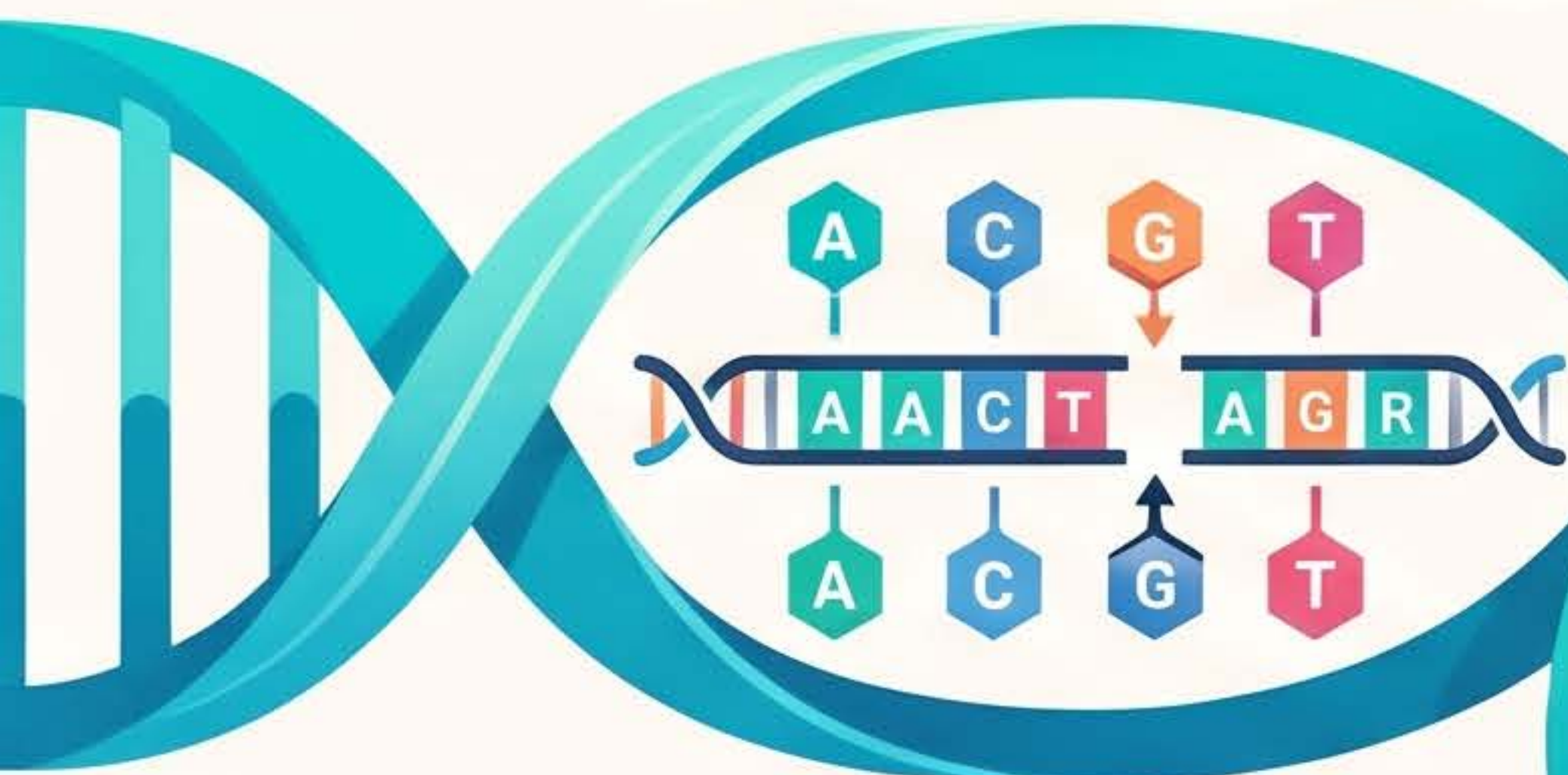


JEPA-DNA：潜在空間での学習が拓く次世代ゲノム解析

JEPA-DNAは、従来のトークン予測と潜在空間予測を組み合わせ、DNAの機能的・意味的な理解を目指すNVIDIAの次世代ゲノム基盤モデルです。

トークンレベル予測による局所学習



配列中の塩基を予測することで、DNAの局所的なパターンを正確に学習します。

【学習目標】トークンレベル
【焦点】局所的な配列（塩基）
【目的】欠損データの復元・パターン認識

潜在空間予測によるグローバルな理解



配列全体の埋め込み表現を潜在空間で予測し、機能的・意味的な理解を深めます。

【学習目標】潜在空間（JEPA）
【焦点】グローバルな表現
【目的】機能的・意味的な情報の獲得

研究ワークフローにおける活用と制限



多彩な評価・抽出タスクに対応
特徴抽出、線形プローブ、配列変化に対するゼロショット評価などをサポートします。



非商用の研究目的限定
Hugging Face上で公開されており、学術・研究目的での利用が想定されています。

DNABERT-2を基盤とした 1.17億パラメータ

既存の強力なモデルを土台に、JEPAの手法を統合して構築されています。