Token distribution plot <s> </s> <unk> <pad> <mask> Α Т C G CTT CTC **♦♦ ♦♦ \*\*\* \*\*\* ⊢** GC CG ₩ ₩ CAAG **AGC** GAA CACT **GTGA** CCTGC CGG CCG CCTGG GTTT Weights CTCCC CGT **ACG** TTTC CCAGG **CTCCT GGGAG TCTGC** AAAAA **CCTGA** CCAGC **ACATG TGCTT** CCTTC **TGAGA** CGA **TACCT AGGTA**  $+\Pi$ +AATTA **♦**⊢**□**⊢ **CCTGAG CTTAG** CTGTC CACTC CAGTA **TACTG GTTGG** Tokens