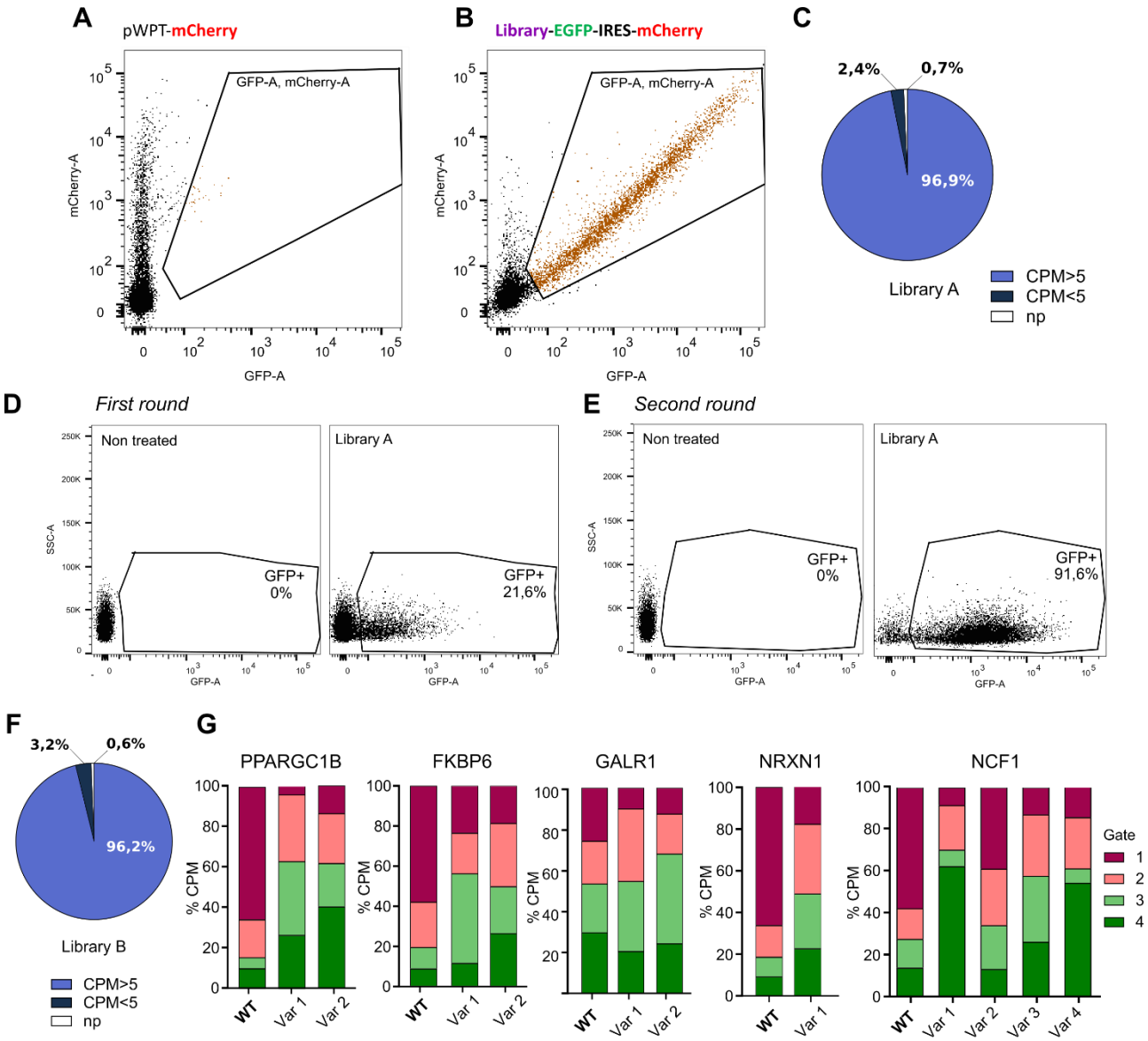
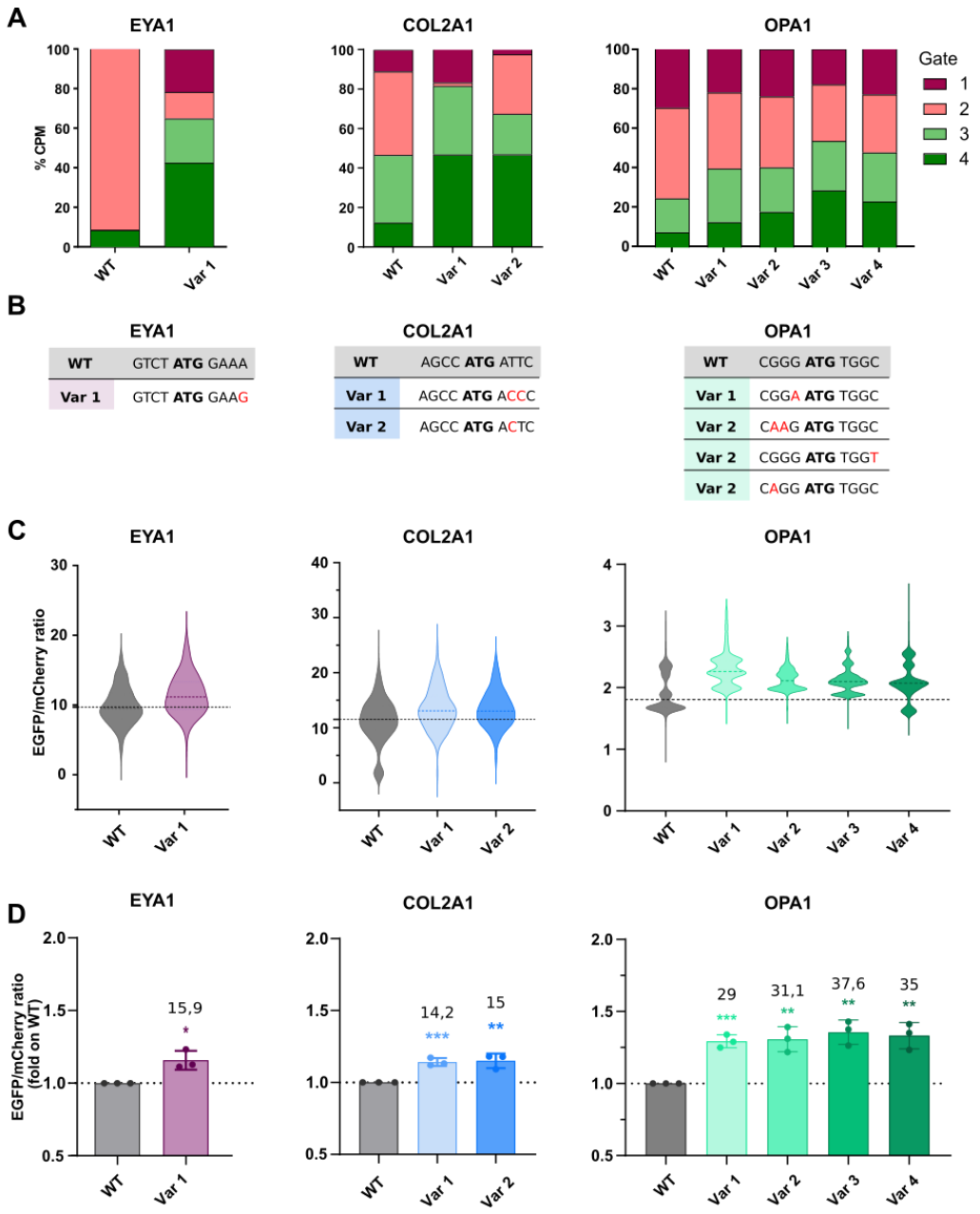
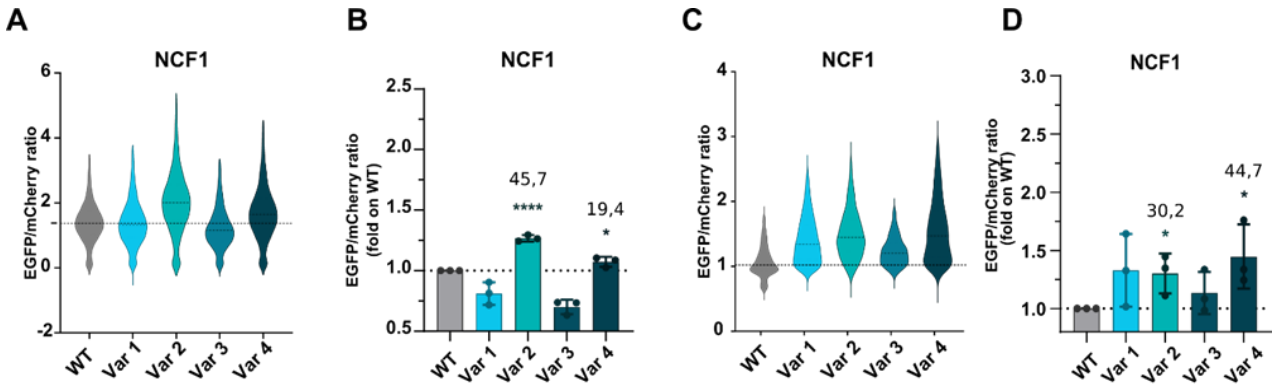


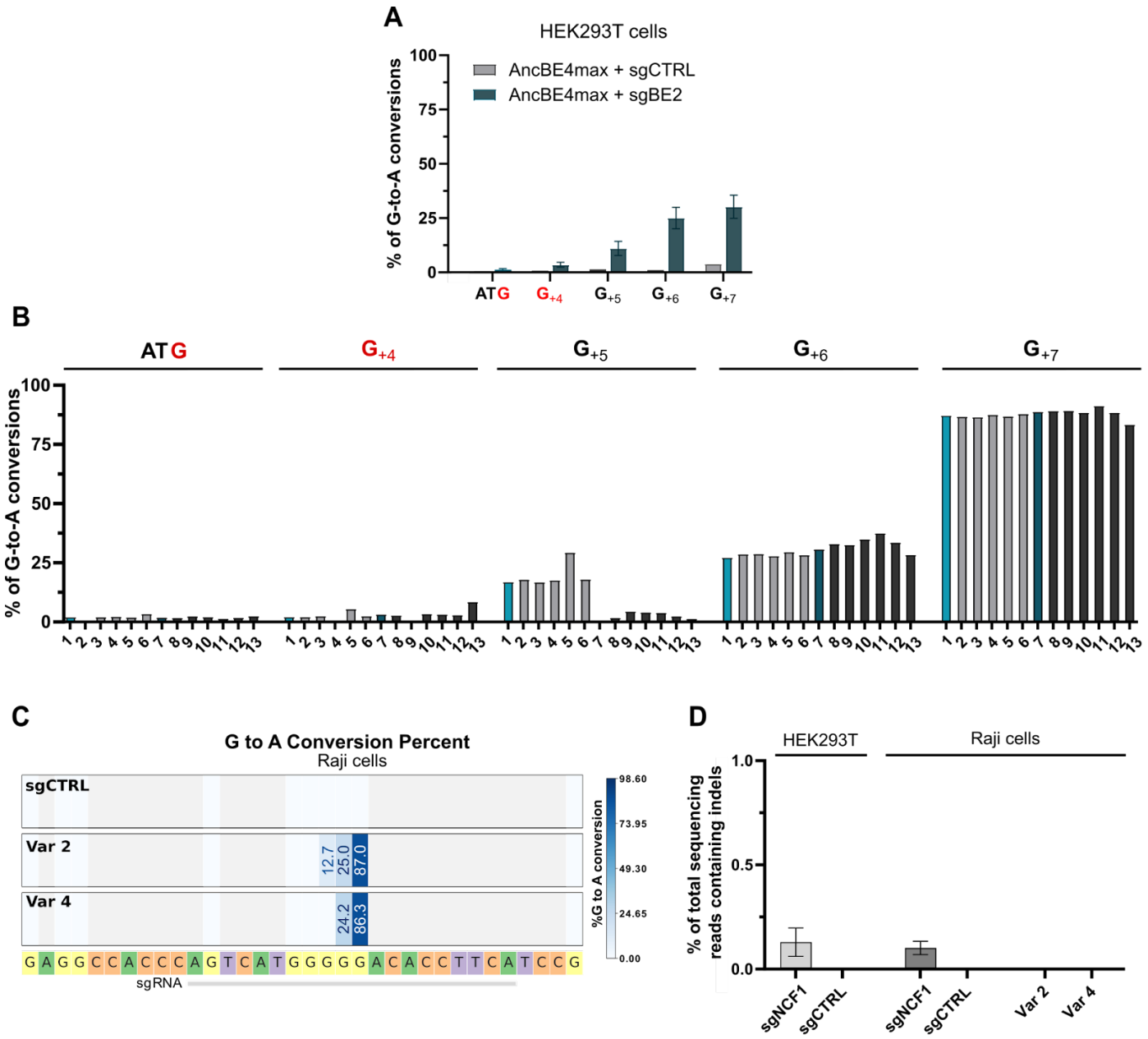
Supplementary Fig. 1. Additional Library A FACS analysis and Library B results. A, B FACS analysis of HEK293T cells transduced with the respective reporter vectors. **C.** Pie charts representing the percentage of sequences identified in the library by deep sequencing. CPM>5: sequences present in the library and well-represented; CPM<5: sequences present but insufficiently represented; np: sequences not present in the library. **D.** EGFP expression of the cells transduced with the Kozak library in the first round of sorting. **E.** EGFP expression of the cells transduced with the Kozak library in the second round of sorting. **F, G.** Library B FACS-seq results. **F.** Pie charts representing the percentage of sequences identified in Library B by deep sequencing, as in C. **G.** Percentage of the count per million reads (CPM) in the 4 gates of the wild type (WT) and the respective variants (Var) of the 5 selected genes in Library B.



Supplementary Fig. 2 Validation of additional actionable hit variants. A. Percentage of the count per million reads (CPM) in the 4 gates of the wild-type (WT) and the respective variants (Var) of the 3 additional selected genes in Library A. **B.** Wild-type (WT) and variants (Var) Kozak sequences of the additional selected hit genes. **C.** Translational enhancement analysed as EGFP/mCherry expression by high content image analysis. The violin plots report the data distribution from n=3 biological replicates. The dashed line indicates the population median. **D.** The histogram represents the mean of the populations analysed by high content image analysis. Data are means ± SD from n=3 biological replicates. The numbers indicate the percentage of mean increase of the variants over the WT. Statistically significant differences were calculated using the unpaired t-test of each variant versus the corresponding WT.



Supplementary Fig. 3 Additional validation of the *NCF1 Kozak* variants. A-B. Validation in HEK293T transduced at low MOI with lentiviral particles of EGFP-IRES-mCherry bearing NCF1 WT, Var 1, Var 2, Var 3, and Var 4 Kozak sequence and analysed 3 days post-transduction by flow cytometry. **A.** Translational enhancement analysed as EGFP/mCherry expression by flow cytometry. The violin plots report the distribution of the data from n=3 biological replicates. The dashed line indicates the population median. **B.** The histograms represent the mean of the populations analysed from n=3 biological replicates analysed by flow cytometry. The numbers indicate the percentage of mean increase of the variants over the WT. **C-D.** NCF1 validation in U2OS cells. **C.** Translational enhancement analysed as EGFP/mCherry expression by high content image analysis. The violin plots report the distribution of the data from n=3 biological replicates. The dashed line indicates the population median. **D.** The histograms represent the median of the populations analysed from n=3 biological replicates analysed by high content image analysis The numbers indicate the percentage of mean increase of the variants over the WT. Statistically significant differences were calculated using the unpaired t-test of each variant versus the corresponding WT.



Supplementary Fig. 4 *NCF1* Kozak sequence base editing. A. Base editing efficiency of NCF1 Kozak sequence in HEK293T cells transiently transfected with AncBE4max and sgNCF1. The percentage of corrected G-to-A conversions (y-axis) is shown for each position in the *NCF1* Kozak sequence (the A of ATG being position +1). **B.** Editing efficiency in the clones (1-13) isolated from the bulk population at target and bystander (in red) guanines, analysed with the EditR software 5 days post-electroporation of AncBE4max and sgNCF1. Clones 1 and 7 were selected for further validation and are referred to in the text as Var 2 and Var 4, respectively. Light grey bars represent clones with the editing pattern reproducing variant 2; dark grey bars represent clones with the editing pattern reproducing variant 4. **C.** Percentage of G-to-A conversion in Var 2 and Var 4 Raji clones as analyzed by targeted deep sequencing of *NCF1* Kozak sequence. **D.** Cas9-mediated indels formation in the target locus, following base editing of NCF1 Kozak sequence, as analyzed by targeted deep sequencing.