

Regular expressions summary

Special characters	Meaning
.	any character, except newline
\n	newline
\t	tab
\s	any whitespace character (space, newline, tab)
\S	any non-whitespace character
\d	any digit character
\D	any non-digit character
*	match 0 or more times preceding character or group
+	match 1 or more times preceding character or group
?	match 0 or 1 time
?	shortest (non-greedy) match
{ }	repetition
{ , }	repetition, min to max
()	capture
()	grouping
\1	recall the first capture
\2	recall the second capture
\n	recall the n-th capture
^	beginning of string
\$	end of string
[]	any one of set of characters
	alternative
\	special or quote

Regular Expression	Example	Meaning
AGA	TAGATC	match AGA anywhere
^AGA	AGATGC	AGA at the beginning
TAA\$	AAGTAA	TAA at the end
A.T	AAACTG	A followed by any single character (except newline), followed by T
A.*T	CATATCT	A followed by any number of characters, followed by T (greedy)
A.*?T	CATATCT	A followed by any number of characters, followed by T (non-greedy)
(A.*?T)	CATATCT	capture A followed by any number of characters, followed by T (non-greedy)
A{5}	TAAAAATC	5 A's
TA{2,4}CG	CTAAACGA	T followed by 2 to 4 A's, followed by CG
[AT]CG	CCTCGA	either of ACG or TCG
[AA CC TT]CG	ACCCGTA	AA or CC or TT, followed by CG
A(CG){3}T	GACGCGCGTA	A followed by 3 CG's, followed by T
((.)(.){3}\2)	GAATTAC	capture 4 consecutive characters, 1st and 4th, and 2nd and 3rd the same
A\.T	TCGA.TAA	A followed by . and T