Supplementary Information

1. Supplementary Tables

2. Supplementary Figures

1. Supplementary Tables.

Supplementary Table S1: Patient characteristics and patient genomic analysis (WES and SNP 6.0)

Supplementary Table S2: Leukemia-initiating cell frequencies of paired diagnosis and relapse patient samples

Supplementary Table S3: Targeted-sequencing analysis of PDX and PairTree predicted mutational population clusters

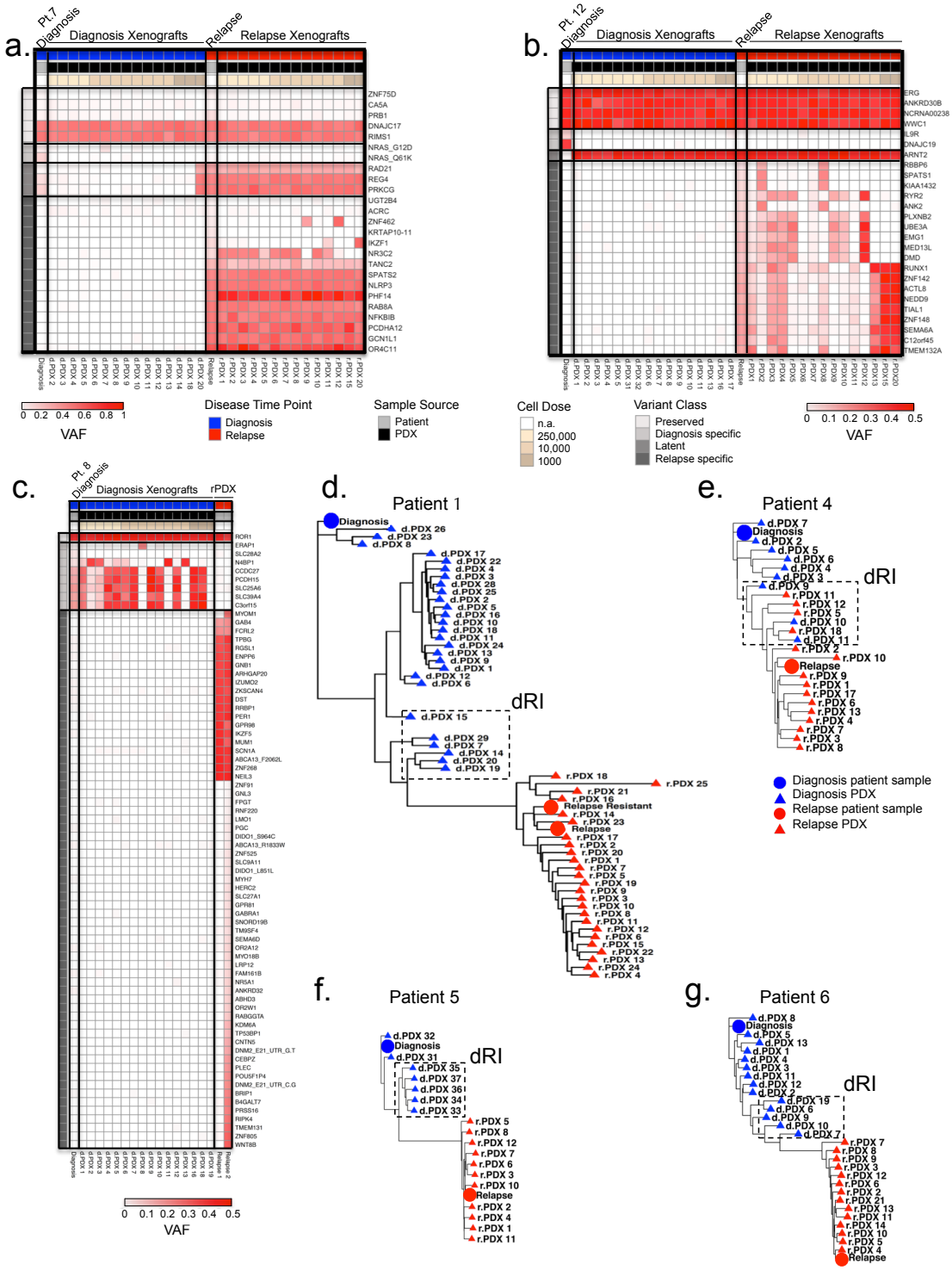
Supplementary Table S4: RNA-sequencing data from Patient 9 PDX

Supplementary Table S5: PDX targeted-sequencing tissue concordance

Supplementary Table S6: dPDX and dRI-PDX leukemia-initiating cell frequencies

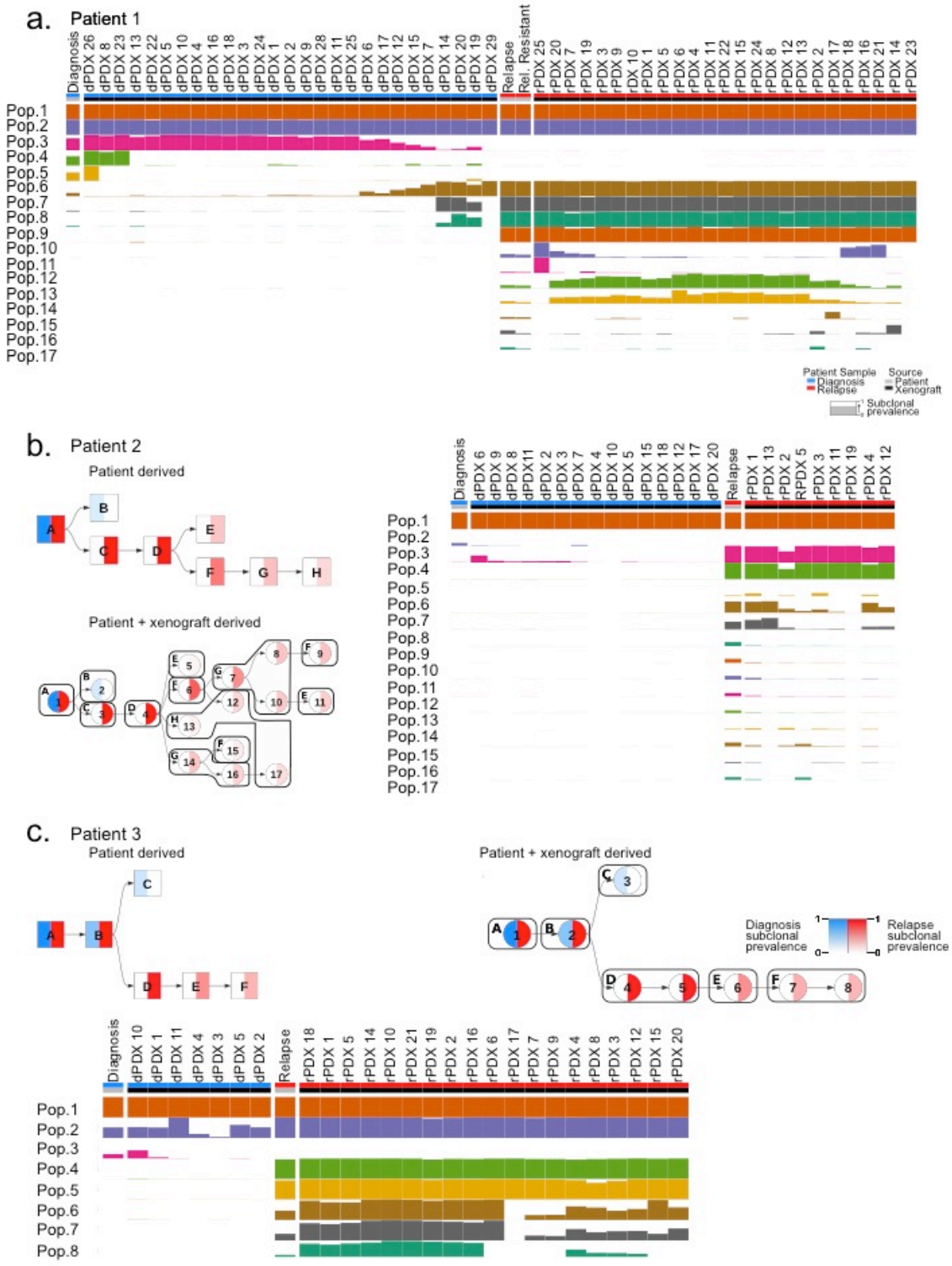
Supplementary Table S7: RNA-sequencing, pathway enrichment (GSEA) reports and GSVA results (including gene list HSC vs B) of PDX and paired patient samples

2. Supplementary Data Figures



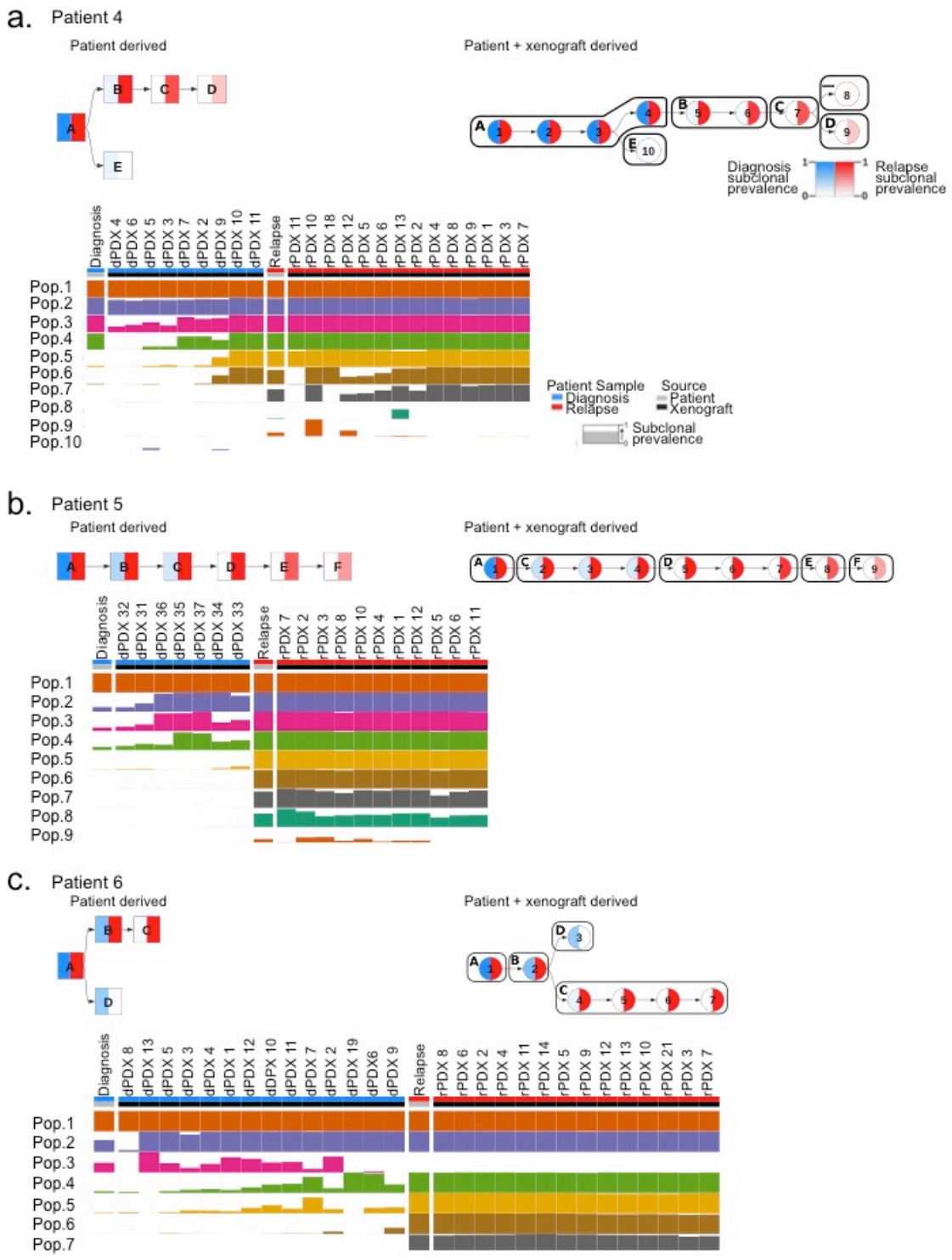
Supplementary Figure S1: PDXs present with clonal skewing towards clinically relevant relapse clones

a-c. Heatmaps of VAF of the leukemic SNV and indel variants identified by whole exome sequencing in diagnosis (top blue bar) and relapse (top red bar) patient samples and PDX. Variants are clustered as preserved (present in diagnosis and relapse patient samples), diagnosis specific (present in diagnosis patient sample and absent in relapse patient sample (VAF < 0.1)), latent (present in diagnosis patient sample with VAF < 0.3 and expanding in relapse sample), and relapse specific (present in relapse patient sample and absent in diagnosis patient sample (VAF < 0.1)). PDX are ordered in decreasing numbers of transplanted cell doses. Despite the presence of diagnosis specific variants in patient samples these variants were not present in diagnosis PDXs for patient 7 (**a**) and patient 12 (**b**). Diagnosis specific variants were detected in patient 8 (**c)**. **d-g**. Phylogenetic analysis showing clonal relationships of xenografts based on VAF of leukemic variants for patients 1,4,5 and 6. The distance between symbols on the trees represent the degrees of relation between them (Minkowski's distance) estimated by a nearest neighbour joining method. Circles represent patient samples and triangles represent PDX; blue represents diagnosis and red represents relapse. Diagnosis clones on the trajectory to relapse were termed diagnosis relapse initiating clones or dRI and are indicated by a box with a hatched border.



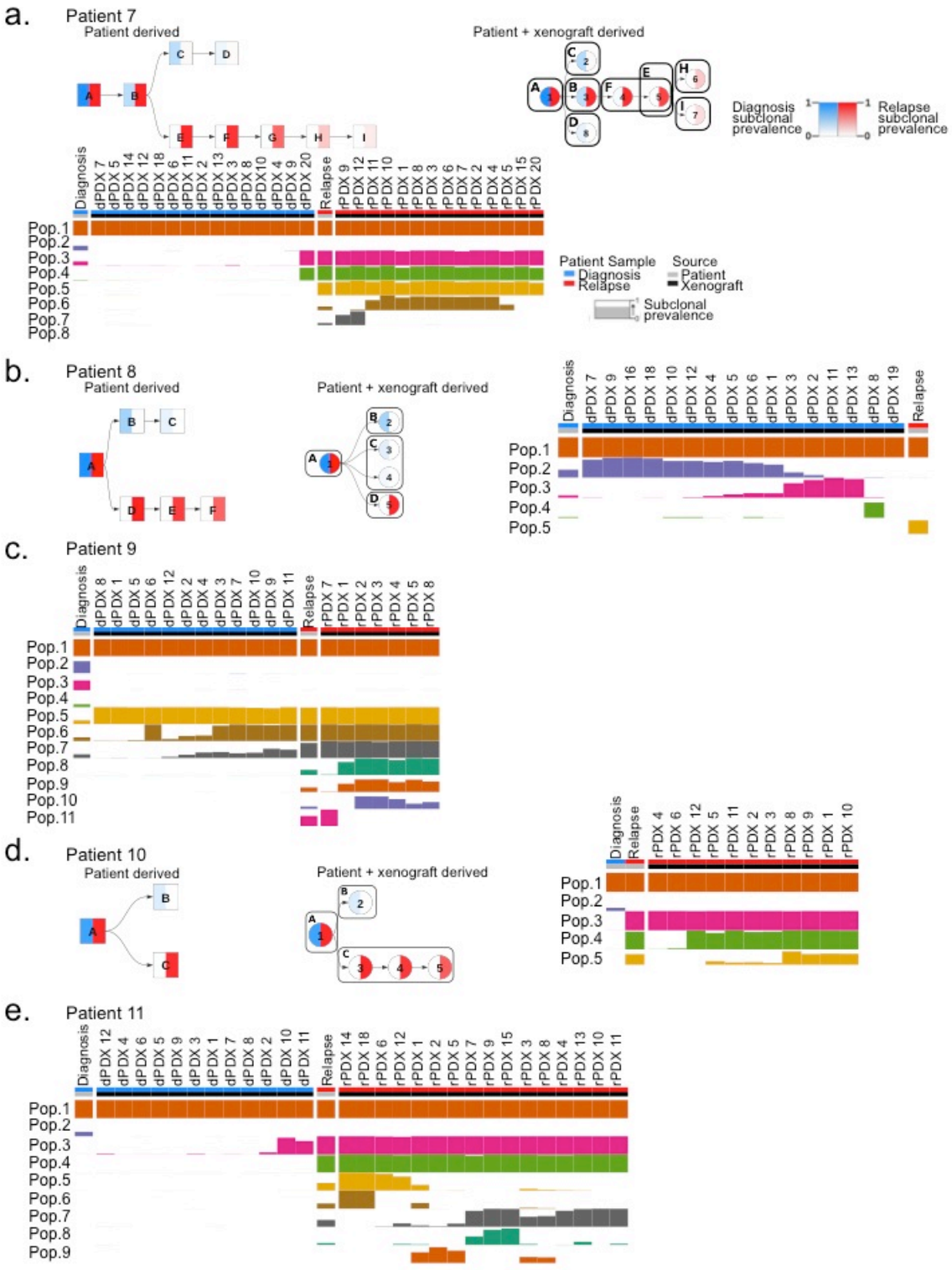
Supplementary Figure S2: Mutational trees derived from patient samples and combined patient samples and xenograft genomic analysis for patients 1-3.

Mutational trees of variants clustered to form populations using the PairTree algorithm. Nodes in mutational trees are divided in half, with the intensity of blue in the left half indicating the frequency of the population's variants at diagnosis, and the intensity of red in the right half showing the frequency of the population's variants at relapse. Mutational populations derived from combine patient and xenograft analysis are indicated by numerically labeled nodes. Mutational populations derived from patient samples alone are labeled alphabetically. Presence of identified mutational populations in patient samples and representative xenografts are displayed in mutational populations tables. Mutational populations (Pop.) are displayed on the y-axis and individual patient samples or xenografts are displayed on the x-axis. The height of the population bar represents the prevalence of the lineage in the patient sample or xenograft. **a**. Complete mutational population table for patient 1 patient samples and xenografts. **b**-**c**. Mutational trees of patients 2 and 3 respectively, uncovered by analysis of patient samples alone or combined analysis of patient samples and xenografts. Mutational population tables for the presence of identified mutational populations in patient samples and representative xenografts of patients are displayed. Mutational populations derived from combine patient and xenograft analysis are indicated by numerically labeled nodes. Mutational populations derived from patient samples alone are labeled alphabetically. Legends in **a** and **c** are applicable to entire figure.



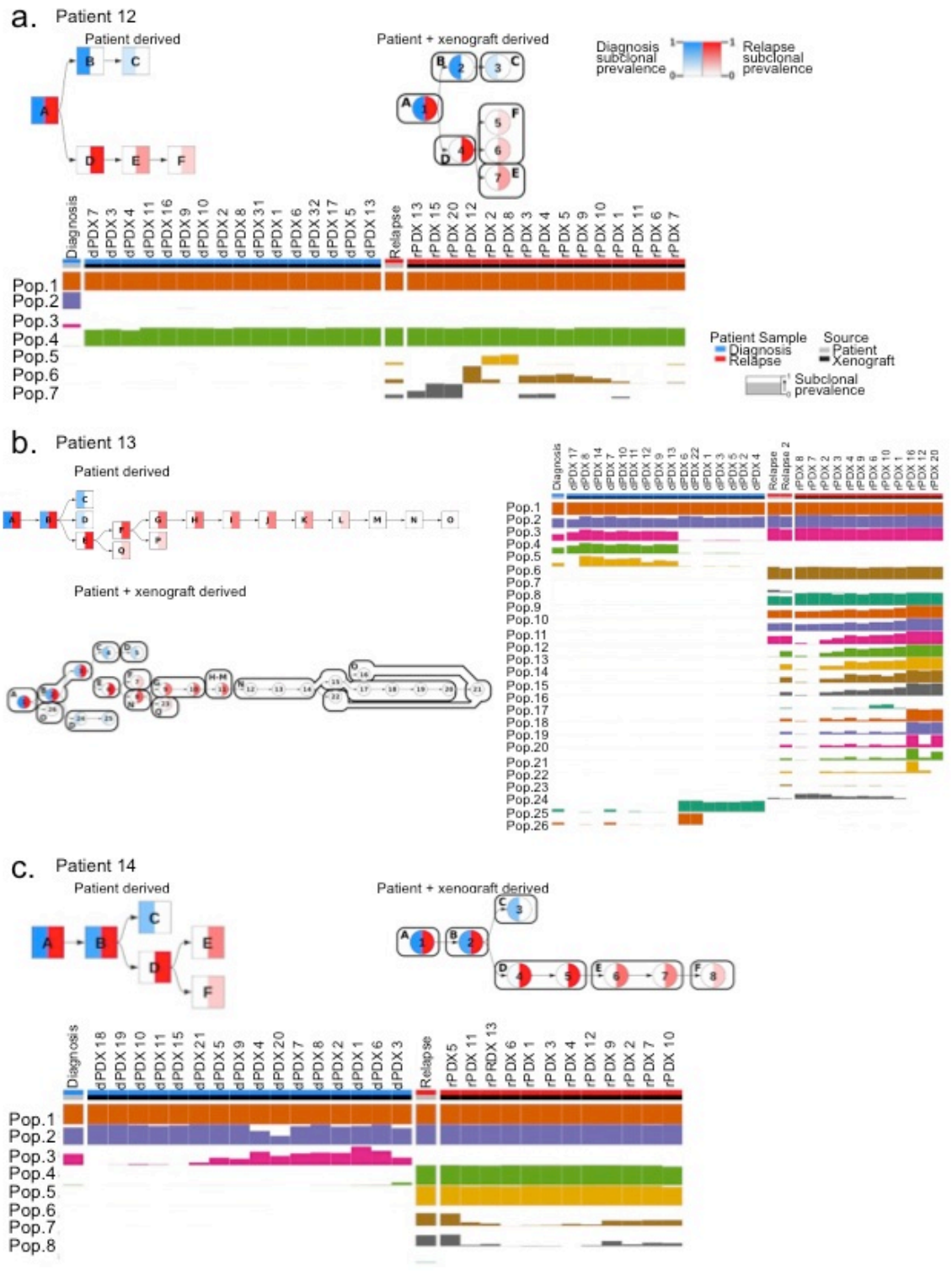
Supplementary Figure S3: Mutational trees derived from patient samples and combined patient samples and xenograft genomic analysis for patients 4-6.

Mutational trees of variants clustered to form populations using the PairTree algorithm. Nodes in mutational trees are divided in half, with the intensity of blue in the left half indicating the frequency of the population's variants at diagnosis, and the intensity of red in the right half showing the frequency of the population's variants at relapse. Mutational populations derived from combine patient and xenograft analysis are indicated by numerically labeled nodes. Mutational populations derived from patient samples alone are labeled alphabetically. Presence of identified mutational populations in patient samples and representative xenografts are displayed in mutational populations tables. Mutational populations (Pop.) are displayed on the y-axis and individual patient samples or xenografts are displayed on the x-axis. The height of the population bar represents the prevalence of the lineage in the patient sample or xenograft. **a**-**c**. Mutational trees of patients 4,5 and 6 respectively, uncovered by analysis of patient samples alone and combined mutational trees derived using patient samples and xenografts. Mutational population tables for the presence of identified mutational populations in patient samples and representative xenografts of patients are displayed. Legends in **a** are applicable to entire figure.



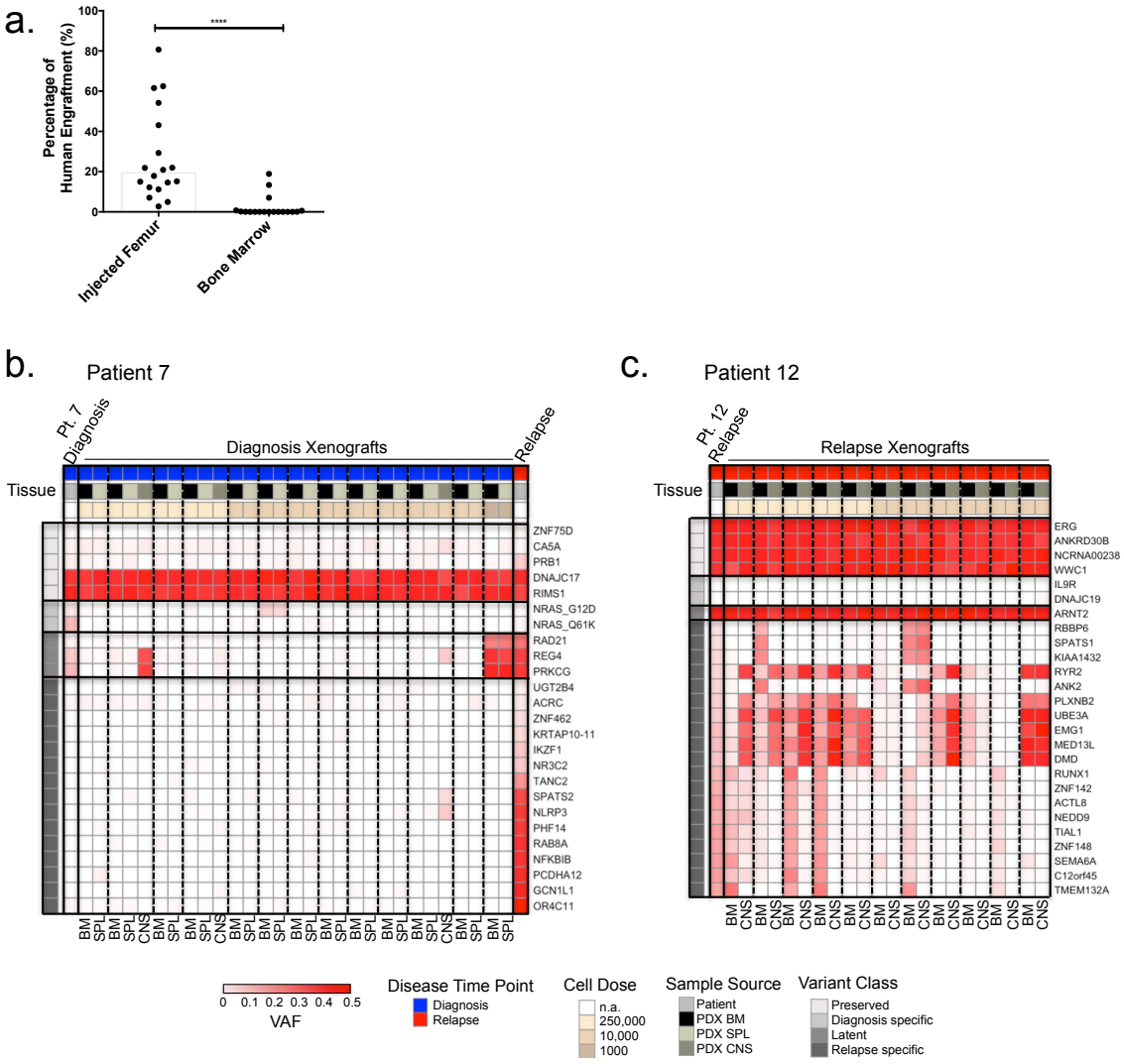
Supplementary Figure S4: Mutational trees derived from patient samples and combined patient samples and xenograft genomic analysis for patients 7-11.

Mutational trees of variants clustered to form populations using the PairTree algorithm. Nodes in mutational trees are divided in half, with the intensity of blue in the left half indicating the frequency of the population's variants at diagnosis, and the intensity of red in the right half showing the frequency of the population's variants at relapse. Mutational populations derived from combine patient and xenograft analysis are indicated by numerically labeled nodes. Mutational populations derived from patient samples alone are labeled alphabetically. Presence of identified mutational populations in patient samples and representative xenografts are displayed in mutational populations tables. Mutational populations (Pop.) are displayed on the y-axis and individual patient samples or xenografts are displayed on the x-axis. The height of the population bar represents the prevalence of the lineage in the patient sample or xenograft. **a**.**b**.**d**. Mutational trees of patients 7,8 and 10 respectively, uncovered by analysis of patient samples alone and combined mutational trees derived using patient samples and xenografts. **a-e** Mutational population tables for the presence of identified mutational populations in patient samples (7-11) and representative xenografts of patients are displayed. Legends in **a** are applicable to entire figure.



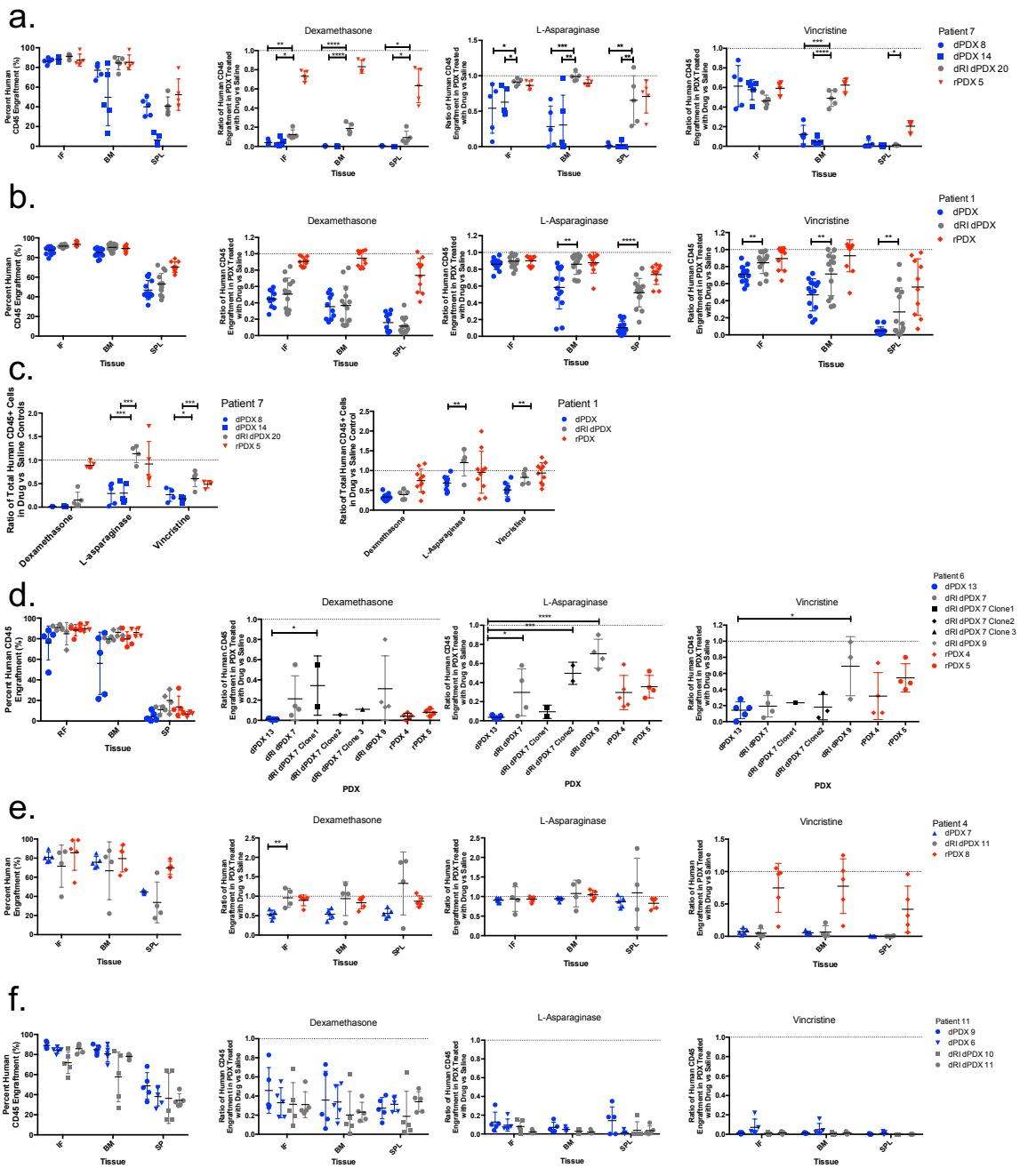
Supplementary Figure S5: Mutational trees derived from patient samples and combined patient samples and xenograft genomic analysis for patients 12-14.

Mutational trees of variants clustered to form populations using the PairTree algorithm. Nodes in mutational trees are divided in half, with the intensity of blue in the left half indicating the frequency of the population's variants at diagnosis, and the intensity of red in the right half showing the frequency of the population's variants at relapse. Mutational populations derived from combine patient and xenograft analysis are indicated by numerically labeled nodes. Mutational populations derived from patient samples alone are labeled alphabetically. Presence of identified mutational populations in patient samples and representative xenografts are displayed in mutational populations tables. Mutational populations (Pop.) are displayed on the y-axis and individual patient samples or xenografts are displayed on the x-axis. The height of the population bar represents the prevalence of the lineage in the patient sample or xenograft. **a**-**c.** Mutational trees of patients 12,13 and 14 respectively, uncovered by analysis of patient samples alone and combined mutational trees derived using patient samples and xenografts. Mutational population tables for the presence of identified mutational populations in patient samples and representative xenografts of patients are displayed. Legends in **a** are applicable to entire figure.



Supplementary Figure S6: Clonal discordance in PDX and PDX tissues.

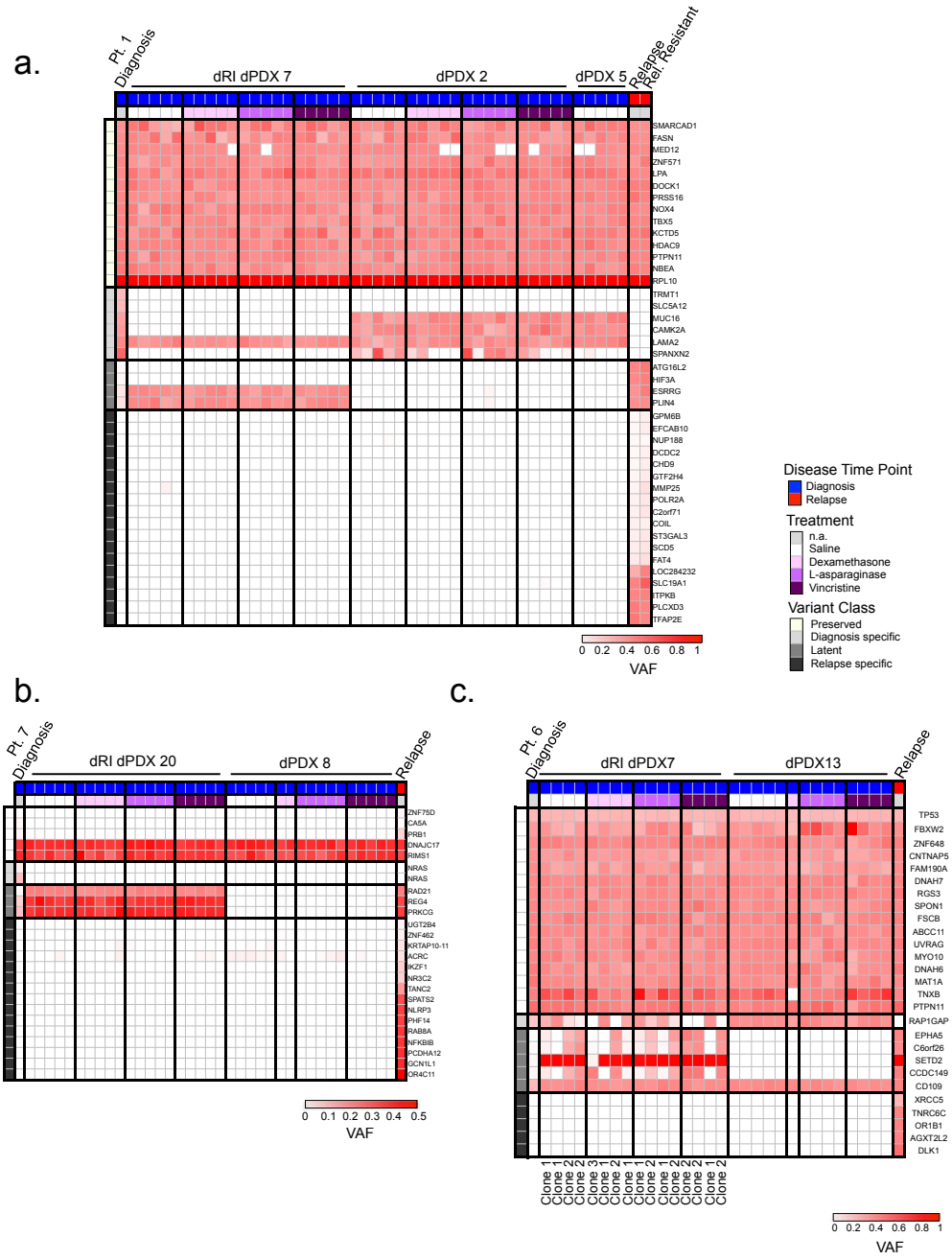
a. Human purified cells from primary dRI-PDX from patient 9 were transplanted into secondary NSG recipients. Mice were analyzed for human chimerism by flow cytometry in the IF and BM. A significantly lower level of human chimerism was observed in the BM in comparison to the IF. Bars represent median. **** $p < 0.0001$, unpaired two-tailed t-test. **b** and **c**. Paired diagnosis and relapse patient samples were transplanted intrafemorally into 30 irradiated NSG mice each in a limiting dilution assay. Mice were sacrificed 20-30 weeks post transplant and their engraftment was assessed by flow cytometry in the bone marrow (BM), spleen (SPL) and central nervous system (CNS). Heatmaps of VAF of the SNV and indel leukemic variants identified by whole exome sequencing in BM, SPL and CNS of patient 7 (**c**) and patient 12 (**d**).



Supplementary Figure S7: Functional heterogeneity of dRI-PDX

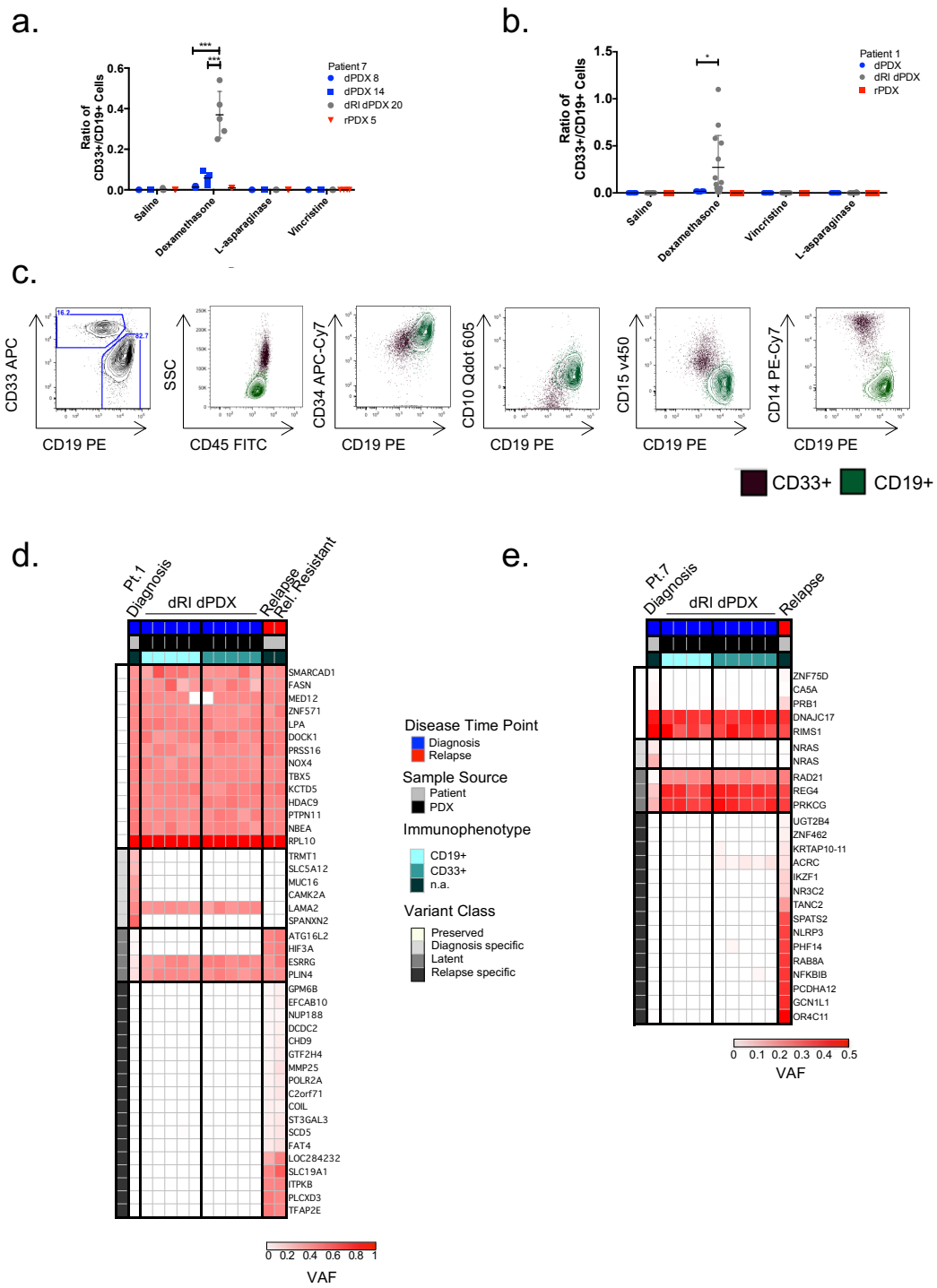
Purified human cells from primary dPDX, dRI-PDX and rPDX (representative relapse genetics) xenografts were injected into secondary NSG mice. Mice were randomized into 4 groups (with 4-5 mice per group) and treated with saline, dexamethasone (DEX),

L-asparagine (L-ASP) or vincristine (VIN). After 4 weeks of treatment mice were sacrificed and engraftment in the IF, BM and SPL were analyzed by flow cytometry. **a-b** Response of PDX to DEX, L-ASP and VIN as determined by the ratio of human engraftment in drug treated PDX in comparison to saline treated controls for patient 7,1, and 4 respectively. For patient 1 (**b**) dPDX represents 3 primary dPDX for L-ASP and VIN (#s 2,5 and 18, n=14 mice) and 2 primary dPDX for DEX (#'s 2 and 5, n=10 mice). dRI-PDX represents 3 primary dRI-PDX (#'s 7,14 and 15, n=13 mice for VIN and n=14 mice for DEX and L-ASP). rPDX represents 2 primary rPDX (#2 and 22, and n=10 mice in each treated group). **c.** Response of total human CD45 cell numbers in PDX to DEX, L-ASP and VIN as determined by the ratio of total human CD45 cells in the IF and BM of drug treated PDX in comparison to saline treated controls for the PDX of patients 7 and 1 as shown in Figure 6 a and b. **d**. Response of patient 6 PDX to DEX, L-ASP and VIN as determine by the ratio of human engraftment in drug treated PDX in comparison to saline treated controls. Response of dRI-PDX 7 secondary recipients are shown grouped or separated based on their clonal composition **e.** Response of PDX to DEX, L-ASP and VIN as determined by the ratio of human engraftment in drug treated PDX in comparison to saline treated controls for patient 4. **f**. Response of patient 11 PDX to DEX, L-ASP and VIN as determined by the ratio of human engraftment in drug treated PDX in comparison to saline treated controls. Bars represent mean and standard deviation. Only significance between dPDX and dRI-PDX are shown. * $p < 0.5$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, unpaired two-sided t-tests.



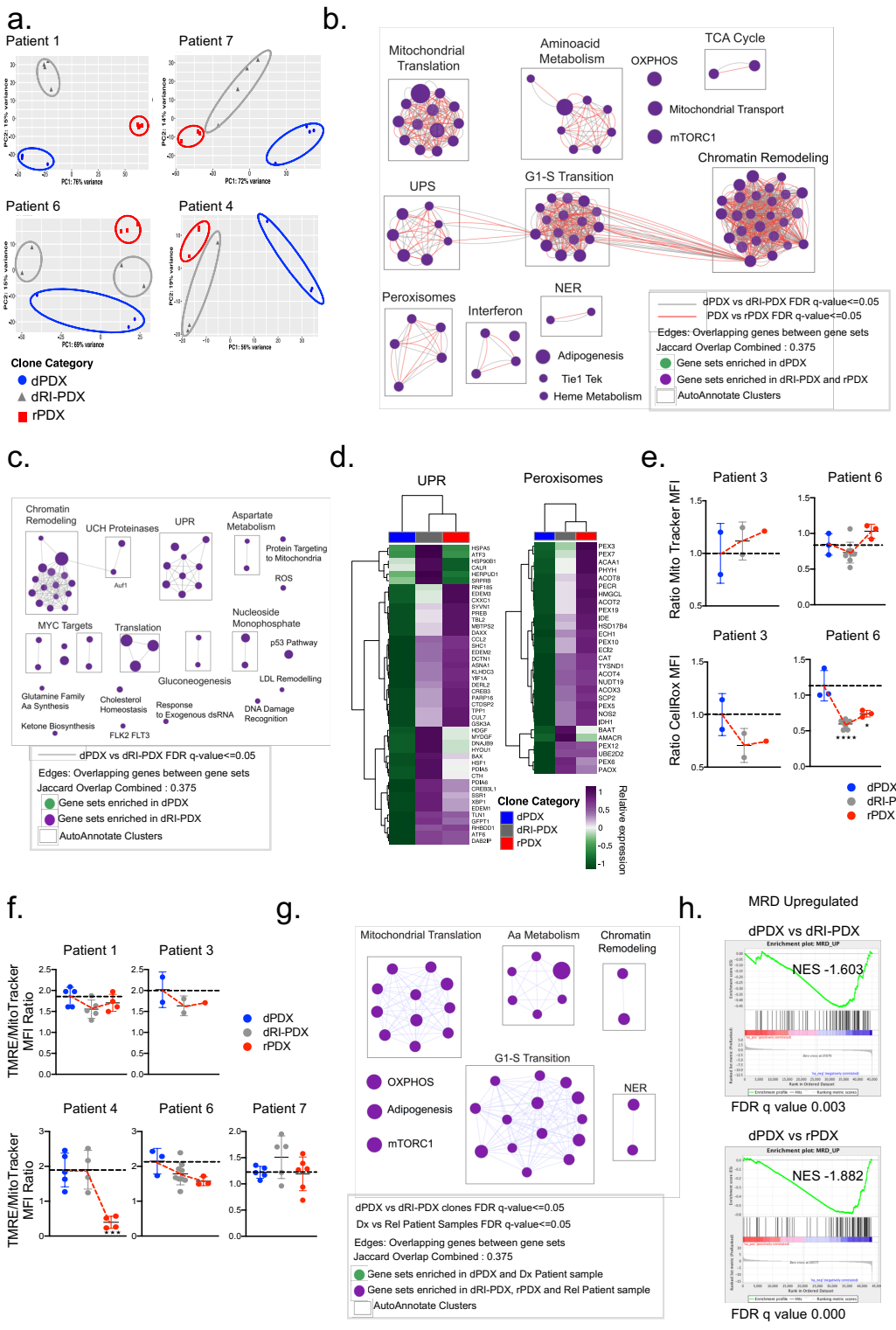
Supplementary Figure S8: Genetic stability in treated secondary PDX

Human cells from secondary PDX treated with saline, dexamethasone, L-asparaginase or vincristine were purified from the BM of mice and subjected to targeted sequencing. Heatmaps display the VAF of leukemic variants previously profiled in the patient samples and primary PDX for patients 1 (**a**), patient 7 (**b**) and patient 6 (**c**). For patient 6 polyclonality of dRI clones was observed in secondary recipients of dRI-PDX 7.



Supplementary Figure S9: Immunophenotypic plasticity upon dexamethasone challenge

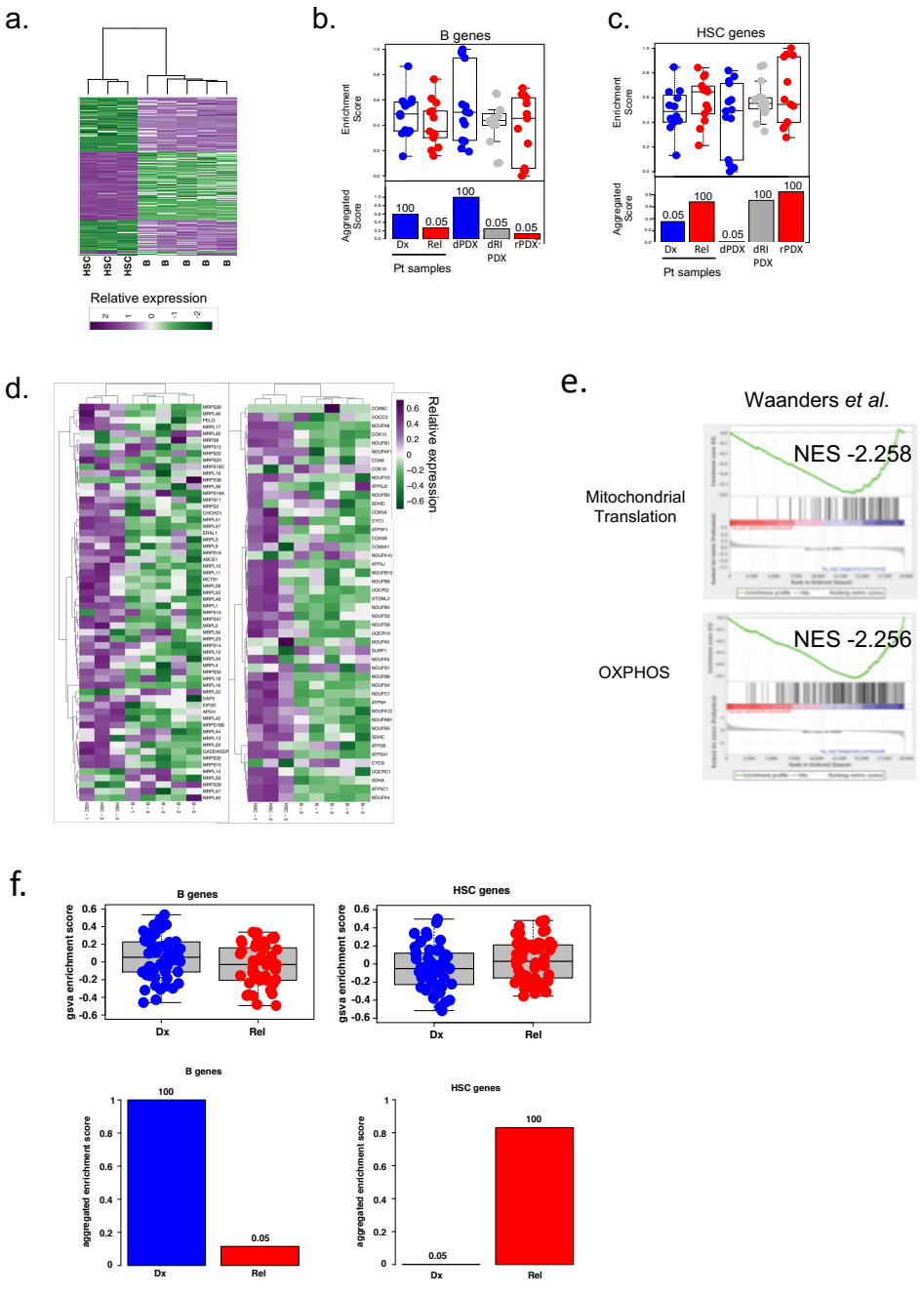
Purified human cells from primary dPDX, dRI-PDX and rPDX (representative relapse PDX) xenografts were injected into secondary NSG mice. Mice were randomized into 4 groups (and treated with either saline, dexamethasone (DEX), L-asparagine (L-ASP) or vincristine (VIN). After 4 weeks of treatment mice were sacrificed and engraftment in the IF, BM and SPL were analyzed by flow cytometry. **a**. and **b**. Cells were analyzed for the presence of $CD33^+$ and $CD19^+$ and the ratio of $CD33^+$ to $CD19^+$ was calculated for patient 7 (n=4-5 mice per group) (**a**) and patient 1 (dPDX represents 2 primary dPDX (#'s 2 and 5, n=10 mice), dRI-PDX represents 3 primary dRI-PDX (#'s 7, 14 and 15, n = 14 mice) and rPDX represent 3 primary rPDX (#'s 2 and 22, n=10 mice) (**b**). **c**. Representative flow cytometry analysis of $CD19^+$ and $CD33^+$ cells in dRI-PDX of patient 7 for lymphoid and myeloid cell surface markers. **d**. and **e**. Fluorescence activated cell sorting was performed to isolate $CD19^+$ and $CD33^+$ cells from secondary dRI-PDX. Heatmaps show results of targeted sequencing for each population for patient 1 and patient 7 respectively. Bars represent mean and standard deviation. Only significance between dPDX and dRI-PDX are shown. * $p < 0.5$, *** $p < 0.001$, unpaired two-sided t-tests.



Supplementary Figure S10: dRI clones are transcriptomic intermediates between diagnosis and relapse clones with a distinct metabolic signature.

a. Principle component analysis (PCA) of the 1000 most variable genes for Patients 1,4, 6 and 7. **b.** Enrichment map of gene sets differentially enriched in dPDX vs dRI-PDX and dPDX vs rPDX (FDR q value ≤ 0.05). Node size is proportional to the number of genes included in each gene set (minimum 10 genes/gene set). Grey and red edges indicate gene overlap. Green node: enrichment in dPDX (positive normalized enrichment score (NES)). Purple node: enrichment in dRI-PDX and rPDX (negative NES). Clusters were automatically annotated using Autoannotate app in Cytoscape (black squares) (UPS: Ubiquitin Proteasome System, OXPHOS: Oxidative Phosphorylation, TCA: Tricarboxylic acid, NER: Nucleotide Excision Repair). **c.** Enrichment map of gene sets differentially enriched in dPDX vs dRI-PDX (FDR q value ≤ 0.05) but not in dPDX vs rPDX (FDR q value ≥ 0.05). Node size is proportional to the number of genes included in each gene set (minimum 10 genes/gene set). Grey edges indicate gene overlap. Green node: enrichment in dPDX (positive NES). Purple node: enrichment in dRI-PDX (negative NES). Clusters were annotated automatically using Autoannotate app in Cytoscape (black squares) (UPR: Unfolded Protein Response, UCH: Ub C-terminal Hydrolase, Aa: Aminoacid, LDL: Low-Density Lipoprotein). **d.** Heatmaps showing the expression of leading-edge genes (GSEA) for selected gene sets from enrichment map in (c). Relative expression was generated from variance stabilized normalized counts. **e.** dPDX, dRI-PDX and rPDX from Patient 3 and Patient 6 were stained with MitoTracker Green and CellROX and analyzed by flow cytometry. MFI for each samples and dye is represented as ratio to dPDX samples for each patient (Patient 3: dPDX n=2, dRI-PDX n=2, rPDX=1; Patient 6: dPDX n=3, dRI-PDX n=9, rPDX=3). **f**. dPDX, dRI PDX and rPDX from Patients 1,3,4,6 and 7 were stained with MitoTracker Green and TMRE dyes and analyzed by flow cytometry. The ratio between TMRE MFI and MitoTracker green MFI was calculated for each sample (Patient 1: dPDX n=5, dRI-PDX n=5, rPDX=4; Patient 3: dPDX n=2, dRI-PDX n=2, rPDX=1; Patient 4: dPDX n=5, dRI-PDX n=4, rPDX=4; Patient 6: dPDX n=3, dRI-PDX n=9, rPDX=3; Patient 7: dPDX n=5, dRI-PDX n=5, rPDX=5). **g**. Enrichment map of gene sets differentially enriched in dPDX vs dRI-PDX (FDR q value ≤ 0.05), dPDX vs rPDX (FDR q value ≥ 0.05) and diagnosis (Dx) vs relapse (Rel) patient samples (FDR q value ≤ 0.05). Node size is proportional to the number of genes included in each gene set (minimum 10 genes/gene set). Purple edges indicate gene overlap. Green node: enrichment in dPDX and diagnosis patient samples (positive NES). Purple node: enrichment in dRI-PDX, rPDX and relapse patient samples (negative NES). Clusters were annotated automatically using Autoannotate app in Cytoscape (black

squares) **h**. From the following comparisons (1) dPDX vs dRI-PDX, (2) dPDX vs rPDX and (3) diagnosis vs relapse patient samples, GSEA enrichment plots were generated for both minimal residual disease (MRD) upregulated and MRD downregulated genes from Ebinger *et al.,* 2016 Cancer Cell (Table S8) (17).



Supplementary Figure S11: dRI and relapse clones share a stem cell profile

a. Heatmap of 2000 differentially expressed genes between B cells and hematopoietic stem cells (HSC) isolated by cell sorting from human umbilical cord blood samples (HSC n=3; B cells n=5). 1000 genes have a higher expression in B cell genes (B genes) and 1000 genes have a higher expression in HSC (HSC genes). **b and c.** Barplot of the aggregated and enrichment GSVA scores for B cell genes (b) and HSC genes (c) in the patient samples used in this study. Aggregated GSVA scores for samples in each category were summed and scaled from 0 to 1. The numbers above the aggregated bars

represent how many times the observed score was higher than random scores obtained in 1000 permutations using a list of 1000 random genes. **d.** Heatmaps showing the expression levels of leading-edge genes (GSEA) Mitochondrial Translation and OXPHOS gene sets in HSC and B cells. **e.** GSEA enrichment plots showing the enrichment of mitochondrial translation and OXPHOS gene sets at relapse in the Waanders *et al.* (30) paired diagnosis and relapse patient cohort. NES scores are indicated (FDR q-value<0.05). **f**. Dotplots and barplots of the aggregated gene set variation analysis (GSVA) scores for B cell genes and HSC genes in the paired diagnosis and relapse patient samples from Waanders *et al.* (30). GSVA scores for samples in each category were summed up and scaled from 0 to 1.