Summary of Regular Expression Syntax

regex	meaning
`TATA`	match four consecutive letters, TATA
`TAG|TGA|TAA`	match TAG or TGA or TAA
`.`	match any character but not a newline character
`..`	match any two characters (independently, not necessarily the same character)
`(.)`	capture (remember) and match any character
`.*`	greedy **match any character** 0 or more times (each is independent of others)
`(.*)`	capture and greedy match any character 0 or more times
`(.*?)`	capture and non-greedy match any character 0 or more times
`.+`	greedy match any character 1 or more times (each is independent of others)
`(.+)`	capture and greedy match any character 1 or more times
`(.+?)`	capture and non-greedy, match any character 1 or more times
`\1`	recall the first captured group
`\2`	recall the second captured group
`\n`	recall the *n*th captured group
`.?`	optional, match any character 0 or 1 time
`T?`	optional, match a T or nothing
`(CAAT)?`	Optional, match CAAT or nothing
`A{3,7}`	greedy match between 3 and 7 As
`A{3,}`	greedy match of 3 or more As
`[CG]`	match any *one* of the characters in the set, a C or a G
`TATA[AT]`	match TATA followed by an A or a T
`[^CG]`	match any *one* character that is *not* in the set, not a C and not a G
`[CG]{5,10}`	greedy match a C or a G between 5 and 10 times
`^ATG`	string begins with ATG
`TAG$`	string ends with TAG
`(?:...)`	cluster-only parentheses, don't capture 3 character match (don't remember 3 characters)
`(?=TAG|TGA|TAA)`	True if the look-ahead assertion succeeds; that is, it does find TAG or TGA or TAA
`(?!TAG|TGA|TAA)`	True if the look-ahead assertion fails; that is, it fails to find TAG or TGA or TAA
`\s`	match any whitespace character (tab, space, newline)
`\S`	match any character that is not whitespace
`\d`	match any character that is a digit, same as `[0123456789]`
`\D`	match any character that is not a digit
`\w`	match any one "word" character (includes alphanumeric, plus '_')
`\W`	match any one nonword character

LeBlanc, Mark D., and Betsey Dexter Dyer. *Perl for Exploring DNA*. New York: Oxford UP, USA, 2007. 230-231.

http://cs.wheatoncollege.edu/~mleblanc/dna/