

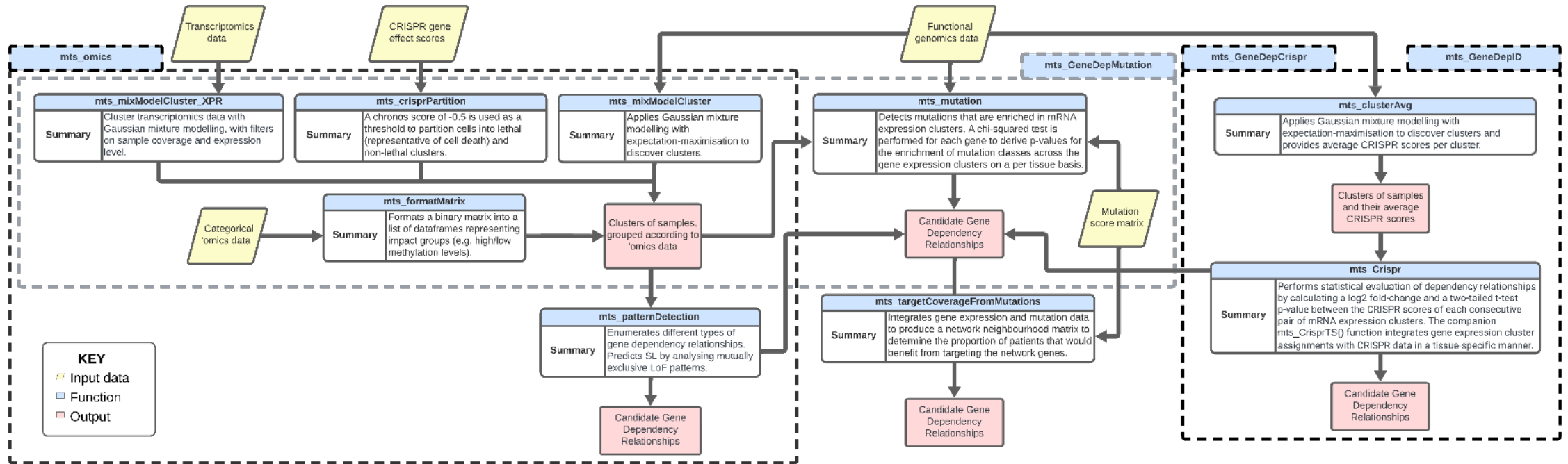


Figure 1: Key MultiSEp Functionality. MultiSEp may be applied to a wide range of data types that contain information about gene functional status in order to predict gene dependency relationships (GDRs), including Synthetic Lethality (SL). Input data (yellow) may be analysed through several different workflows. Dashed boxes outline the wrapper functions, which provide end-to-end pipelines for convenience. A central step is grouping samples to represent different gene function states (clusters); achieved by Gaussian mixture modelling, threshold-based partitioning or from data categories such as mutational status. Analysis of samples' distribution across the functional states for gene pairs reveals patterns that evidence GDRs. For example, SL gene pairs canonically have a mutually exclusive loss of function (LoF) pattern, corresponding to depletion of samples when both genes have LoF. The “Functional genomics data” input (top) could be many different data types (for example, transcriptomics, CRISPR, proteomics, methylation, mutations). Similarly, the “Categorical ‘omics data” input (left-middle) could be multiple data types such as mutations, histone marks, methylation. Common use cases analyse transcriptomics (XPR), CRISPR with XPR, or mutation with XPR and functions are provided for these, such as *mts_mixModelCluster XPR()*.