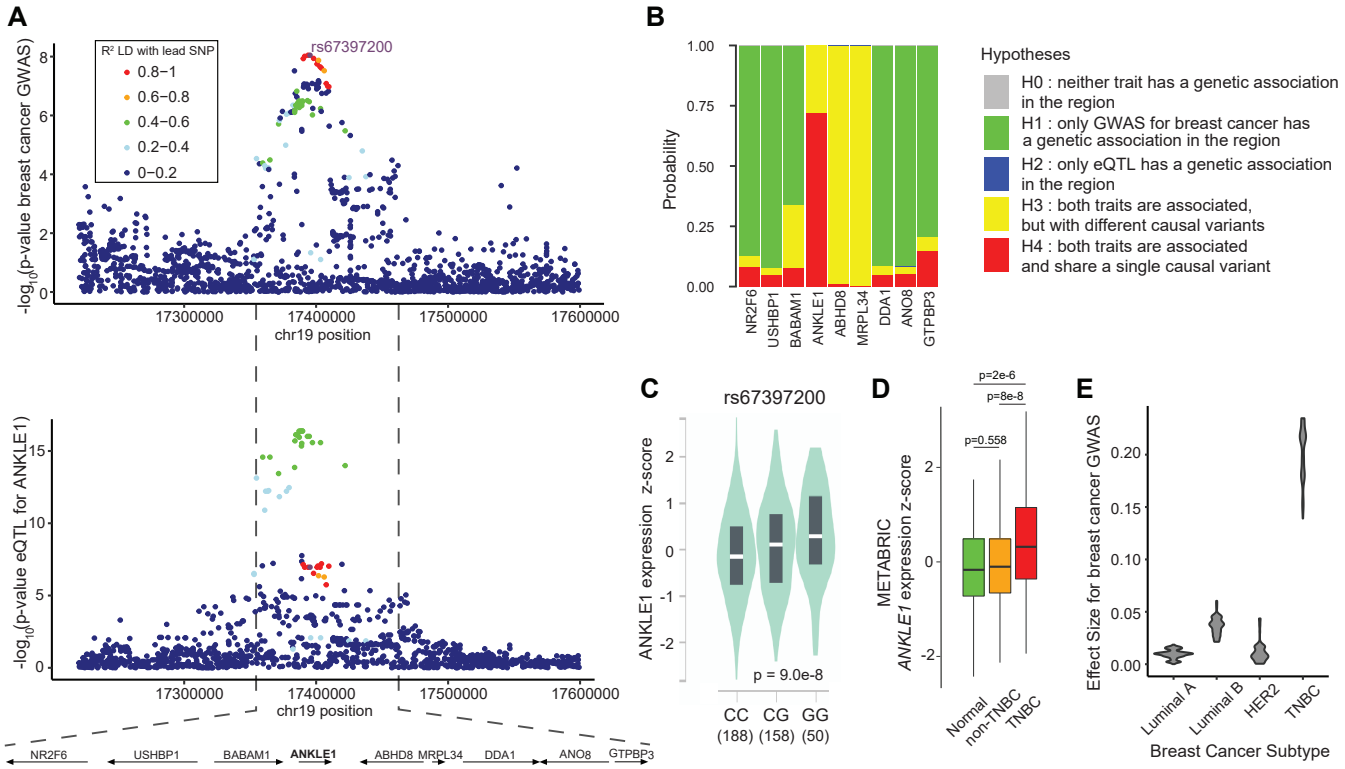

Fig. 1. *ANKLE1* is the most likely causal gene for Triple-Negative Breast Cancer risk in the 19p13.1 region. A) The breast cancer susceptibility GWAS variants (upper panel) and *ANKLE1* eQTL variants (lower panel) colocalize. B) Genetic colocalization analysis indicates that both ANKLE1 expression and breast cancer GWAS are associated and likely share a single causal variant. C) The G allele of rs67397200, which is associated with increased breast and ovarian cancer risk, is associated with higher expression of *ANKLE1* in breast tissue. The number of individuals with each genotype is indicated in parentheses. D) *ANKLE1* expression as measured by METABRIC is higher in Triple Negative Breast Cancer (TNBC) than in Normal (tumor-adjacent) or non-TNBC (Luminal A, Luminal B, HER2) tissue (p-values are calculated with a two-tailed t-test). E) Violin plots illustrate that TNBC is the predominant breast cancer subtype associated with the alleles within the 19p13.1 locus.

(Figure S2E) and contain normal hemoglobin levels (Figure S2F).

ANKLE1 localizes to the mitochondria to promote mtDNA degradation and mitophagy.

In addition to removing their nucleus, erythrocytes also digest their mitochondria and degrade mtDNA. Mitochondrial DNA is circular, but it is also distinct from nuclear DNA with respect to the lack of nucleosomes and a unique structural feature known as a displacement loop (D-loop) (Kasamatsu et al. 1971). The D-loop structure contains two branched DNA sites; ANKLE1 cleaves branched DNA orders of magnitude more efficiently than B-form DNA (Song et al. 2020). We hypothesized that ANKLE1 functions to cleave mitochondrial DNA in red blood cell development to facilitate mtDNA degradation. We found that mitochondrial DNA copy number is higher in *ANKLE1* KO lines throughout differentiation (Figure 2D). Mitophagy, or autophagy of the mitochondrial organelle, and enucleation occur during erythropoiesis in humans and in the K562 model (Moras et al. 2017) (Figure S2D). Reduced mtDNA is not accompanied by reduction in mitochondrial mass (Figure S2G), which suggests that mitophagy is not impaired in *ANKLE1* KO cells.

Since ANKLE1 has a role in regulating mtDNA copy number in the erythropoiesis model, we leveraged genomic data to determine whether ANKLE1 expression inversely correlates with mtDNA level. Two independent studies found that the rs2288464-A allele, which is located at the chr19p13.1 locus, is associated with higher mtDNA level in human blood (Ganel et al. 2021; Guyatt et al. 2019). Consistent with our hypothesis, the rs2288464-A variant correlates with decreased *ANKLE1* expression in human blood (Figure S3A).

Mitochondrial DNA is depleted in breast cancer tissue relative to normal tissue (Reznik et al. 2016). We show that mtDNA content in breast cancer is inversely correlated with *ANKLE1* expression (Figure S3B). These findings are consistent with the proposed role of ANKLE1, which is to decrease the level of mtDNA in human cells.

The *ANKLE1* knockout model of erythropoiesis, expression and mtDNA quantification data, and the preferential cleavage of branched DNA all suggest a role for ANKLE1 in regulating mtDNA abundance through D-loop cleavage. We directly characterized the phenotypes associated with overexpression of ANKLE1 in HEK293T cells. We sorted

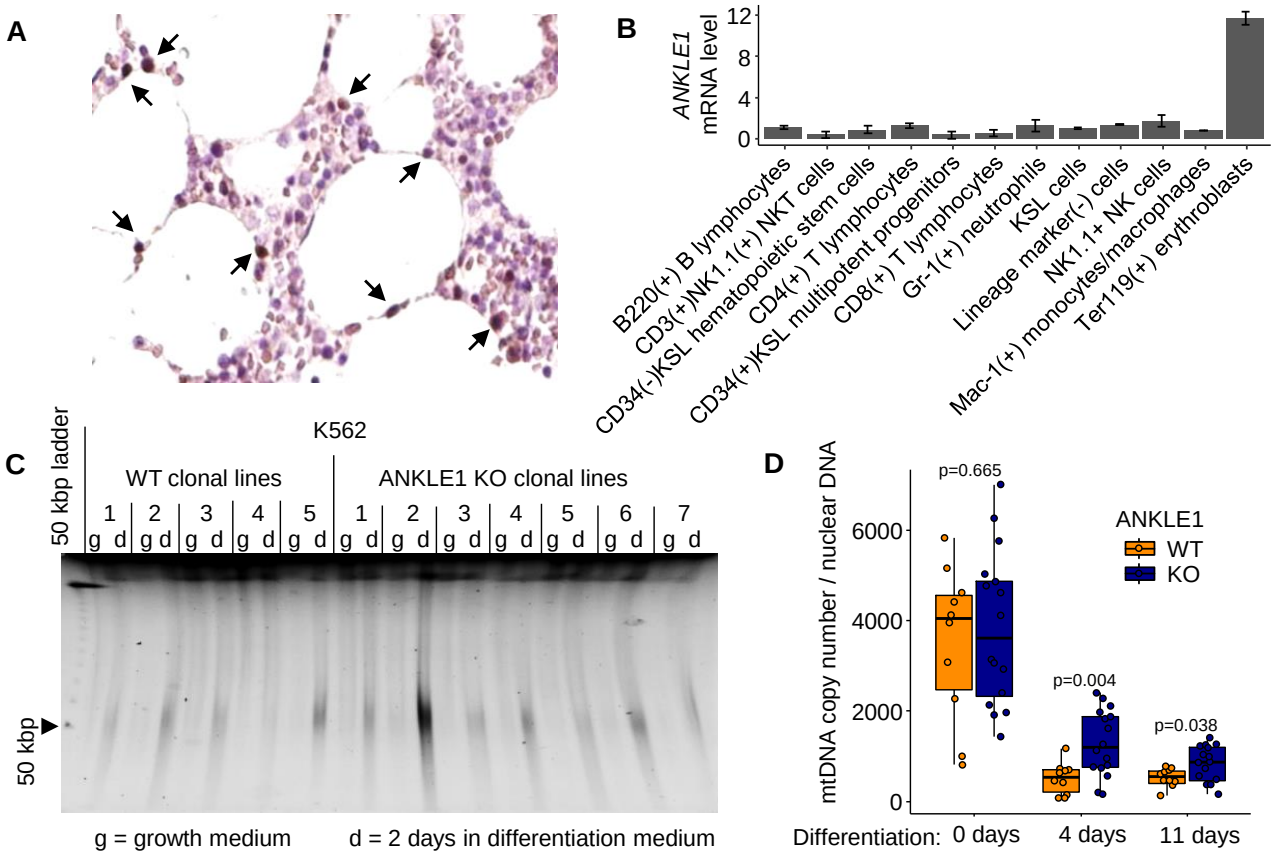

Fig. 2. The role of ANKLE1 in development is limited to mtDNA degradation in differentiating erythroblasts. A) Immunohistochemistry (brown) for ANKLE1 in human bone marrow tissue highlights rare population of cells expressing ANKLE1, as shown by the arrows. Images were reproduced with permission from `https://www.proteinatlas.org/ENSG00000160117-ANKLE1/tissue/bone+marrow#img`. B) Microarray data (GDS3997, probe 1443978at ANKLE1) from human bone marrow shows that *ANKLE1* is specifically expressed in the erythroblast lineage (Konuma et al. 2011). C) Nuclear DNA fragmentation is no different between wild-type (WT) and knock-out (KO) *ANKLE1* K562 cells, as shown by pulse field gel electrophoresis (PFGE) analysis of DNA isolated from undifferentiated and differentiated (2 days) K562 cells. D) Quantitative PCR with primers specific to nuclear and mitochondrial DNA shows that mitochondrial DNA copy number is higher in *ANKLE1* KO K562 clonal lines compared to WT lines during differentiation (p-values are calculated with a two-tailed t-test).

cells to select for non-apoptotic cells that express either GFP-ANKLE1 or GFP alone (Figure S3C). We found that ANKLE1 expression reduced the levels of mtDNA by over two-fold (Figure 3A) and decreased mitochondrial mass by 25% (Figure 3B). Next, we aimed to determine whether ANKLE1 induces mitophagy by quantifying the colocalization of lysosomes and mitochondria, which is indicative of active mitophagy. We used confocal microscopy to show that mitochondria and lysosomes stained by either mitotracker and lysotracker or Mito-CFP and Lyso-RFP colocalize (Figure 3C&D). Moreover, ANKLE1 also colocalizes with staining of lysosome and mitochondria organelles (Figure 3C&D). We further quantified mitophagy by imaging flow cytometry of cells with either GFP or GFP-ANKLE1 that were stained with mitotracker, lysotracker and DAPI (Figure 3E). Bright detail similarity analysis confirmed that ANKLE1 colocalizes with mitochondria (Figure 3F) and, to smaller extent, with lysosomes (Figure 3G). We found that ANKLE1 is uniformly distributed in cytoplasm in some cells, so we compared these cells to cells characterized by punctate ANKLE1 distribution. Cells with punctate distribution of ANKLE1 have reduced mitochondrial mass (Fig-

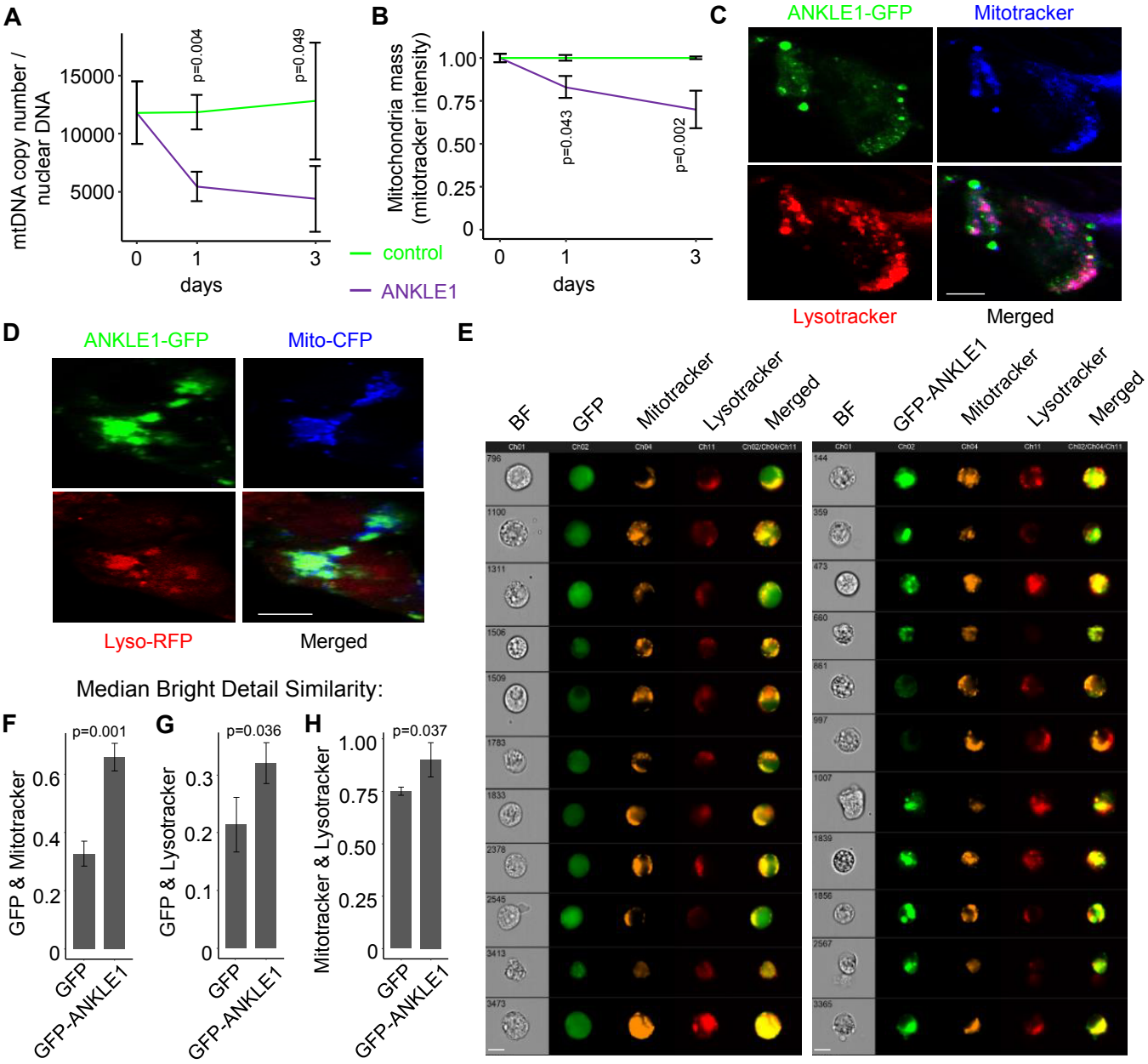

Fig. 3. ANKLE1 localizes to mitochondria, degrades mtDNA, and leads to mitophagy. A-B) Overexpression of ANKLE1 decreases mtDNA level (A) and decreases mitochondria mass (B) in HEK293T cells compared to the GFP expressing controls. C) Confocal microscopy images (scale bar 10 µm) of HEK293T cells overexpressing GFP-ANKLE1, stained with mitotracker and lysotracker show colocalization of ANKLE1, mitochondria, and lysosomes. D) Confocal microscopy imaging (scale bar 10 µm) with flourescent proteins, as an orthogonal method to panel (C), also shows colocalization of ANKLE1, mitochondria, and lysosomes. E) Representative cells from imaging flow cytometry of HEK293T cells overexpressing GFP (control) or GFP-ANKLE1, stained with DAPI (to exclude dead cells), mitotracker, and lysotracker (scale bar 10 µm) show colocalization of ANKLE1, mitochondria, and lysosomes. F-G) We quantified all the imaging flow cytometry data with bright detail similarity analysis of GFP, mitotracker, and lysotracker which more rigorously shows that ANKLE1 colocalizes with both mitochondria (F) and lysosomes (G). H) The same analysis shows that mitochondria and lysosomes colocalize more in the presence of ANKLE1, indicating an increase in mitophagy (p-values are calculated with a two-tailed t-test).

ure S3D& E), which suggests that ANKLE1 translocates to the mitochondria and precipitates mitochondrial degradation. Lastly, mitochondria colocalize with lysosomes more frequently in the presence of ANKLE1, indicating that mitophagy is increased in GFP-ANKLE1 cells compared to GFP-alone (Figure 3H).

Expression of ANKLE1 leads to mitochondria degradation, DNA damage, Epithelial to Mesenchymal transition, and STAT1 activation.

Since the data suggest that ANKLE1 function is mediated through the mitochondria in normal and disease states, we hypothesized that ANKLE1 overexpression would lead to changes in expression of energy production and respiration genes. We performed genomic transcriptome profiling (RNA-seq) after overexpression of ANKLE1 for 1, 3, and 7 days. Several gene set and gene ontology terms related to metabolic changes and mitchondrial regulation are enriched in the differentially expressed genes (Figure 4A&B and Figure S4D&E). *Hypoxia* and *glycolysis* are enriched gene sets among the activated genes (Figure 4A and Figure S4B). *Electron transport chain: OXPHOS system in mitochondria* is the most significant ontology term for repressed genes (Figure 4B).

Previous work found that despite harboring endonuclease activity and both a nuclear import and export signal, ANKLE1 is not nuclearly localized and does not cause an extensive DNA-damage response (Brachner et al. 2012; Zlopasa et al. 2016). However, we observed that DNA damage response gene sets are activated upon ANKLE1 overexpression (Figure 4A and Figure S4B&D). We used a more sensitive assay, Pulse Field Gel Electrophoresis (PFGE), to assay ANKLE1's ability to cut nuclear DNA when overexpressed. We excluded DNA breaks caused by apoptosis by selecting non-apoptotic cells (Figure S3C). We found that ANKLE1 cuts DNA with low frequency, resulting in fragments between 50 kbp and 200 kbp (Figure 4C).

Epithelial-mesenchymal transition (EMT) is the most significantly enriched gene set among ANKLE1 regulated genes (Figure 4A and Figure S4B). EMT is also a defining feature of breast cancer transformation (Micalizzi and Ford 2009; Prat and Perou 2011). Several enriched gene sets represent hallmarks of EMT, such as TGF beta signaling, apical junction, and KRAS signaling (Figure 4A).

We found that many gene sets that relate to STAT1 signaling were enriched among the gene sets (Figure 4A&B and Figure S4B,D&E). Previous work found that mtDNA double strand breaks leads to phosphorylation of STAT1, which leads to genomic transcriptional changes (Tigano et al. 2021). Taken together with the mitochondria degradation phenotypes (Figure 3A&B), we hypothesized that mtDNA cleavage by ANKLE1 leads to STAT1 phosphorylation. Indeed, we found that STAT1 is phosphorylated in ANKLE1-expressing cells and this is not accompanied by activation of STAT3, STAT5, or NF-$\varkappa$B (Figure 4D).

ANKLE1 expression in normal breast epithelium cells drives the Warburg effect and apoptosis resistance in TP53 mutant cells.

Recall that ANKLE1 expressing cells exhibit hallmarks of gene expression that indicate a switch from oxidative phosphorylation to glycolysis. In an effort to recapitu-

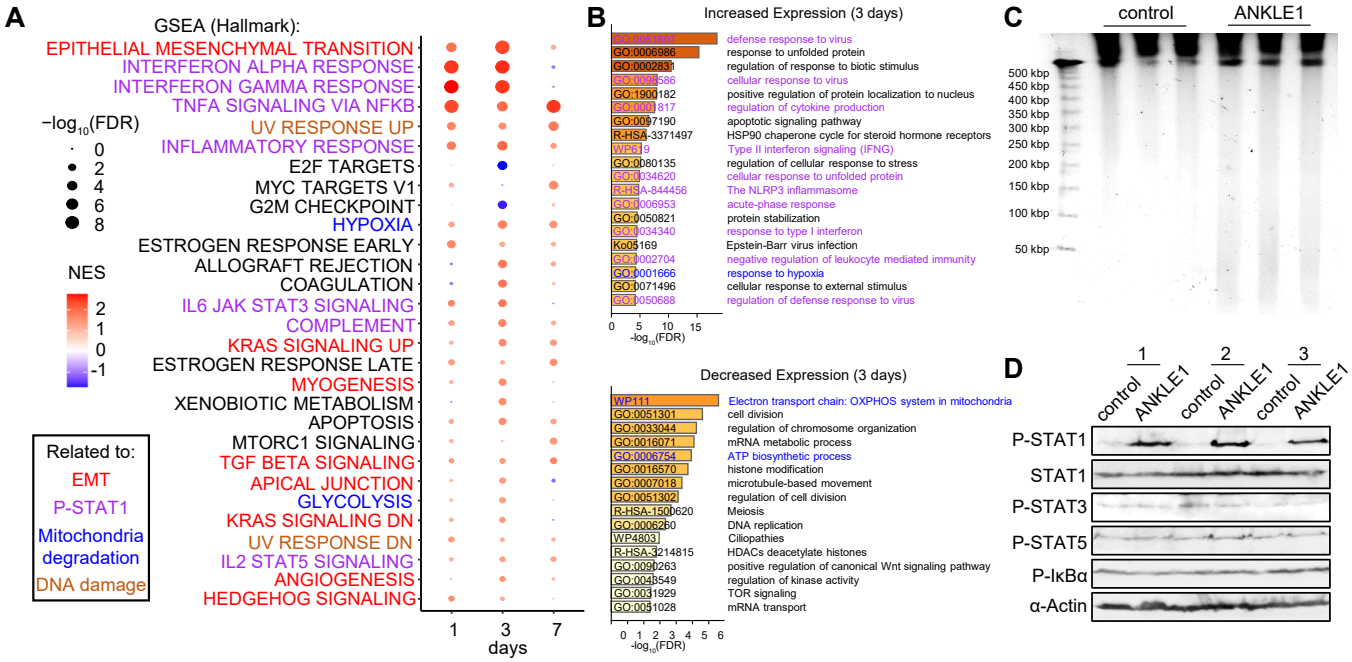

Fig. 4. Expression of ANKLE1 leads to STAT1 activation, mitochondria degradation, DNA damage, and Epithelial to Mesenchymal transition. A) Gene Set Enrichment Analysis (GSEA) of RNA-seq results for ANKLE1 overexpressing cells versus control cells identifies hallmarks of STAT1 activation, mitochondria degradation, DNA damage, and Epithelial to Mesenchymal transition. B) Gene ontology analysis for differentially expressed genes (Figure S4C) after 3-days post-transfection of ANKLE1 cells versus control cells show enrichment for STAT1 activation and mitochondria degradation. C) ANKLE1 cuts nuclear DNA in non-apoptotic cells, as shown by DNA Pulse Field Gel Electrophoresis (PFGE) from non-apoptotic cells that overexpress ANKLE1 (24 hours post-transfection). D) ANKLE1 leads to STAT1 activation after 24 hours of ANKLE1-transfection, as shown by western blots probed for P-STAT1, STAT1, P-STAT3, P-STAT5, P-IκB and α-Actin.

late the processes that may lead to transformation of normal breast cells, we began experiments in the MCF10A-5E cell line. MCF10A-5E cells are clonally-derived from non-tumorigenic epithelial MCF10A cells (ATCC). These cells exhibit homogeneous behavior in 3D cultures and express many of the hallmarks of epithelial breast cells, such as epithelial sialomucins, cytokeratins, and milk fat globule antigen (Janes et al. 2010; Tait et al. 1990). Similarly to what we observed in HEK293T cells, ANKLE1 induces mitophagy in MCF10A-5E cells (Figure 5A). We found transient *ANKLE1* expression leads to chronic phenotypic changes within MCF10A-5E cells. We performed a 22 day time course experiment and we found that *ANKLE1* expression peaks within a day of transfection (Figure 5B). *ANKLE1* expression gradually decreases to background levels between 11-22 days after transfection (Figure S5A). mtDNA is decreased to approximately 50% and persists at this level for the remainder of the time course (Figure 5B). These data indicate that transient *ANKLE1* expression can lead to stable changes in cell metabolism.

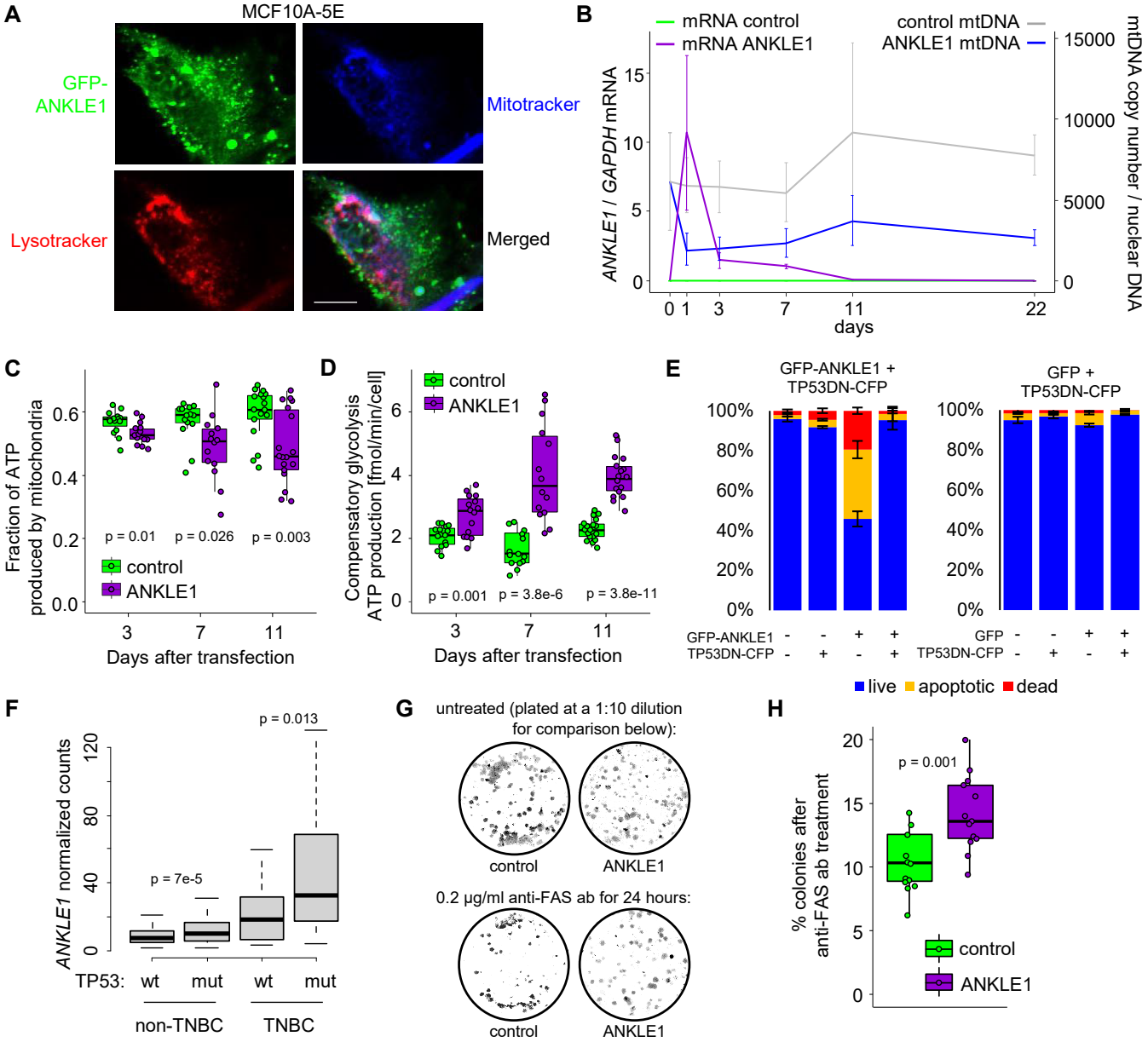

Fig. 5. ANKLE1-induced mitophagy in normal breast cancer cells shifts the metabolism to glycolysis and increases resistance to apoptosis in TP53 negative cells. A) Confocal microscopy images (scale bar = 10 µm) of MCF10A-5E cells overexpressing GFP-ANKLE1 stained with mitotracker and lysotracker show a colocalization of ANKLE1, mitochondria, and lysosomes. B) We quantified relative levels of *ANKLE1* mRNA and mtDNA over a 22 day time course of transient *ANKLE1* expression. Although *ANKLE1* levels return to baseline (Figure S5A), the decrease of mtDNA level is maintained. C-D) We used the Seahorse XF Real-Time ATP Rate Assay to show that ANKLE1 decreases the fraction of ATP produced by mitochondria (C) and ANKLE1 increases maximum compensatory glycolysis in MCF10A-5E (D). E) Flow cytometry analysis quantification of Figure S5E shows that ANKLE1 induces apoptosis in MCF10A-5E cells and apoptosis is abolished by a TP53 dominant negative mutant. F) The level of *ANKLE1* mRNA in human noTNBC and TNBC samples isolated from TP53 wild-type or mutant tumors indicates higher expression when TP53 is mutated. G-H) ANKLE1 overexpression produced more colonies when MCF10A-5E cells were treated with anti-FAS activating antibodies, even though ANKLE1 transient expression occurred 11 days prior to anti-FAS treatment (p-values are calculated with a two-tailed t-test).

We hypothesized that ANKLE1 expression contributes to a decrease in ATP produced by oxidative phosphorylation relative to ATP produced by glycolysis, which is known as the Warburg effect. We simultaneously measured total ATP production and ATP produced in the mitochondria of ANKLE1-transfected MCF10A-5E cells compared to GFP-transfected control cells. We found that, on average, the fraction of ATP produced by the mitochondria is lower at days 3, 7, and 11 post-transfection (Figure 5C). Next, we blocked oxidative phosphorylation and quantified ATP produced by glycolysis in ANKLE1 vs. GFP control cells. We found that ANKLE1-transfected cells can produce ATP by glycolysis at a faster rate when oxidative phosphorylation is blocked, this is termed *compensatory glycolysis* (Figure 5D, Figure S5B). We hypothesize that ANKLE1-mediated mitochondria degradation leads to a demand in energy resources that triggers activation of hypoxia (Figure 4A) and HIF1 pathways (Figure S4E), which lead to an increase in glycolysis. We repeated these experiments in HEK293T cells and we again found that ANKLE1 reduces the fraction of ATP produced in by the mitochondria and increases compensatory glycolysis (Figure S5C&D).

Throughout the experiments with ANKLE1 expressing MCF10A-5E cells, we empirically noticed that the fraction of apoptotic cells was higher than what we observed with ANKLE1 overexpression in HEK293T cells. The main genetic difference between these cell lines is the presence of large T antigen that deactivates TP53 in HEK293T cells. MCF10A-5E cells contain active TP53. Taken together with the fact that ANKLE1-mediated risk of breast cancer is primarily through its effect on the TNBC subtype, of which 80% are mutant for TP53 (Koboldt et al. 2012), we hypothesized that TP53 mutations allow cells that express ANKLE1 ectopically to evade apoptosis. We tested this hypothesis by transfecting MCF10A-5E cells with vectors carrying either GFP or GFP-ANKLE1 along with a CFP-TP53 dominant negative R175H mutant (TP53DN). We quantified apoptosis after 24 hours with AnnexinV-PE and 7-AAD (Figure S5E). We found that ANKLE1-alone induces apoptosis in 60% of MCF10A-5E cells (Figure 5E). ANKLE1-driven apoptosis in MCF10A-5E cells is completely abolished in the presence of dominant negative TP53. In contrast, ANKLE1 does not increase apoptosis compared to the GFP control in HEK293T cells (Figure S5F). If mutant TP53 protects against ANKLE1-induced apoptosis, we would expect that TP53-mutant tumors have higher levels of ANKLE1 expression. We queried the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) database and we found that breast cancer tissues with mutated TP53 have significantly higher ANKLE1 expression irrespective of their TNBC status (Figure S5F).

Mitochondria are known to play an essential role in apoptosis by regulating the release of proteins from the intermembrane space, which activates caspases (Wang and Youle 2009). To further highlight the importance of the mitochondria in apoptosis, previous work showed that modest differences in cellular mitochondrial content in clonal cell lines can determine the apoptotic fate upon activation of death

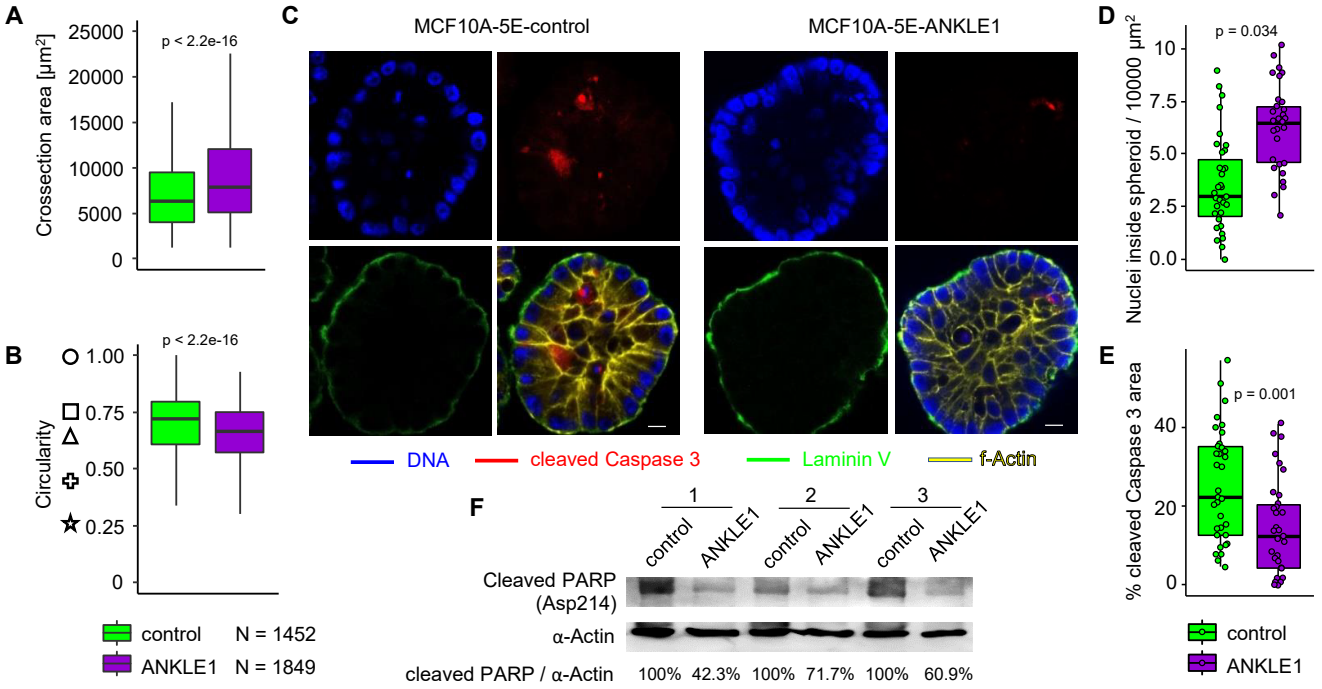

Fig. 6. Spheroids formed by normal breast epithelium cells after ANKLE1 expression exhibit precancerous phenotypes. A) MCF10A-5E cells that transiently expressed ANKLE1 form larger spheroids. B) MCF10A-5E cells that transiently expressed ANKLE1 form less circular spheroids. Note that the symbols to the left are examples of circularity at the indicated y-value. C) Representative confocal microscopy images (scale bar = 10 µm) of spheroids stained for DNA, cleaved Caspase 3, and f-Actin reveal the effect of transient ANKLE1 expression. D-E) Spheroids derived from MCF10A-5E cells transiently transfected with ANKLE1 contain more intact nuclei (D) and exhibit less cleaved Caspase 3 staining (E) inside the spheroid. F) Western blot shows a lower level of cleaved PARP in protein extracts from spheroids derived from MCF10A-5E cells transiently expressing ANKLE1. We used densitometry to quantify the percent of cleaved PARP relative to the control.

receptor (Márquez-Jurado et al. 2018). We designed an experiment to directly test whether ANKLE1-induced depletion of the mitochondria (Figure 5B) increases the resistance of death receptor-triggered apoptosis in MCF10A-5E cells. Recall that at day 11 ANKLE1 expression is at background levels, but mitochondrial DNA copy number is reduced by 50% (Figure S5A). We found that transiently expressed ANKLE1 prior to day 11 are more resistant to death receptor-induced apoptosis and able to proliferate and form colonies (Figure 5G&H).

Transient, ectopic expression of ANKLE1 induces phenotypes consistent with hallmarks of carcinogenesis.

Thus far we have shown that ANKLE1 promotes the Warburg effect and resistance to apoptosis within HEK293T and MCF10A-5E cells. Recall that the MCF10A-5E clone was selected for its ability to form homogenous 3D spheroids on the matrigel surface, which makes the model amenable for measuring features that are associated with transformation and cancer progression. MCF10A-5E cells cultured on matrigel form round hollow spheroids due to apoptosis of internally localized cells and this orderly growth is disrupted upon the introduction of an oncogene (Debnath et al. 2003; Janes et al. 2010). We transiently transfected ANKLE1 into MCF10A-5E cells and 11 days later we seeded them onto matrigel for an additional 11 days. We quantified the size and circularity of spheroids with OrganoSeg (Borten et al. 2018) (Figure S6A). Although MCF10A-5E cells no longer express ANKLE1 when they are seeded on matrigel, the spheroids that form are larger (Figure 6A) and less circular (Figure 6B and Figure S6B). We hypothesized that the spheroids that previously expressed ANKLE1 have a defect in apoptosis of internally localized cells, so we stained for cleaved Caspase 3 (a marker of apoptosis) and DNA (Figure 6C). Indeed, spheroids formed from acute ANKLE1-expressing cells have more internally located intact nuclei (Figure 6D) and less apoptosis (Figure 6E). We confirmed less apoptosis in ANKLE1 spheroids by quantifying cleaved PARP protein levels by western blotting (Figure 6F). These data indicate that transient, ectopic expression of ANKLE1 in breast epithelium provides resistance to apotosis, which can trigger spheroid phenotypes that are consistent with cancer progression.

Discussion

Although there are relatively few reports on the function of ANKLE1, the data suggest that ANKLE1 originally evolved as a means to resolve DNA bridges (Hong et al. 2018a,b). Resolving DNA bridges is consistent with ANKLE1's preferred specificity to cleave branched DNA (Song et al. 2020). While the molecular function of cleaving branched DNA remains unchanged in vertebrates, the ANKLE1 protein may have been functionally repurposed for mtDNA degradation in mammals. Mammals have evolved to degrade both their nuclei and mitochondria during erythropoiesis. The nucleus is actively expelled from the cell in a process known as enucleation and mitochondria are digested by mitophagy. Our data support a complementary mechanism to ensure degradation of the mtDNA in differentiating erythroblasts by ANKLE1 (Figure 2D and Figure S2G). The mitochondrial chromosome is circular and more stable because, without free DNA ends, circular DNA is refractory to digestion by exonucleases. Since mtDNA harbors a D-loop with branched DNA structures (Kasamatsu et al. 1971), *ANKLE1* cleavage of mtDNA may provide a mechanism to linearize the mitochodiral chromosome to ensure its faithful degradation (Figure 7A). ANKLE1-mediated mtDNA digestion may prevent accumulation of cytosolic mtDNA that escaped from autophagosomes, as this phenomenon is associated with multiple metabolic diseases (Pérez-Treviño et al. 2020).

ANKLE1 is not expressed in normal breast tissue and we hypothesize that the ectopic expression of ANKLE1 leads to changes in metabolism, transcription, and cellular homeostasis. In support of this assertion, we show that ANKLE1 triggers cleavage of nuclear DNA in cells without triggering apoptosis (Figure 4C) and ANKLE1 induces mtDNA degradation and mitophagy (Figure 3). The mitochondria phenotypes may drive STAT1 phosphorylation (Figure 4D) and a metabolic switch from oxidative phosphorylation to glycolysis (Figure 5).

In the 1920s Otto Warburg showed that cultured tumor tissues have high rates of glucose uptake and lactate secretion, even in the presence of oxygen (aerobic glycolysis) (Warburg et al. 1927). Although cancer is caused by mutations within tumor suppressors and oncogenes, the Warburg effect represents a critical metabolic switch for cancer cells to survive and proliferate. There are several postulated roles of the Warburg effect in cancer cells. For instance a higher rate of ATP production by glycolysis may provide a selective advantage when competing for shared and limited energy resources. The Warburg effect has also been proposed to be an adaptation mechanism to support the biosynthetic requirements of uncontrolled proliferation. The increased glucose consumption is used as a carbon source for anabolic processes needed to support cell proliferation, such as the *de novo* generation of nucleotides, lipids, and proteins. As the tumor microenvironment has been more recently studied and appreciated to have a role in cancer biology, it has been proposed that the Warburg effect may present an advantage for cell growth in a relatively acidic and low oxygen microenvironment. Acidification of the microenvironment by lactate may allow for enhanced invasiveness and is a contributor to tumor-associated macrophage polarization that support tumor growth (Liberti and Locasale 2016). We suggest that ANKLE1-induced degradation of the mitochondria facilitates the Warburg effect to promote cancer growth.

ANKLE1-induced mitophagy may also promote survival of cancer cells by avoiding apoptosis. We show that ANKLE1 expression leads to lower mitochondria content and resistance to apoptosis caused by death receptor activation (Figure 5H&I). Activation of the death receptor is the primary mechanism by which the immune system removes

dysregulated or mutated cells (O'Reilly et al. 2016). In support of this apoptosis-resistance hypothesis, previous work has shown that cells with fewer mitochondria are more resistant to apoptosis induced by death receptor activation (Márquez-Jurado et al. 2018). Another possible mechanism by which ectopic ANKLE1 expression may promote tumorigenesis in TP53 mutant cancers, which represent over 80% of TNBCs, is to increase mutational burden. We found that ANKLE1 cuts DNA in the nucleus without causing apoptosis in TP53 mutant cells, which will increase the mutation rate.

In summary, the biological role of ANKLE1 is limited to degrading mtDNA in differentiating erythroblasts (Figure 7B). Ectopic ANKLE1 expression in normal breast cells leads to genome instability and degradation of mtDNA, which causes mitophagy, activation of STAT1, resistance to apoptosis, and a shift to aerobic glycolysis as a means of energy production (Figure 7C).

Methods

Analysis of publicly available data sets and statistical analysis.

GWAS data were obtained from published metaanalysis summary statistics (Zhang et al. 2020). Expression QTL data were obtained from GTEx GCS bucket (https://console.cloud.google.com/storage/browser/gtex-resources): dbGaP accession number phs000424.v8.p2 (Consortium et al. 2020). Violin plots of *ANKLE1* expression for different genotypes in breast (Figure 1 and Figure S1) or blood (Figure S3) tissue were obtained from the GTEx Portal. Colocalization analysis was performed

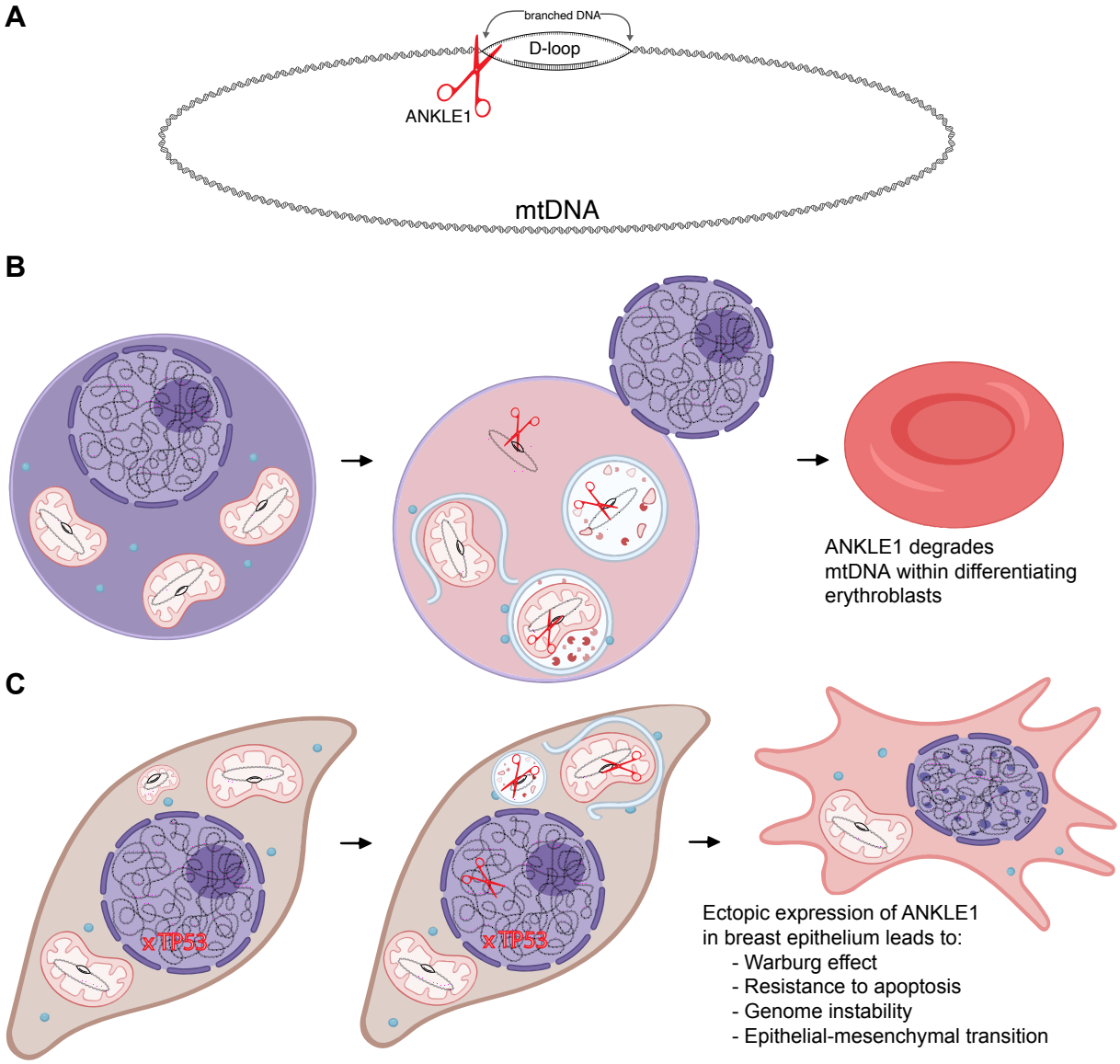

Fig. 7. A model of ANKLE1 function in erythroblast differentiation and breast cancer. A) ANKLE1 preferentially cuts branched DNA and we speculate that ANKLE1-mediated cleavage of the mtDNA D-loop facilitates degradation of mitochondrial DNA. B) We propose that the role of ANKLE1 in erythroblast differentiation is limited to mtDNA degradation. C) Ectopic expression of ANKLE1 in breast epithelium also causes mtDNA degradation, and this leads to mitophagy, the Warburg effect and resistance to apoptosis. Additionally, ANKLE1 in the nucleus may contribute to genome instability. Our RNA-seq data indicate that these phenotypes may lead to transcriptional changes that are consistent with the Epithelial-Mesenchymal transition.

with the `coloc` package (Wallace 2020) using GWAS and QTL p-values to assess colocalization. The effect sizes from the summary statistics were partitioned for different breast cancer subtypes and only the most significant ($p<0.0001$) GWAS single nucleotide polymorphisms (SNPs) were queried. *ANKLE1* expression in bone marrow subpopulations was analyzed based on microarray data: GDS3997, probe 1443978at (Konuma et al. 2011). *ANKLE1* expression during erythropoiesis was quantified by RNA-seq (GSE115684) (Ludwig et al. 2019). mtDNA copy number was previously measured by integrating whole exome and whole genome sequencing data (Reznik et al. 2016). Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) data (Curtis et al. 2012) were accessed through the Xena platform (Goldman et al. 2020) and used to obtain TNBC, non-TNBC, normal adjacent, and TP53 mutation status, as well as *ANKLE1* expression (p-values were calculated with a two-tailed t-test). Circularity was calculated as: 4π *Area / Perimeter ^{2}.

Plasmid construction.

The pEGFP-C1-ANKLE1 plasmid that was used for all ANKLE1 expression experiments was kindly provided by Roland Foisner (Brachner et al. 2012). The pEGFP-C1 plasmid was used as the control. sgRNA constructs for *ANKLE1* knockout were generated by inserting oligonucleotides containing the targeted sequences (5′- TTCAGGGCACAGCCTAGAAC -3′ and 5′- GATTCTGCCCTAGCCCCACC -3′) into the pX458 vector (Addgene Plasmid #48138 (Ran et al. 2013)). MitoCFP plasmid was obtained from Addgene (Plasmid #58426 (Mishra et al. 2014)). LysoRFP plasmid was obtained from Addgene (Plasmid #58061, donated by Michael Davidson). TP53DN(R175H)-CFP was constructed by cloning CFP from mitoCFP plasmid into AgeI site of p53 (dominant negative R175H mutant)-pcw107-V5 (Addgene Plasmid #64638 (Martz et al. 2014)), using 5′- GGGTTAGGGATAGGCTTACCACCGGTTTACTTGTACAGCTCGTCCATGC -3′ and 5′- CTTGTACAAAGTGGTTACCGGAGGATCCGGTGGTGTGAGCAAGGGCGAGGAGCTG -3′ primers for PCR and In-Fusion Cloning (Takara Bio).

Cell culture, treatment, transfection, and generation of stable cell lines.

K562 and HEK293T cells were purchased from ATCC. MCF10A-5E cells were kindly provided by Kevin Janes. K562 cells were cultured in RPMI supplemented with 10% FBS (growth condition) or in RPMI (-glutamine) supplemented with 10% FBS and 1mM sodium butyrate (differentiation medium) (Canh Hiep et al. 2012). HEK293T cells were cultured in DMEM supplemented with 10% FBS. MCF10A-5E cells were cultured in DMEM/F12 supplemented with 5% horse serum, EGF (20 ng/ml), hydrocortisone (0.5 µg/ml), cholera toxin (0.1 µg/ml stock), and insulin (10 µg/ml) (Debnath et al. 2003). Cells were first transient transfected and cultured for 11 days, then the cells were seeded on the surface of matrigel (Thermo Fisher Scientific) in DMEM/F12 supplemented with 2% horse serum, EGF (2 ng/ml), hydrocortisone (0.5 µg/ml), cholera toxin (0.1 µg/ml stock), insulin (10 µg/ml), and 2% matrigel for another 11 days to form spheroids (Debnath et al. 2003). Cells were treated with MitoTracker™ Red CMXRos (100 nM) and LysoTracker™ Deep Red (50 nM) 30 min before fixation or live imaging. All cell lines were transfected with Lipofectamine 3000 reagent. K562 *ANKLE1* KO lines were obtained by co-transfecting two plasmids carrying sgRNAs for *ANKLE1* or by transfecting pX458 as a control. The GFP-positive cells were sorted into 96-well plates one day after transfection. Clones were validated by PCR with the following primers: 5′- GGTTAGTCTTCCCAGGGCAC -3′ and 5′- GCCTCCCGTGTATAAGCCTC -3′ (1838 bp PCR product for WT clones vs. 270 bp PCR product for KO clones). Hemoglobin was measured by absorbance at 425 nm of 10^5 cells suspended in 0.1 ml of PBS. For the colony formation assay, cells were treated for 24 h with 500 ng/ml Anti-Fas (human, activating) clone CH11 antibody (Milipore) and seeded on 6-well plates at the specified dilutions. In the top panel of Figure 5G, only 10% of untreated cells were imaged to ensure that the number of colonies were visually comparable to the antibody-treated cells. Colonies were fixed with methanol and stained with Enhanced Gram Crystal Violet Ethanol Solution (ThermoFisher Scientific).

qPCR.

Genomic DNA was isolated with QuickExtract™ DNA Extraction Solution. RNA isolation, cDNA synthesis, and qPCR were performed with Power SYBR™ Green Cells-to-CT™ Kit using StepOne™ Real-Time PCR System (Applied Biosystems). Mitochondria content was measured by primers specific for human mtDNA (primers: 5′- CACCCAAGAACAGGGTTTGT -3′ and 5′- TGGCAATGGGTATGTTGTTA -3′) and genomic DNA (primers: 5′- GAGGCAGGACTCAGGACAAG -3′ and 5′- GGATGCCTCAGGGACCAG -3′). *ANKLE1* RNA level was measured by primers specific for *ANKLE1* (primers: 5′-ACACCCTTCACCAGGCAGTT -3′ and 5′-AAAACTCTGGGCCAGGAGCAA -3′) and we used the following *GAPDH* primers: 5′- TGCACCACCAACTGCTTAGC -3′ and 5′- GGCATGGACTGTGGTCATGAG -3′.

Immunofluorescence and microscopy.

Cells or spheroids were fixed with 4% PFA (20% Paraformaldehyde Solution (Electron Microscopy Science)) for 1 h. Spheroids were blocked and permabilized with 0.5 % Triton, 100mM Glycine, 10 % FBS in PBS for 30 min, then stained overnight with Laminin-5 Alexa Fluor 488 Conjugated antibodies (Milipore), Cleaved Caspase-3 (Asp175) (D3E9) Alexa Fluor 467 antibodies (Cell Signaling) and ActinRed 555 (ReadyProbes, Invitrogen) overnight in 0.2 % Triton, 10 % FPS in PBS. Stained spheroids were mounted in ProLong Gold antifade reagent with DAPI (Invitrogen) on glass coverslips and imaged with Zeiss LSM 710 Multiphoton confocal microscope. Images were quantified using the Fiji software (Schindelin et al. 2012).

Flow cytometry.

Enucleation was measured by staining with DRAQ5 (Fisher Scientific) and Phalloidin-FITC (TOCRIS) and assessing the fraction of DRAQ5 negative cells. Apoptosis was detected by staining with 7-AAD and Anexin V-PE (Apoptosis Detection Kit I - BD Pharmingen). Attune NxT flow cytometer (Life Technologies) was used for all flow cytometry experiments. For imaging flow cytometry live cells were stained with DAPI (Sigma) and run on Amnis ImageStreamX Mark II (Luminex). Cells were sorted with Influx Cell Sorter (BD Biosciences). Flow cytometry data were analysed with FCS express software.

Western blot.

Cells were lysed in IPH buffer (50 mM Tris-Cl, 0.5% NP-40%, 50 mM EDTA). Protein lysates were run on 10% polyacrylamide SDS-PAGE gels and transferred to nitrocellulose membranes (Amersham Protran 0.45 µm NC - GE Healthcare Life Sciences). Membranes were blocked for 30 minutes in 5% milk in TBST buffer and incubated overnight with primary antibody (1:1000). Secondary antibody (1:5000) incubation was carried out for 1 hour after washing with TBST, and before washing and incubation with SuperSignal™ West Pico PLUS Chemiluminescent Substrate (ThermoFisher). The following primary antibodies were used: Anti-Phospho-Stat1 (Tyr701) (58D6) antibody (Cell Signaling), Anti-Stat1 (D1K9Y) antibody (Cell Signaling), Anti-Phospho-Stat3 (Tyr705) (D3A7) antibody (Cell Signaling), Anti-Phospho-Stat5 (Tyr694) (D47E7) antibody (Cell Signaling), Anti-Phospho-IκBα (Ser32) (14D4) antibody (Cell Signaling), Anti-Cleaved PARP (Asp214) (D64E10) antibody (Cell Signaling) and Anti-β-actin (AC-74) antibody (Sigma Aldrich). Densitometry was performed with Fiji (Schindelin et al. 2012).

ATP rate assay.

Percent ATP produced by mitochondria and the level of compensatory glycolysis were measured using a Seahorse XF Real-Time ATP Rate Assay Kit and XFe96 Analyzers (Agilent) according to manufacturer instructions.

Pulse Field Gel Electrophoresis (PFGE).

Live cells were embedded in 1% LMP agarose in 50mM EDTA (pH 8). Solidified plugs were incubated in 1% LDS (RPI Research Products International), 100 mM EDTA, 10 mM Tris pH 8.0 overnight and then washed 5 times (30 min each wash) in 50 mM EDTA (pH 8). Plugs were then secured in 1% Certified Megabase Agarose (BIO-RAD) gel in 0.5x TBE buffer and run with Lambda Ladder (ProMega-Markers) using CHEF-MAPPER system (BIO-RAD).

RNA-seq.

Live, nonapoptotic, GFP-positive cells were sorted 1, 3 and 7 days after transfection with pEGFP-C1-ANKLE1 or pEGFP-C1 plasmids. RNA was isolated by TRIzol extraction using Direct-zol RNA MicroPrep Kit including DNase treatment (ZymoResearch). rRNA depletion was performed using a NEBNext rRNA Depletion Kit v2 (Human/Mouse/Rat) and RNA was purified using Agencourt RNAClean XP Beads (New England Biolabs). Libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs). Concentrations of libraries were measured with Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific) and fragment size distribution were measured with TapeStation (Agilent). Libraries were pooled together and sequenced at the University of Virginia Genome Analysis and Technology Core using the Illumina NextSeq2000 instrument. RNA-seq data were aligned to the human assembly hg38 (Gencode v33) using HISAT2 (Kim et al. 2015) and quantified by HTSeq (Anders et al. 2015). We applied `DESeq2` (Love et al. 2014) to identify differentially expressed genes with a false discovery rate less than 0.05 (Figure S4C). Principal component analysis was performed to ensure that replicates group together and that variation is observed between control and ANKLE1 expressing cells and time points, rather than between batches (Figure S4A). Gene Set Enrichment Analysis (GSEA) (Subramanian et al. 2005) was performed in `R` with the `fgsea` package (Korotkevich et al. 2021) using $\log_2$(Fold Change)*-$\log_{10}$(FDR) to rank all genes. Gene ontology analysis was done using Metascape (Zhou et al. 2019) on significantly (FDR < 0.05) activated or repressed genes. All RNA-seq library data files are available under GEO accession number GSE186393.

ACKNOWLEDGEMENTS

This work was funded by R35-GM128635 to MJG. RP was supported by an NCI Predoctoral to Postdoctoral Fellow Transition Award (F99/K00-CA253732). We want to thank University of Virginia Flow Cytometry Core (RRID: SCR_017829), University of Virginia Genome Analysis and Technology Core (RRID: SCR_018883) and University of Virginia Advanced Microscopy Facility (RRID: SCR_018736).

AUTHOR CONTRIBUTIONS

PP and MJG conceptualized and developed the project. PP performed all experiments with the exception of the RNA-seq experiment performed by RP. PP and MJG analyzed all the data and wrote the manuscript.

Bibliography

Alves LR, Costa ES, Sorgine MH, Nascimento-Silva MCL, Teodosio C, Bárcena P, Castro-Faria-Neto HC, Bozza PT, Orfao A, Oliveira PL, et al.. 2011. Heme-oxygenases during erythropoiesis in k562 and human bone marrow cells. *PLoS One* **6**: e21358.

Anders S, Pyl PT, and Huber W. 2015. Htseq—a python framework to work with high-throughput sequencing data. *bioinformatics* **31**: 166–169.

Antoniou AC, Wang X, Fredericksen ZS, McGuffog L, Tarrell R, Sinilnikova OM, Healey S, Morrison J, Kartsonaki C, Lesnick T, et al.. 2010. A locus on 19p13 modifies risk of breast cancer in brca1 mutation carriers and is associated with hormone receptor–negative breast cancer in the general population. *Nature genetics* **42**: 885–892.

Bolton KL, Tyrer J, Song H, Ramus SJ, Notaridou M, Jones C, Sher T, Gentry-Maharaj A, Wozniak E, Tsai YY, et al.. 2010. Common variants at 19p13 are associated with susceptibility to ovarian cancer. *Nature genetics* **42**: 880–884.

Borten MA, Bajikar SS, Sasaki N, Clevers H, and Janes KA. 2018. Automated brightfield morphometry of 3d organoid populations by organoseg. *Scientific reports* **8**: 1–10.

Brachner A, Braun J, Ghodgaonkar M, Castor D, Zlopasa L, Ehrlich V, Jiricny J, Gotzmann J, Knasmüller S, and Foisner R. 2012. The endonuclease ankle1 requires its lem and giy-yig motifs for dna cleavage in vivo. *Journal of cell science* **125**: 1048–1057.

Braun J, Meixner A, Brachner A, and Foisner R. 2016. The giy-yig type endonuclease ankyrin repeat and lem domain-containing protein 1 (ankle1) is dispensable for mouse hematopoiesis. *PLoS One* **11**: e0152278.

Canh Hiep N, Kinohira S, Furuyama K, and Taketani S. 2012. Depletion of glutamine enhances sodium butyrate-induced erythroid differentiation of k562 cells. *The journal of biochemistry* **152**: 509–519.

Consortium G et al.. 2020. The gtex consortium atlas of genetic regulatory effects across human tissues. *Science* **369**: 1318–1330.

Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, et al.. 2012. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**: 346–352.

Debnath J, Muthuswamy SK, and Brugge JS. 2003. Morphogenesis and oncogenesis of mcf-10a mammary epithelial acini grown in three-dimensional basement membrane cultures. *Methods* **30**: 256–268.

Ferreira MA, Gamazon ER, Al-Ejeh F, Aittomäki K, Andrulis IL, Anton-Culver H, Arason A, Arndt V, Aronson KJ, Arun BK, et al.. 2019. Genome-wide association and transcriptome studies identify target genes and risk loci for breast cancer. *Nature communications* **10**: 1–18.

Ganel L, Chen L, Christ R, Vangipurapu J, Young E, Das I, Kanchi K, Larson D, Regier A, Abel H, et al.. 2021. Mitochondrial genome copy number measured by dna sequencing in human blood is strongly associated with metabolic traits via cell-type composition differences. *Human genomics* **15**: 1–17.

Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, Banerjee A, Luo Y, Rogers D, Brooks AN, et al.. 2020. Visualizing and interpreting cancer genomics data via the xena platform. *Nature biotechnology* **38**: 675–678.

Guyatt AL, Brennan RR, Burrows K, Guthrie PA, Ascione R, Ring SM, Gaunt TR, Pyle A, Cordell HJ, Lawlor DA, et al.. 2019. A genome-wide association study of mitochondrial dna copy number in two population-based cohorts. *Human genomics* **13**: 1–17.

Hoffman JD, Graff RE, Emami NC, Tai CG, Passarelli MN, Hu D, Huntsman S, Hadley D, Leong L, Majumdar A, et al.. 2017. Cis-eqtl-based trans-ethnic meta-analysis reveals novel genes associated with breast cancer risk. *PLoS genetics* **13**: e1006690.

Hong Y, Sonneville R, Wang B, Scheidt V, Meier B, Woglar A, Demetriou S, Labib K, Jantsch V, and Gartner A. 2018a. Lem-3 is a midbody-tethered dna nuclease that resolves chromatin bridges during late mitosis. *Nature communications* **9**: 1–11.

Hong Y, Velkova M, Silva N, Jagut M, Scheidt V, Labib K, Jantsch V, and Gartner A. 2018b. The conserved lem-3/ankle1 nuclease is involved in the combinatorial regulation of meiotic recombination repair and chromosome segregation in caenorhabditis elegans. *PLoS genetics* **14**: e1007453.

Janes KA, Wang CC, Holmberg KJ, Cabral K, and Brugge JS. 2010. Identifying single-cell molecular programs by stochastic profiling. *Nature methods* **7**: 311–317.

Kasamatsu H, Robberson DL, and Vinograd J. 1971. A novel closed-circular mitochondrial dna with properties of a replicating intermediate. *Proceedings of the National Academy of Sciences* **68**: 2252–2257.

Kim D, Langmead B, and Salzberg SL. 2015. Hisat: a fast spliced aligner with low memory requirements. *Nature methods* **12**: 357–360.

Koboldt D, Fulton R, McLellan M, Schmidt H, Kalicki-Veizer J, McMichael J, Fulton L, Dooling D, Ding L, Mardis E, et al.. 2012. Comprehensive molecular portraits of human breast tumours. *Nature* **490**: 61–70.

Konuma T, Nakamura S, Miyagi S, Negishi M, Chiba T, Oguro H, Yuan J, Mochizuki-Kashio M, Ichikawa H, Miyoshi H, et al.. 2011. Forced expression of the histone demethylase fbxl10 maintains self-renewing hematopoietic stem cells. *Experimental hematology* **39**: 697–709.

Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, and Sergushichev A. 2021. Fast gene set enrichment analysis. *BioRxiv* p. 060012.

Lawrenson K, Kar S, McCue K, Kuchenbaeker K, Michailidou K, Tyrer J, Beesley J, Ramus SJ, Li Q, Delgado MK, et al.. 2016. Functional mechanisms underlying pleiotropic risk alleles at the 19p13. 1 breast–ovarian cancer susceptibility locus. *Nature communications* **7**: 1–22.

Liberti MV and Locasale JW. 2016. The warburg effect: how does it benefit cancer cells? *Trends in biochemical sciences* **41**: 211–218.

Liu Y, Walavalkar NM, Dozmorov MG, Rich SS, Civelek M, and Guertin MJ. 2017. Identification of breast cancer associated variants that modulate transcription factor binding. *PLoS genetics* **13**: e1006761.

Love MI, Huber W, and Anders S. 2014. Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome biology* **15**: 1–21.

Ludwig LS, Lareau CA, Bao EL, Nandakumar SK, Muus C, Ulirsch JC, Chowdhary K, Buenrostro JD, Mohandas N, An X, et al.. 2019. Transcriptional states and chromatin accessibility underlying human erythropoiesis. *Cell reports* **27**: 3228–3240.

Márquez-Jurado S, Díaz-Colunga J, das Neves RP, Martinez-Lorente A, Almazán F, Guantes R, and Iborra FJ. 2018. Mitochondrial levels determine variability in cell death by modulating apoptotic gene expression. *Nature communications* **9**: 1–11.

Martz CA, Ottina KA, Singleton KR, Jasper JS, Wardell SE, Peraza-Penton A, Anderson GR, Winter PS, Wang T, Alley HM, et al.. 2014. Systematic identification of signaling pathways with potential to confer anticancer drug resistance. *Science signaling* **7**: ra121–ra121.

Micalizzi DS and Ford HL. 2009. Epithelial–mesenchymal transition in development and cancer. *Future oncology* **5**: 1129–1143.

Mishra P, Carelli V, Manfredi G, and Chan DC. 2014. Proteolytic cleavage of opa1 stimulates mitochondrial inner membrane fusion and couples fusion to oxidative phosphorylation. *Cell metabolism* **19**: 630–641.

Moras M, Lefevre SD, and Ostuni MA. 2017. From erythroblasts to mature red blood cells: organelle clearance in mammals. *Frontiers in physiology* **8**: 1076.

O'Reilly E, Tirincsi A, Logue SE, and Szegezdi E. 2016. The janus face of death receptor signaling during tumor immunoediting. *Frontiers in immunology* **7**: 446.

Pérez-Treviño P, Velásquez M, and García N. 2020. Mechanisms of mitochondrial dna escape and its relationship with different metabolic diseases. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* **1866**: 165761.

Prat A and Perou CM. 2011. Deconstructing the molecular portraits of breast cancer. *Molecular oncology* **5**: 5–23.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, and Zhang F. 2013. Genome engineering using the crispr-cas9 system. *Nature protocols* **8**: 2281–2308.

Reznik E, Miller ML, Şenbabaoğlu Y, Riaz N, Sarungbam J, Tickoo SK, Al-Ahmadie HA, Lee W, Seshan VE, Hakimi AA, et al.. 2016. Mitochondrial dna copy number variation across human cancers. *elife* **5**: e10769.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al.. 2012. Fiji: an open-source platform for biological-image analysis. *Nature methods* **9**: 676–682.

Song J, Freeman AD, Knebel A, Gartner A, and Lilley DM. 2020. Human ankle1 is a nuclease specific for branched dna. *Journal of molecular biology* **432**: 5825–5834.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al.. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **102**: 15545–15550.

Tait L, Soule HD, and Russo J. 1990. Ultrastructural and immunocytochemical characterization of an immortalized human breast epithelial cell line, mcf-10. *Cancer research* **50**: 6087–6094.

Thul PJ, Åkesson L, Wiking M, Mahdessian D, Geladaki A, Blal HA, Alm T, Asplund A, Björk L, Breckels LM, et al.. 2017. A subcellular map of the human proteome. *Science* **356**.

Tigano M, Vargas DC, Tremblay-Belzile S, Fu Y, and Sfeir A. 2021. Nuclear sensing of breaks in mitochondrial dna enhances immune surveillance. *Nature* **591**: 477–481.

Wallace C. 2020. Eliciting priors and relaxing the single causal variant assumption in colocalisation analyses. *PLoS genetics* **16**: e1008720.

Wang C and Youle RJ. 2009. The role of mitochondria in apoptosis. *Annual review of genetics* **43**: 95–118.

Warburg O, Wind F, and Negelein E. 1927. The metabolism of tumors in the body. *Journal of General Physiology* **8**: 519–530.

Witt O, Sand K, and Pekrun A. 2000. Butyrate-induced erythroid differentiation of human k562 leukemia cells involves inhibition of erk and activation of p38 map kinase pathways. *Blood, The Journal of the American Society of Hematology* **95**: 2391–2396.

Zhang H, Ahearn TU, Lecarpentier J, Barnes D, Beesley J, Qi G, Jiang X, O'Mara TA, Zhao N, Bolla MK, et al.. 2020. Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nature genetics* **52**: 572–581.

Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, and Chanda SK. 2019. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature communications* **10**: 1–10.

Zlopasa L, Brachner A, and Foisner R. 2016. Nucleo-cytoplasmic shuttling of the endonuclease ankyrin repeats and lem domain-containing protein 1 (ankle1) is mediated by canonical nuclear export-and nuclear import signals. *BMC cell biology* **17**: 1–12.

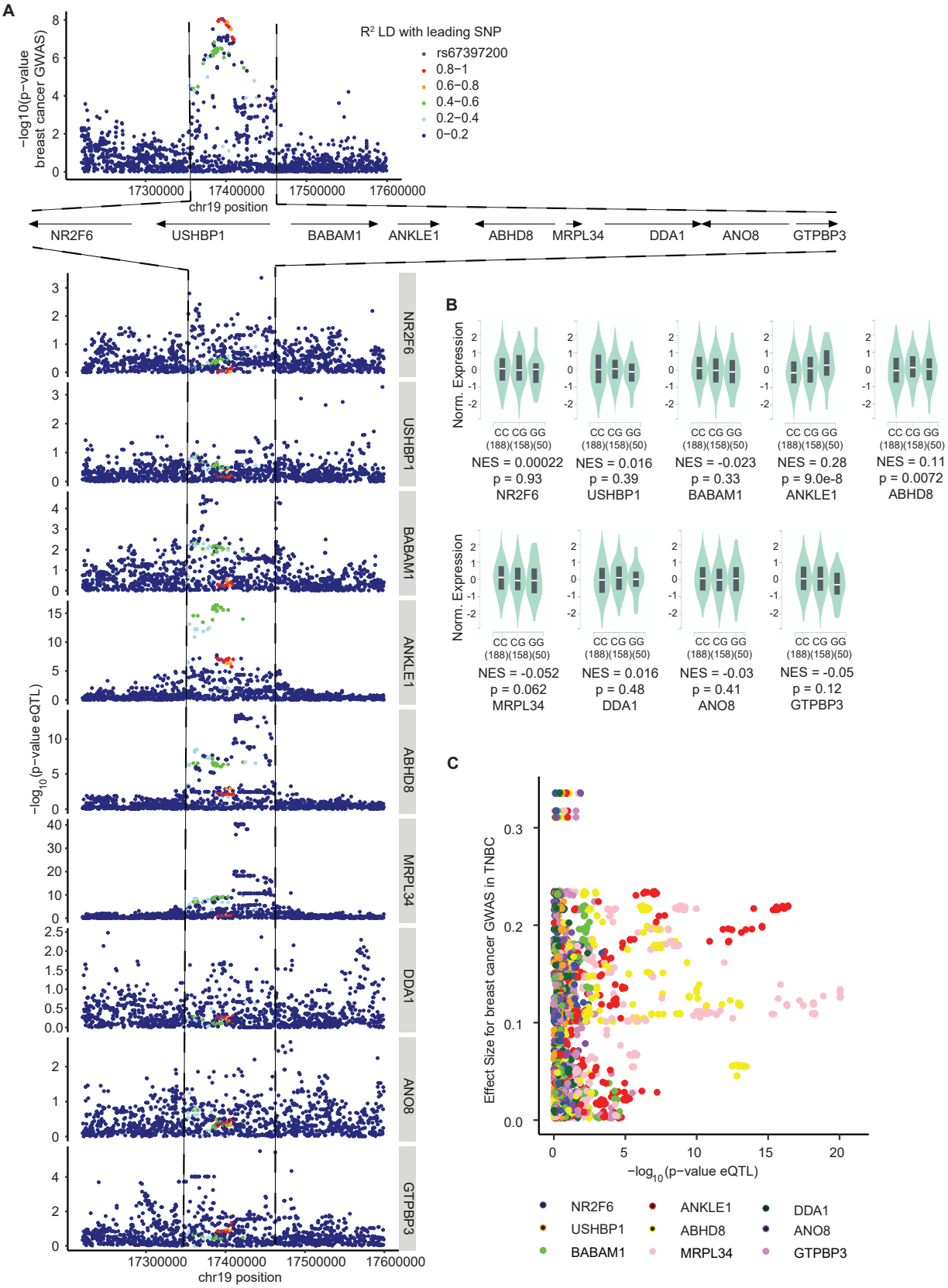

Fig. S1. Comparison of *ANKLE1* with other genes in the locus confirms that *ANKLE1* is the most likely causal breast cancer susceptibility gene within chr19p13.1. A) The breast cancer susceptibility GWAS variants within the chr19p13.1 locus (higher panel) colocalize predominantly with *ANKLE1* eQTL variants as opposed to the other local genes (lower panel). B) The G allele of rs67397200, which is associated with increased breast cancer risk, is associated with higher expression of *ANKLE1* in blood. The risk allele does not associate with the expression of other genes in the locus, except *ABHD8* albeit with a less significant p-value. C) The *ANKLE1* eQTL p-value correlates best with TNBC GWAS effect size.

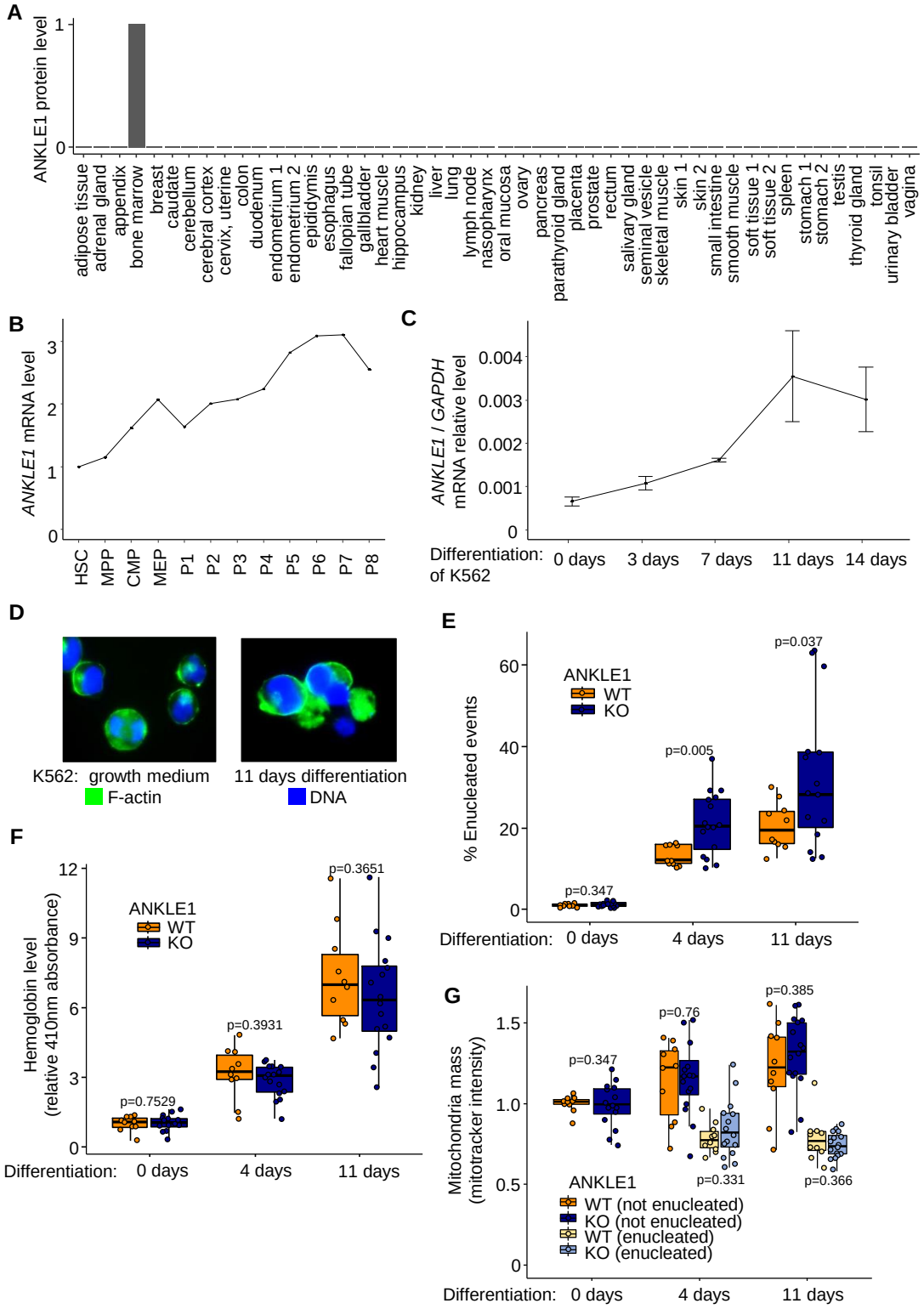

Fig. S2. *ANKLE1* expression is only within the bone marrow and increases throughout erythropoiesis, but knockout of *ANKLE1* has only modest phenotypes. A) ANKLE1 protein expression in human tissues is limited to bone marrow (https://www.proteinatlas.org/). B) *ANKLE1* mRNA level increases throughout erythroid differentiation (RNA-seq data: GSE115684 (Ludwig et al. 2019)): hematopoietic stem cells (HSC), multipotential progenitors (MPP), common myeloid progenitors (CMP), megakaryocyte erythroid progenitors (MEP), myeloid progenitor (P1) to reticulocyte (P8). C) *ANKLE1* mRNA increases in differentiating K562 cells. D) A representative picture of K562 cells shows enucleation after 11 days in differentiation medium (right panel) compared to undifferentiated cells (left panel). E) The fraction of enucleated cells is increased in *ANKLE1* KO K562 cells during differentiation. F) Hemoglobin level is unchanged upon *ANKLE1* KO in K562 cells during differentiation. G) *ANKLE1* KO does not affect mitochondria mass during K562 cell differentiation (p-values are calculated with a two-tailed t-test).

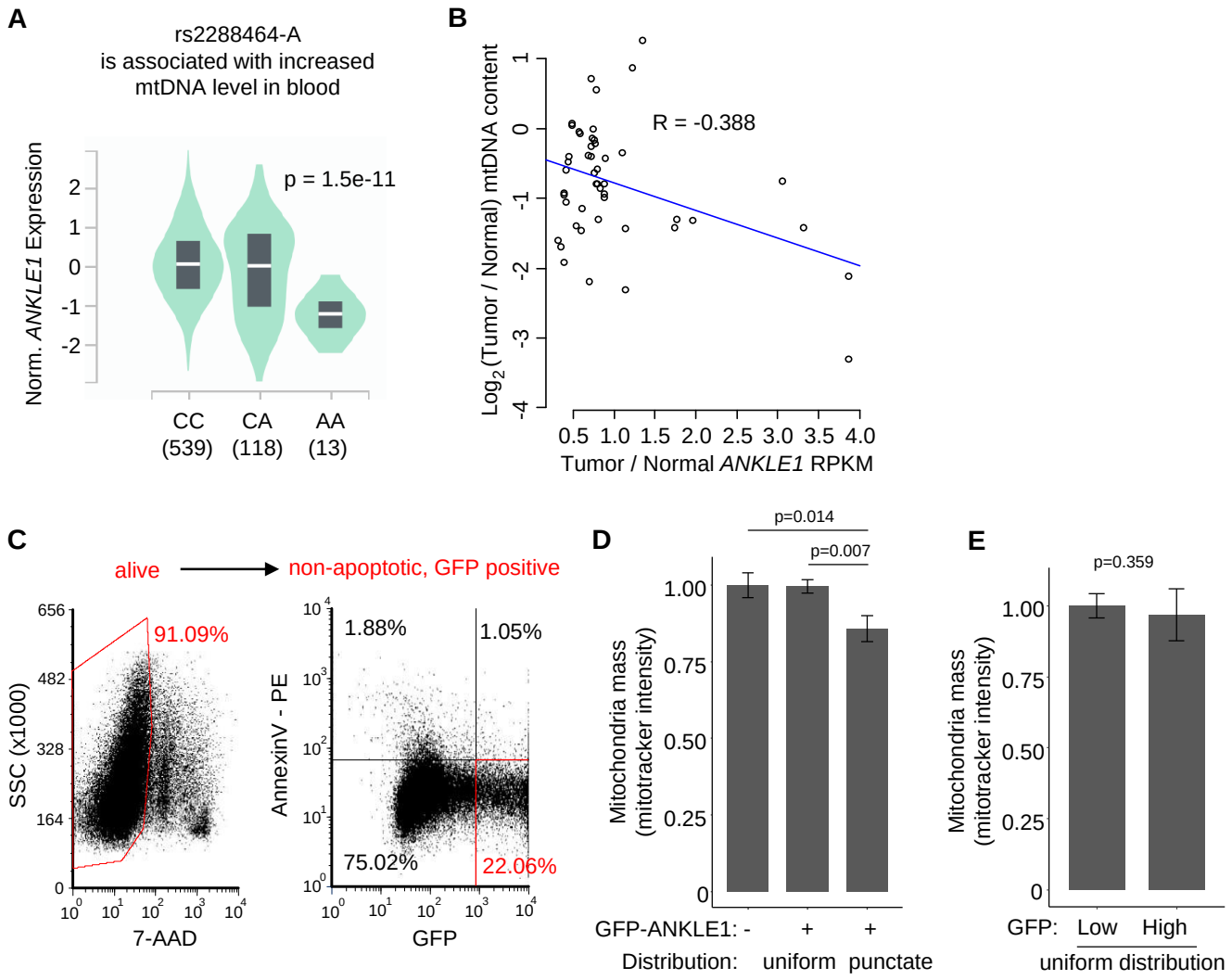

Fig. S3. *ANKLE1* expression correlates with decreased mtDNA and ANKLE1 protein localizes to the mitochondria. A) The rs2288464-A allele is associated with increased mtDNA level in blood and lower *ANKLE1* expression. B) *ANKLE1* expression negatively correlates with mtDNA fold change between tumor and normal tissue in breast cancer (Reznik et al. 2016). C) Alive, non-apoptotic and GFP-positive cells were selected using the illustrated gating strategy. D) Imaging flow cytometry of HEK293T cells overexpressing GFP-ANKLE1 indicates that ANKLE1 expression causes lower mitochondria content (measured by mitotracker fluorescence) only when ANKLE1 is distributed in puncta. E) In contrast, differences in ANKLE1 intensity does not cause lower mitochondria content when ANKLE1 is uniformly distributed throughout the cell (p-values are calculated with a two-tailed t-test).

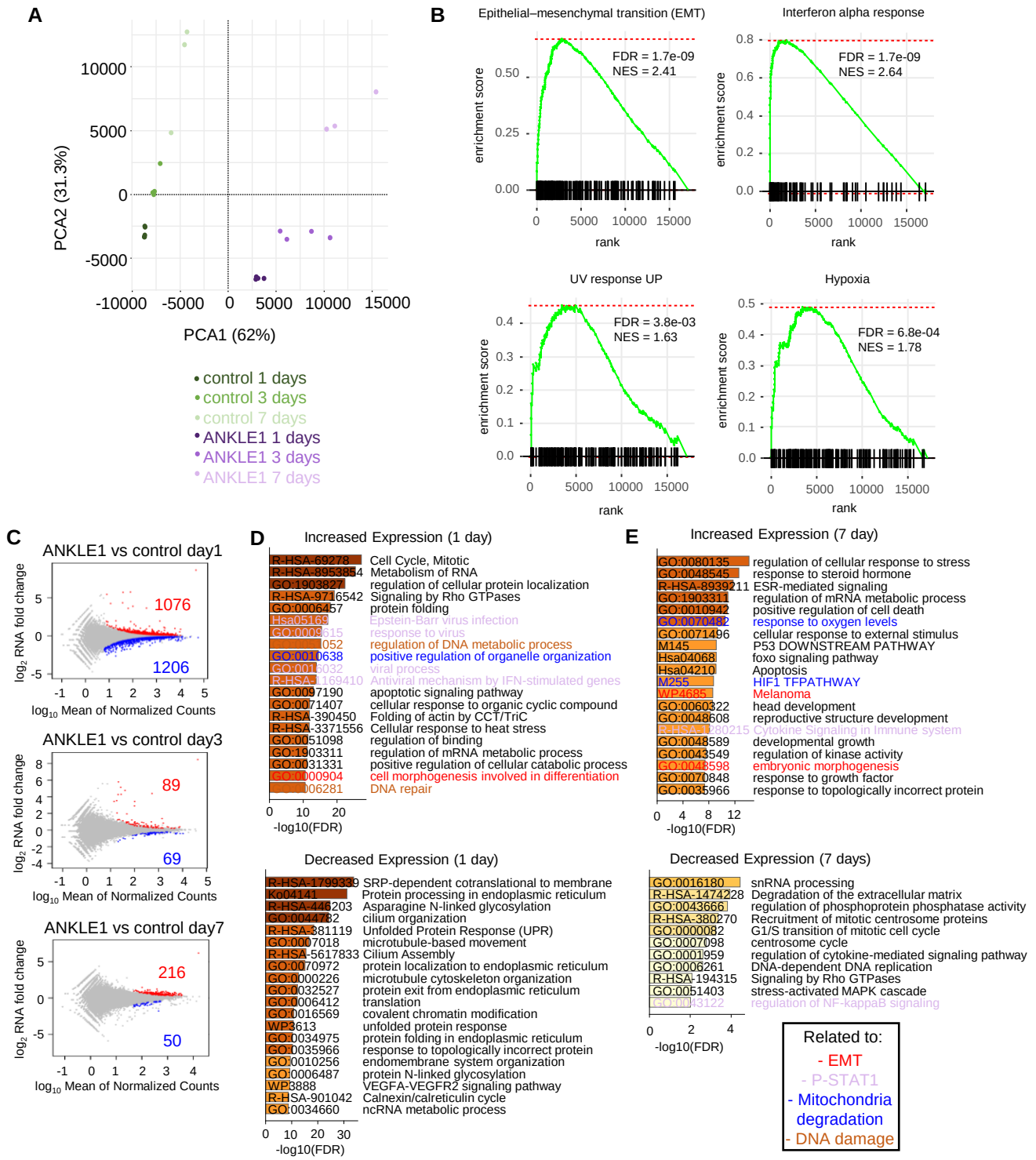

Fig. S4. RNA-seq analysis of all time points indicates that ANKLE1 expression leads to changes in EMT, STAT1 activation, mitochondria degradation, and DNA damage. A) Principal component analysis of the RNA-seq data illustrates that biological replicates cluster together. B) GSEA plots show enrichment of genes engaged in Epithelial-Mesenchymal transition, Interferon alpha response, UV response UP, and Hypoxia. C) These plots compare the change in expression of all genes upon *ANKLE1* transfection for the indicated times. Red genes are activated upon ANKLE1 expression, blue genes are repressed, and light-gray points represent all other genes. The numbers of activated and repressed genes are indicated by text with the corresponding color. D-E) Gene ontology analysis for differentially expressed genes (Figure S4C) for the 3 day (D) and 7 day (E) time points after transfection. ANKLE1 overexpressing cells versus control cells show enrichment for STAT1 activation DNA damage and mitochondria degradation.

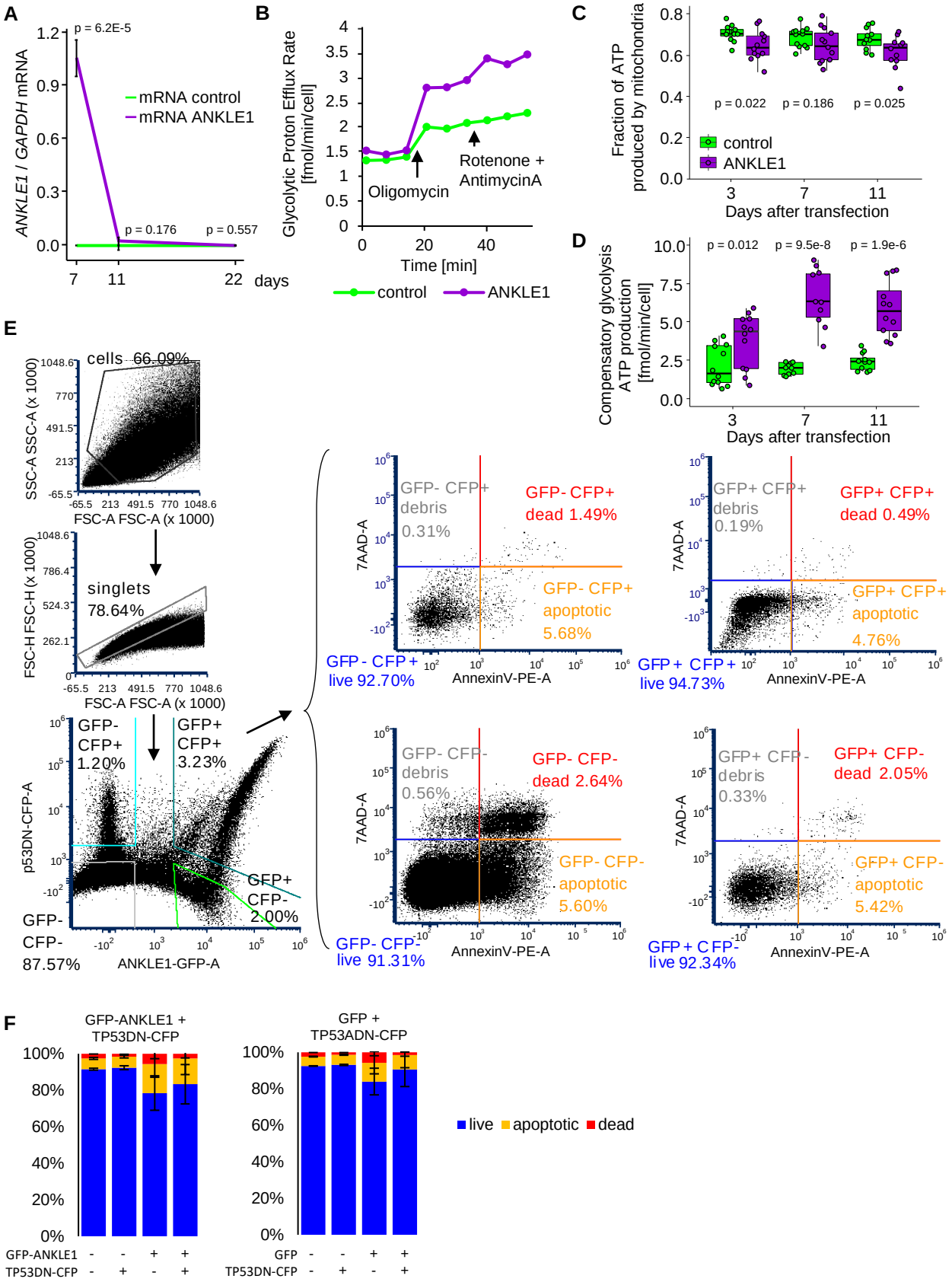

Fig. S5. ANKLE1 induces mitophagy, which results in a change in energy production and apoptosis resistance that is dependent upon TP53 inactivation. A) *ANKLE1* mRNA recovers to pre-transfection levels 11 days after ANKLE1 overexpression. B) An example ATP-rate assay for control and ANKLE1-overexpressing cells shows that glycolytic proton efflux rate is higher for ANKLE1. C) ANKLE1 decreases the fraction of ATP produced by mitochondria in HEK293T cells. D) ANKLE1 increases maximum compensatory glycolysis in HEK293T (p-values are calculated with a two-tailed t-test). E) An example gating strategy isolates GFP-ANKLE1 (or control) and CFP-TP53 to assess the level of apoptosis in cells overexpressing ANKLE1 and/or a TP53 dominant negative mutant. F) The level of apoptosis in HEK293T is not affected by ANKLE1 overexpression, presumably because TP53 is already inactive in HEK293T cells.

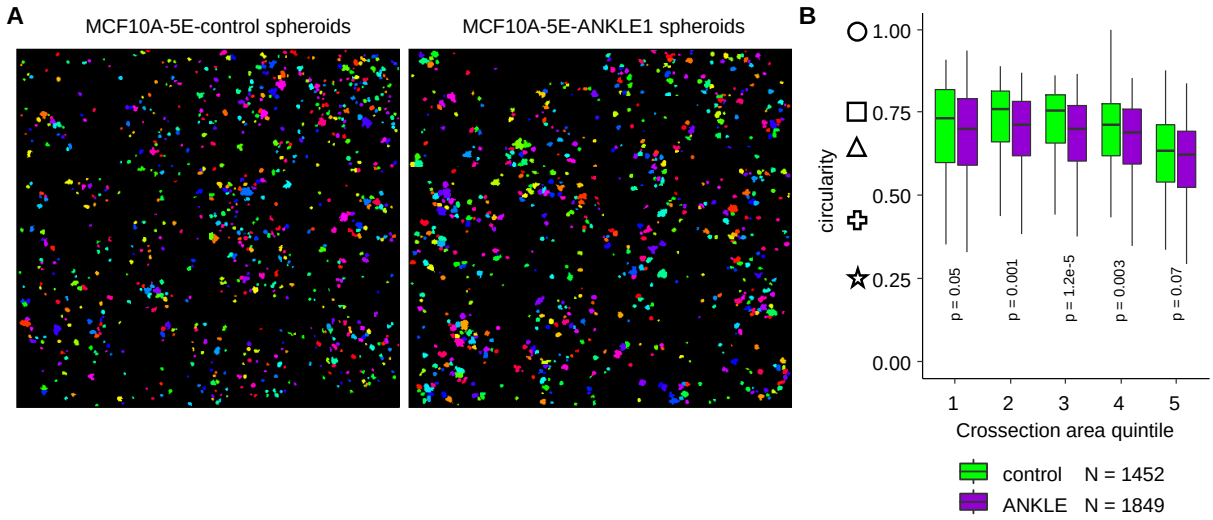

Fig. S6. ANKLE1 induces distortion of MCF10A-5E spheroids, irrespective of spheroid size. A) Images that were quantified by OrganoSeg software are distinguishable as individual spheroids. B) We separated spheroids into quintiles based upon their circularity. Note that quintiles are defined by the pooling of control and ANKLE1 cells, so the ranges of adjacent box plots are identical. These results confirm that while larger spheroids tend to be less circular, lower circularity among ANKLE1-expressing spheroids is not driven by their larger size (p-values are calculated with a two-tailed t-test).