Gene expression

Data, defined by gene identifiers and associated statistics, are mapped to NCBI Gene unique ids. Imported ids can be any of most standard reference identifiers (e.g., from refseq, ensembl, uniprot, genbank, etc.) and number of commercial platform identifiers (e.g., Illumina beadchip and other array platforms.) Data is ranked on fold change values when present. If fold change is not specified, Correlation Engine ranks on the first recognized statistical column or a user specified rank column. Public data is ranked by fold change.

	Header Name	Notes
Required	gene	Numeric or character gene identifiers. *Alternate names: id, accession number, accession, entrez gene, feature, gene, genbank id, genbank, gene id, gene symbol, gene name, probe set id, probesetid, probe set name, refseq id, refseq, symbol, unigene id, unigene, imported id, protein, protein id, protein name*
Recommended	fold change	Values are unlogged signed +(test/control) or -(control/test)
	fold change (0-n)	Values are unsigned, unlogged test/control will be converted to +/- fold change
	fold change (log2)	Values are signed +log2(test/control) or -log2(control/test) converted to +/- fold change
	p-value	*Alternate names: pvalue, p-value, p value, t-test p-value, t-test*
	Test expression	*Alternate names: average expression, average intensity, expression level, intensity, median expression, test expression, test1 expression, testexpression*
	Control expression	*Alternate names: control expression, control expr, controlexpression*
Optional	Custom	*Custom column header of the users choice for additional statistic data, a total of 5 custom statistical columns can be added. Only the first 7 recommended and custom columns will be displayed in the UI.*

miRNA expression

Data, defined by gene identifiers and associated statistics, are mapped to NCBI Gene unique ids. Imported ids can be any of most standard reference identifiers (e.g., from refseq, ensembl, uniprot, genbank, etc.) and number of commercial platform identifiers (e.g., Illumina beadchip and other array platforms.) Data is ranked on fold change values when present. If fold change is not specified, Correlation Engine ranks on the first recognized statistical column or a user specified rank column. Public data is ranked by fold change.

	Header Name	Notes
Required	gene	Numeric or character gene identifiers. *Alternate names: id, accession number, accession, entrez gene, feature, gene, genbank id, genbank, gene id, gene symbol, gene name, probe set id, probesetid, probe set name, refseq id, refseq, symbol, unigene id, unigene, imported id, protein, protein id, protein name*
Recommended	fold change	Values are unlogged signed +(test/control) or -(control/test)
	fold change (0-n)	Values are unsigned, unlogged test/control will be converted to +/- fold change
	fold change (log2)	Values are signed +log2(test/control) or -log2(control/test) converted to +/- fold change
	p-value	*Alternate names: pvalue, p-value, p value, t-test p-value, t-test*
	Test expression	*Alternate names: average expression, average intensity, expression level, intensity, median expression, test expression, test1 expression, testexpression*
	Control expression	*Alternate names: control expression, control expr, controlexpression*
Optional	Custom	*Custom column header of the users choice for additional statistic data, a total of 5 custom statistical columns can be added. Only the first 7 recommended and custom columns will be displayed in the UI.*

Protein expression

Data, defined by gene identifiers and associated statistics, are mapped to NCBI Gene unique ids. Imported ids can be any of most standard reference identifiers (e.g., from refseq, ensembl, uniprot, genbank, etc.) and number of commercial platform identifiers (e.g., Illumina beadchip and other array platforms.) Data is ranked on fold change values when present. If fold change is not specified, Correlation Engine ranks on the first recognized statistical column or a user specified rank column. Public data is ranked by fold change.

	Header Name	Notes
Required	gene	Numeric or character gene identifiers. *Alternate names: id, accession number, accession, entrez gene, feature, gene, genbank id, genbank, gene id, gene symbol, gene name, probe set id, probesetid, probe set name, refseq id, refseq, symbol, unigene id, unigene, imported id, protein, protein id, protein name*
Recommended	fold change	Values are unlogged signed +(test/control) or -(control/test)
	fold change (0-n)	Values are unsigned, unlogged test/control will be converted to +/- fold change
	fold change (log2)	Values are signed +log2(test/control) or -log2(control/test) converted to +/- fold change
	p-value	*Alternate names: pvalue, p-value, p value, t-test p-value, t-test*
	Test expression	*Alternate names: average expression, average intensity, expression level, intensity, median expression, test expression, test1 expression, testexpression*
	Control expression	*Alternate names: control expression, control expr, controlexpression*
Optional	Custom	*Custom column header of the users choice for additional statistic data, up to a total of 10 custom columns. Only the first 10 recommended and custom statistical columns will be displayed in the UI.*

DNA-methylation

Data is defined by sequence coordinates comprised of chromosome, start, and stop. By default the first statistical column after the coordinates are defined (column 3 or 4) is used for ranking unless a rank column is specified by the user. Public data is typically ranked by a differential statistic, unless context determines otherwise.

	Header Name	Content format
Required	chrid	*Alternate names: chromosome id, chromosome, chromosomeid, chrid, chr id, chr*
	start	*Alternate names: start coordinate, start, start_position, start position, begin, begin coordinate, begin_coordinate, begin_position, beginposition*
	stop	*Alternate names: stop coordinate, stop, end, end coordinate, end_position, end position*
Recommended	% methylated group 1	*Alternate names: % methylated group 1, % methylated 1, % group 1, % meth group 1, % methylated test, % meth test*
	% methylated group 2	*Alternate names: % methylated group 2, % methylated 2, % group 2, % meth group 2, % methylated normals, % meth normals*
	% differential	*Alternate names: % differential, % diff, differential*
	t-statistic	*Alternate names: t-statistic, t statistic, t-stat, t stat*
	p-value	*Alternate names: p-value, p value, pvalue*
	Q-Value	*Alternate names: q-value, p value, qvalue*
Optional	Custom	*Custom column header of the users choice for additional statistic data, up to a total of 10 custom columns. Only the first 10 recommended and custom statistical columns will be displayed in the UI.*

Protein-DNA binding

Data is defined by sequence coordinates comprised of chromosome, start, and stop. By default the first statistical column after the coordinates are defined (column 3 or 4) is used for ranking unless a rank column is specified by the user. Public data is typically ranked by a differential statistic, unless context determines otherwise. Note, log fold change values are recorded as is and not transformed by the

	Header Name	Notes
Required	chrid	*Alternate names: chromosome id, chromosome, chromosomeid, chrid, chr id, chr*
	start	*Alternate names: start coordinate, start, start_position, start position, begin, begin coordinate, begin_coordinate, begin_position, beginposition*
	stop	*Alternate names: stop coordinate, stop, end, end coordinate, end_position, end position*
Recommended	Max Position	*Alternate names: max position, max_position, max pos, max_pos, maxpos, max*
	AUC	*Alternate names: auc, area*
	Fold change	*Alternate names: fold change, fold_change, foldchange, fold change (log2), fold_change (log2), foldchange (log2), fold change (0-n), fold_change (0-n), foldchange (0-n), fold change (+/- mode), fold_change (+/- mode), fold change (log2 mode), fold_change (log2 mode), foldchange (log2 mode), fold change (0-n mode), fold_change (0-n mode), foldchange (0-n mode)*
Optional	Custom	*Custom column header of the users choice for additional statistic data, up to a total of 10 custom columns. Only the first 10 recommended and custom statistical columns will be displayed in the UI.*

Histone methylation binding

Data is defined by sequence coordinates comprised of chromosome, start, and stop. By default the first statistical column after the coordinates are defined (column 3 or 4) is used for ranking unless a rank column is specified by the user. Public data is typically ranked by a differential statistic, unless context determines otherwise. Note, log fold change values are recorded as is and not transformed by the

	Header Name	Notes
Required	chrid	*Alternate names: chromosome id, chromosome, chromosomeid, chrid, chr id, chr*
	start	*Alternate names: start coordinate, start, start_position, start position, begin, begin coordinate, begin_coordinate, begin_position, beginposition*
	stop	*Alternate names: stop coordinate, stop, end, end coordinate, end_position, end position*
Recommended	Max Position	*Alternate names: max position, max_position, max pos, max_pos, maxpos, max*
	AUC	*Alternate names: auc, area*
	Fold change	*Alternate names: fold change, fold_change, foldchange, fold change (log2), fold_change (log2), foldchange (log2), fold change (0-n), fold_change (0-n), foldchange (0-n), fold change (+/- mode), fold_change (+/- mode), fold change (log2 mode), fold_change (log2 mode), foldchange (log2 mode), fold change (0-n mode), fold_change (0-n mode), foldchange (0-n mode)*
Optional	Custom	*Custom column header of the users choice for additional statistic data, up to a total of 10 custom columns. Only the first 10 recommended and custom statistical columns will be displayed in the UI.*

Histone acetylation binding

Data is defined by sequence coordinates comprised of chromosome, start, and stop. By default the first statistical column after the coordinates are defined (column 3 or 4) is used for ranking unless a rank column is specified by the user. Public data is typically ranked by a differential statistic, unless context determines otherwise. Note, log fold change values are recorded as is and not transformed by the

	Header Name	Notes
Required	chrid	*Alternate names: chromosome id, chromosome, chromosomeid, chrid, chr id, chr*
	start	*Alternate names: start coordinate, start, start_position, start position, begin, begin coordinate, begin_coordinate, begin_position, beginposition*
	stop	*Alternate names: stop coordinate, stop, end, end coordinate, end_position, end position*
Recommended	Max Position	*Alternate names: max position, max_position, max pos, max_pos, maxpos, max*
	AUC	*Alternate names: auc, area*
	Fold change	*Alternate names: fold change, fold_change, foldchange, fold change (log2), fold_change (log2), foldchange (log2), fold change (0-n), fold_change (0-n), foldchange (0-n), fold change (+/- mode), fold_change (+/- mode), fold change (log2 mode), fold_change (log2 mode), foldchange (log2 mode), fold change (0-n mode), fold_change (0-n mode), foldchange (0-n mode)*
Optional	Custom	*Custom column header of the users choice for additional statistic data, up to a total of 10 custom columns. Only the first 10 recommended and custom statistical columns will be displayed in the UI.*

ATAC-Seq

Data is defined by sequence coordinates comprised of chromosome, start, and stop. By default the first statistical column after the coordinates are defined (column 3 or 4) is used for ranking unless a rank column is specified by the user. Public data is typically ranked by a differential statistic, unless context determines otherwise. Note, log fold change values are recorded as is and not transformed by the

	Header Name	Notes
Required	chrid	*Alternate names: chromosome id, chromosome, chromosomeid, chrid, chr id, chr*
	start	*Alternate names: start coordinate, start, start_position, start position, begin, begin coordinate, begin_coordinate, begin_position, beginposition*
	stop	*Alternate names: stop coordinate, stop, end, end coordinate, end_position, end position*
Recommended	Test Mean Reads	*Alternate names: test mean reads, test_mean_reads, test reads, test_reads, mean reads, mean_reads*
	Control Mean Reads	*Alternate names: control mean reads, control_mean_reads, control reads, control_reads*
	Mean Total Reads	*Alternate names: mean total reads, mean_total_reads, total reads, total_reads*
	Fold change	*Alternate names: fold change, fold_change, foldchange, fold change (log2), fold_change (log2), foldchange (log2), fold change (0-n), fold_change (0-n), foldchange (0-n), fold change (+/- mode), fold_change (+/- mode), fold change (log2 mode), fold_change (log2 mode), foldchange (log2 mode), fold change (0-n mode), fold_change (0-n mode), foldchange (0-n mode)*
	FDR	*Alternate names: fdr*
	p-value	*Alternate names: p value, p value, pvalue*
Optional	Custom	*Custom column header of the users choice for additional statistic data, up to a total of 10 custom columns. Only the first 10 recommended and custom statistical columns will be displayed in the UI.*

DNA copy number

Data is defined by sequence coordinates comprised of chromosome, start, and stop. By default the first statistical column after the coordinates are defined (column 3 or 4) is used for ranking unless a rank column is specified by the user. Public data is typically ranked by a differential statistic, unless context determines otherwise. Note, log fold change values are recorded as is and not transformed by the

	Header Name	Notes
Required	chrid	*Alternate names: chromosome id, chromosome, chromosomeid, chrid, chr id, chr*
	start	*Alternate names: start coordinate, start, start_position, start position, begin, begin coordinate, begin_coordinate, begin_position, beginposition*
	stop	*Alternate names: stop coordinate, stop, end, end coordinate, end_position, end position*
Recommended	Copy-number change	*Alternate names: copy number change, copy_number chg, copy-number change, copy-number chg, copynumberchange, cnchange, cnvchange*
	Segment Mean	*Alternate names: segment mean, segment_mean, segmean, seg_mean, sm,*
	Z-Score	*Alternate names: z-score*
	Samples with gain	*Alternate names: sample with gain, samples with gain, gain_samples, #samples +, #samples+, #sample +, #sample+*
	Samples with loss	*Alternate names: sample with loss, samples with loss, loss samples, loss_samples, #samples-, #samples -, #sample-, #sample -*
	Number of normal samples	*Alternate names: normal samples, normalsamples*
	Number of probes	*Alternate names: probes, number_of_probes, probe count, probe_count, number of probes*
	Aberration type	Example, Gain or Loss. *Alternate names: aberration type, aberration_type, aberration*
	Fold change	*Alternate names: fold change, fold_change, foldchange, fold change (log2), fold_change (log2), foldchange (log2), fold change (0-n), fold_change (0-n), foldchange (0-n), fold change (+/- mode), fold_change (+/- mode), fold change (log2 mode), fold_change (log2 mode), foldchange (log2 mode), fold change (0-n mode), fold_change (0-n mode), foldchange (0-n mode)*
Optional	Custom	*Custom column header of the users choice for additional statistic data, up to a total of 10 custom columns. Only the first 10 recommended and custom statistical columns will be displayed in the UI.*

Somatic mutation sequencing

Data is defined by sequence coordinates comprised of chromosome, start, stop, allele1, allele2 and refallele.By default the first statistical column after the coordinates are defined (column 4) is used for ranking unless a rank column is specified by the user. If no statistic is provided for mutation data, a rank based on computed impact will be applied.

	Header Name	Notes
Required	chrid	*Alternate names: chromosome id, chromosome, chromosomeid, chrid, chr id, chr*
	start	*Alternate names: start coordinate, start, start_position, start position, begin, begin coordinate, begin_coordinate, begin_position, beginposition*
	stop	*Alternate names: stop coordinate, stop, end, end coordinate, end_position, end position*
	allele1	*Alternate names: allele 1, allele1, allele #1, mut1, mut 1*
	allele2	*Alternate names: allele 2, allele2, allele #2, mut2, mut 2*
	refallele	*Alternate names: reference, reference allele, reference_allele, referenceallele, ref allele, refallele, ref_allele, ref*
Recommended	p-value	*Alternate names: p-value, p value, pvalue*
Optional	Custom	*Custom column header of the users choice for additional statistic data. For somatic mutation, up to 31 recommended and custom columns can be displayed in the UI.*

SNP genotyping

Data is defined by reference SNP cluster IDs and associated statistical columns. The first column after the snp column is used by default for ranking unless the user specifies a rank column. Public data is typically ranked on the p-value statistic measurring the association with the relevant phenotype.

	Header Name	Notes
Required	snp	Contains reference SNP cluster ID. *Alternate names: snp, id, snpid, snp id, marker, marker id, markerid, probeset id, probesetid, probeid, probe id*
Recommended	p-value	*Alternate names: p-value, p value, pvalue*
	allele1	*Alternate names: allele 1, allele1, allele #1, mut1, mut 1*
	allele2	*Alternate names: allele 2, allele2, allele #2, mut2, mut 2*
	Standard Error	*Alternate names: ?*
	allele 1 effect size	*Alternate names: allele 1 freq (case), allele 1 freq, allele 1 frequency (case), allele 1 (case), case*
	allele 2 effect size	*Alternate names: allele 2 freq (control), allele 2 freq, allele 2 frequency (control), allele 2 (control), control*
	allele 1 freq	*Alternate names: allele 1 or 95%ci*
	allele 2 freq	*Alternate names: allele 2 or 95%ci*
	Markertype	*Alternate names: markertype, marker type, marker_type*
	Confidence	*Alternate names: confidence*
Optional	Custom	Numeric or character, total columns = #

Custom

Data, defined by gene identifiers and associated statistics, are mapped to NCBI Gene unique ids. Imported ids can be any of most standard

	Header Name	Notes
Required	gene	Numeric or character gene identifiers. *Alternate names: id, accession number, accession, entrez gene, feature, gene, genbank id, genbank, gene id, gene symbol, gene name, probe set id, probesetid, probe set name, refseq id, refseq, symbol, unigene id, unigene, imported id, protein, protein id, protein name*
Recommended	fold change	Values are unlogged signed +(test/control) or -(control/test)
	fold change (0-n)	Values are unsigned, unlogged test/control will be converted to +/- fold change
	fold change (log2)	Values are signed +log2(test/control) or -log2(control/test) converted to +/- fold change
	p-value	*Alternate names: pvalue, p-value, p value, t-test p-value, t-test*
	Test expression	*Alternate names: average expression, average intensity, expression level, intensity, median expression, test expression, test1 expression, testexpression*
	Control expression	*Alternate names: control expression, control expr, controlexpression*
Optional	Custom	*Custom column header of the users choice for additional statistic data, a total of 5 custom statistical columns can be added. Only the first 7 recommended and custom columns will be displayed in the UI.*