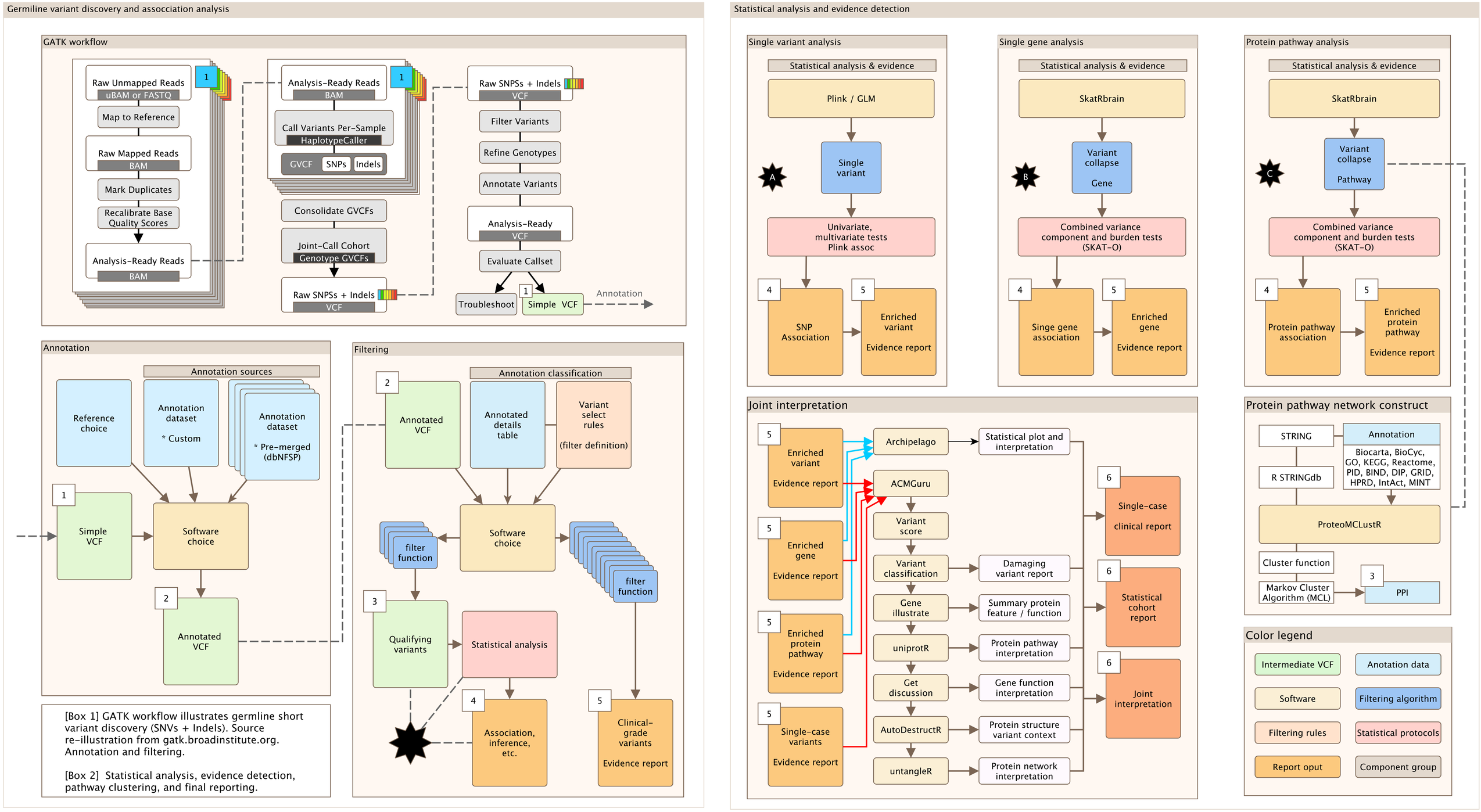
Germiline variant discovery and assocciation analysis
GATK workflow
Raw Unmapped Reads
uBAM or FASTQ
Map to Reference
Raw Mapped Reads
BAM
Mark Duplicates
Recalibrate Base Quality Scores
Analysis-Ready Reads
BAM
1
Analysis-Ready Reads
BAM
Call Variants Per-Sample
HaplotypeCaller
GVCF
SNPs
Indels
Consolidate GVCFs
Joint-Call Cohort
Genotype GVCFs
Raw SNPSs + Indels
VCF
Raw SNPSs + Indels
VCF
Filter Variants
Refine Genotypes
Annotate Variants
Analysis-Ready
VCF
Evaluate Callset
Troubleshoot
Simple VCF
Annotation
Annotation
Annotation sources
Reference choice
Annotation dataset
* Custom
Annotation dataset
* Pre-merged (dbNFSP)
Simple VCF
Software choice
Annotated VCF
Filtering
Annotation classification
Annotated VCF
Annotated details table
Variant select rules
(filter definition)
filter function
Software choice
filter function
Qualifying variants
Statistical analysis
Association, inference, etc.
Clinical-grade variants
Evidence report
[Box 1] GATK workflow illustrates germline short variant discovery (SNVs + Indels). Source re-illustration from gatk.broadinstitute.org. Annotation and filtering.
[Box 2] Statistical analysis, evidence detection, pathway clustering, and final reporting.
Statistical analysis and evidence detection
Single variant analysis
Statistical analysis & evidence
Plink / GLM
Single variant
Univariate, multivariate tests Plink assoc
SNP Association
Enriched variant
Evidence report
Single gene analysis
Statistical analysis & evidence
SkatRbrain
Variant collapse
Gene
Combined variance component and burden tests (SKAT-O)
Singe gene association
Enriched gene
Evidence report
Protein pathway analysis
Statistical analysis & evidence
SkatRbrain
Variant collapse
Pathway
Combined variance component and burden tests (SKAT-O)
Protein pathway association
Enriched protein pathway
Evidence report
Joint interpretation
Enriched variant
Evidence report
Enriched gene
Evidence report
Enriched protein pathway
Evidence report
Single-case variants
Evidence report
Archipelago
ACMGuru
Variant score
Variant classification
Gene illustrate
uniprotR
Get discussion
AutoDestructR
untangleR
Statistical plot and interpretation
Damaging variant report
Summary protein feature / function
Protein pathway interpretation
Gene function interpretation
Protein structure variant context
Protein network interpretation
Single-case clinical report
Statistical cohort report
Joint interpretation
Protein pathway network construct
STRING
R STRINGdb
Annotation
Biocarta, BioCyc, GO, KEGG, Reactome, PID, BIND, DIP, GRID, HPRD, IntAct, MINT
ProteoMCLustR
Cluster function
Markov Cluster Algorithm (MCL)
PPI
Color legend
Intermediate VCF
Anotation data
Software
Filtering algorithm
Filtering rules
Statistical protocols
Report oput
Component group