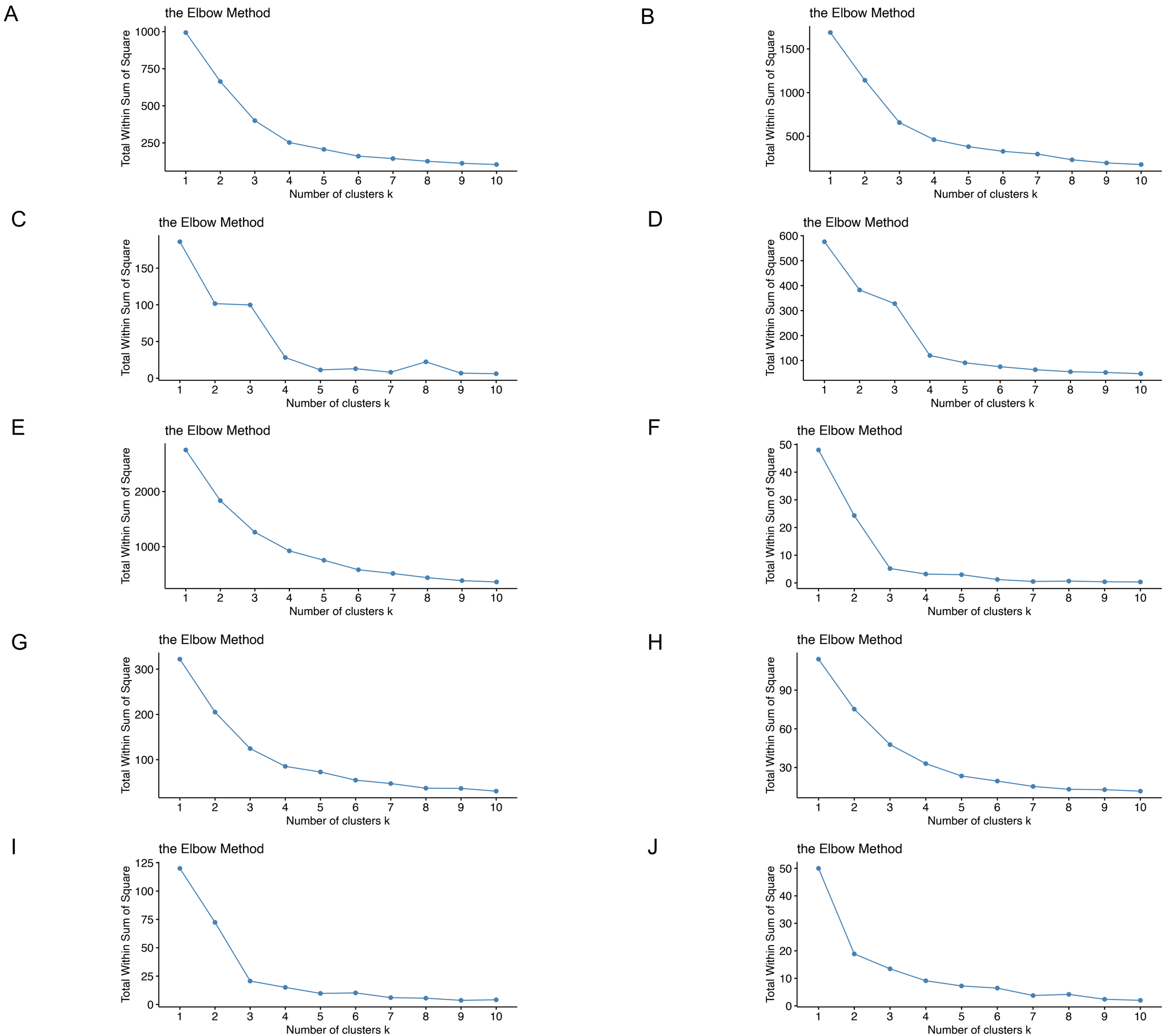

Supplemental Figure 9 Examining optimal K number across the datasets using the Elbow Method **(A)** LeafWorks Inc. dataset **(B)** Phylos Biosciences dataset (n=845) **(C)** Soorni dataset (n=94) **(D)** Medicinal Genomics StrainSEEK V1 (n=289) **(E)** Phylos Biosciences dataset (n=1378) **(F)** Sunrise Genetics (n=25) **(G)** Colorado dataset (n=162) **(H)** Courtagen dataset (n=58) **(I)** Kannapedia 61 dataset (n=61) **(J)** LeafWorks Inc. landrace samples (n=14).