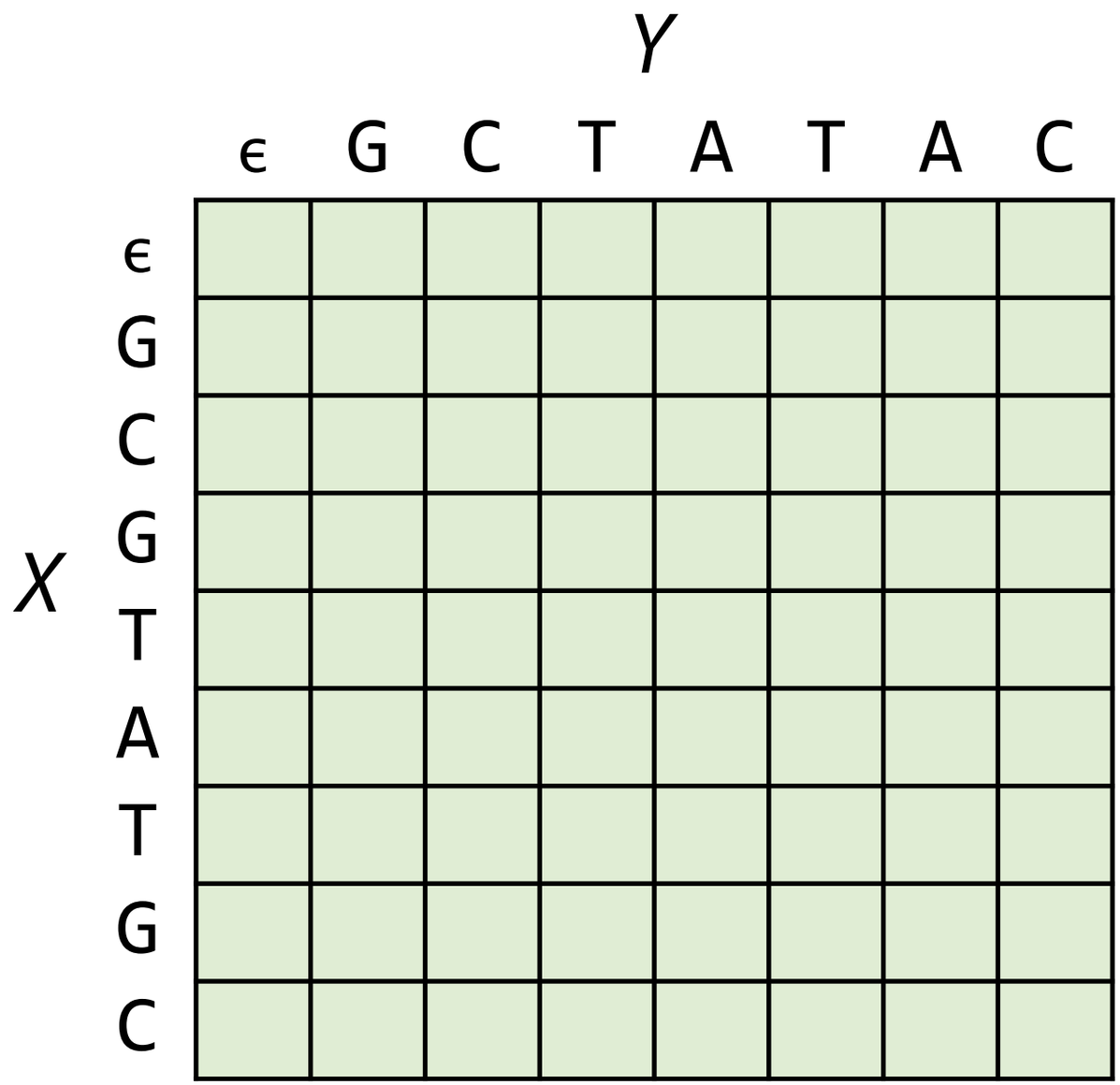
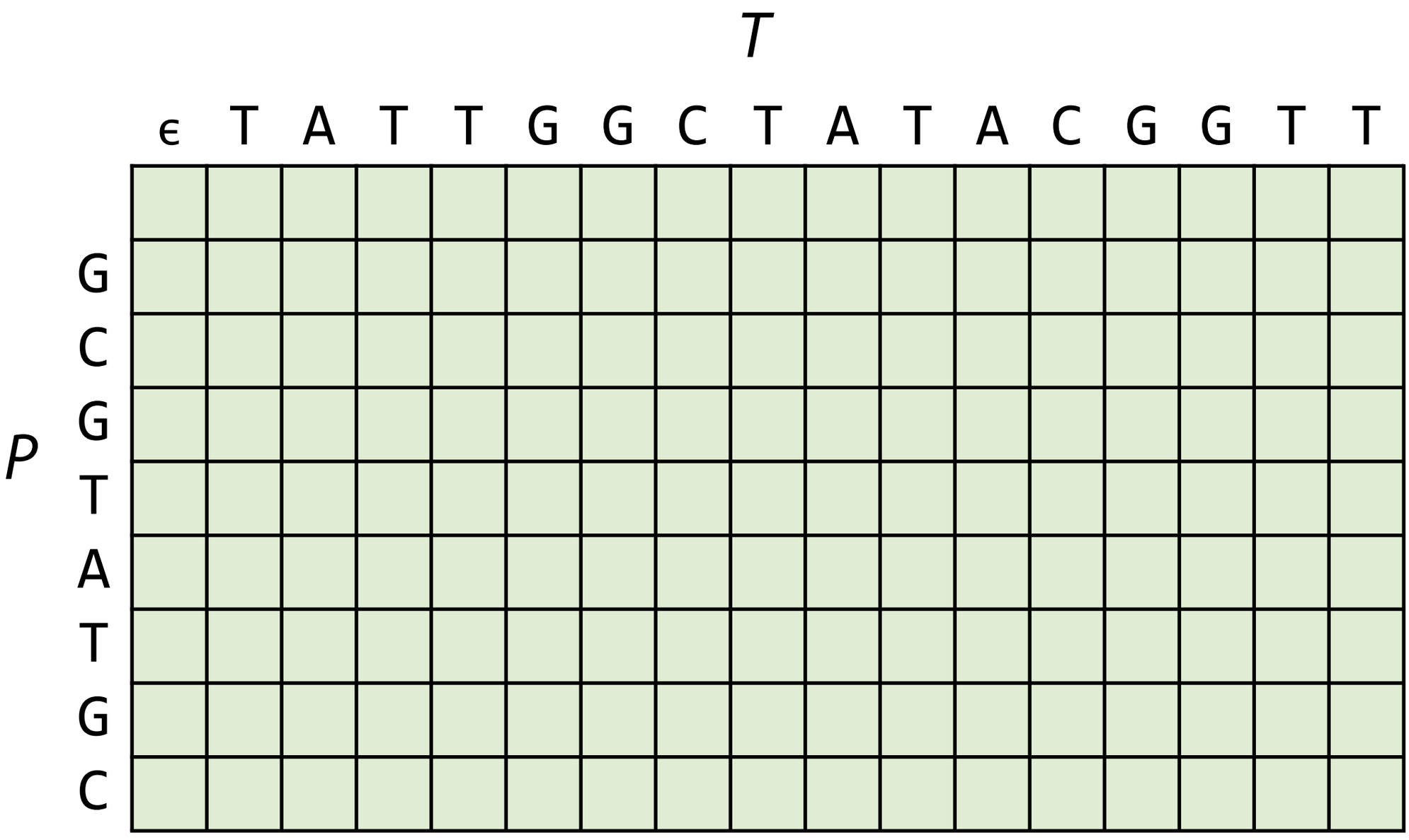
Edit distance



Edit distance & approximate matching



			Y					
		ϵ	G	C	T	A	T	A	C
	ϵ	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
	G	**1**							
	C	**2**							
	G	**3**							
X	T	**4**							
	A	**5**							
	T	**6**							
	G	**7**							
	C	**8**							

T

		ϵ	T	A	T	T	G	G	C	T	A	T	A	C	G	G	T	T
offset →		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	1																
	C	2																
	G	3																
P	T	4																
	A	5																
	T	6																
	G	7																
	C	8																

T

		ϵ	T	A	T	T	G	G	C	T	A	T	A	C	G	G	T	T
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1
	C	2	2	2	2	2	1	1	0	1	2	2	2	1	1	1	1	2
	G	3	3	3	3	3	2	1	1	1	2	3	3	2	1	1	2	2
P	T	4	3	4	3	3	3	2	2	1	2	2	3	3	2	2	1	2
	A	5	4	3	4	4	4	3	3	2	1	2	2	3	3	3	2	2
	T	6	5	4	3	4	5	4	4	3	2	1	2	3	4	4	3	2
	G	7	6	5	4	4	4	5	5	4	3	2	2	3	3	4	4	3
	C	8	7	6	5	5	5	5	5	5	4	3	3	2	3	4	5	4

T

		ϵ	T	A	T	T	G	G	C	T	A	T	A	C	G	G	T	T
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1
	C	2	2	2	2	2	1	1	0	1	2	2	2	1	1	1	1	2
	G	3	3	3	3	3	2	1	1	1	2	3	3	2	1	1	2	2
P	T	4	3	4	3	3	3	2	2	1	2	2	3	3	2	2	1	2
	A	5	4	3	4	4	4	3	3	2	1	2	2	3	3	3	2	2
	T	6	5	4	3	4	5	4	4	3	2	1	2	3	4	4	3	2
	G	7	6	5	4	4	4	5	5	4	3	2	2	3	3	4	4	3
	C	8	7	6	5	5	5	5	5	5	4	3	3	2	3	4	5	4

P occurs in *T* with 2 edits

T

		ϵ	T	A	T	T	G	G	C	T	A	T	A	C	G	G	T	T
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1
	C	2	2	2	2	2	1	1	0	1	2	2	2	1	1	1	1	2
	G	3	3	3	3	3	2	1	1	1	2	3	3	2	1	1	2	2
P	T	4	3	4	3	3	3	2	2	1	2	2	3	3	2	2	1	2
	A	5	4	3	4	4	4	3	3	2	1	2	2	3	3	3	2	2
	T	6	5	4	3	4	5	4	4	3	2	1	2	3	4	4	3	2
	G	7	6	5	4	4	4	5	5	4	3	2	2	3	3	4	4	3
	C	8	7	6	5	5	5	5	5	5	4	3	3	2	3	4	5	4

How did I get here?

T

		ϵ	T	A	T	T	G	G	C	T	A	T	A	C	G	G	T	T
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1
	C	2	2	2	2	2	1	1	0	1	2	2	2	1	1	1	1	2
	G	3	3	3	3	3	2	1	1	1	2	3	3	2	1	1	2	2
P	T	4	3	4	3	3	3	2	2	1	2	2	3	3	2	2	1	2
	A	5	4	3	4	4	4	3	3	2	1	2	2	3	3	3	2	2
	T	6	5	4	3	4	5	4	4	3	2	1	2	3	4	4	3	2
	G	7	6	5	4	4	4	5	5	4	3	2	2	3	3	4	4	3
	C	8	7	6	5	5	5	5	5	5	4	3	3	2	3	4	5	4

T

		ϵ	T	A	T	T	G	G	C	T	A	T	A	C	G	G	T	T
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1
	C	2	2	2	2	2	1	1	0	1	2	2	2	1	1	1	1	2
	G	3	3	3	3	3	2	1	1	1	2	3	3	2	1	1	2	2
P	T	4	3	4	3	3	3	2	2	1	2	2	3	3	2	2	1	2
	A	5	4	3	4	4	4	3	3	2	1	2	2	3	3	3	2	2
	T	6	5	4	3	4	5	4	4	3	2	1	2	3	4	4	3	2
	G	7	6	5	4	4	4	5	5	4	3	2	2	3	3	4	4	3
	C	8	7	6	5	5	5	5	5	5	4	3	3	2	3	4	5	4

How did I get here?

T

		ϵ	T	A	T	T	G	G	C	T	A	T	A	C	G	G	T	T
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1
	C	2	2	2	2	2	1	1	0	1	2	2	2	1	1	1	1	2
	G	3	3	3	3	3	2	1	1	1	2	3	3	2	1	1	2	2
P	T	4	3	4	3	3	3	2	2	1	2	2	3	3	2	2	1	2
	A	5	4	3	4	4	4	3	3	2	1	2	2	3	3	3	2	2
	T	6	5	4	3	4	5	4	4	3	2	1	2	3	4	4	3	2
	G	7	6	5	4	4	4	5	5	4	3	2	2	3	3	4	4	3
	C	8	7	6	5	5	5	5	5	5	4	3	3	2	3	4	5	4

T

P	ϵ	T	A	T	T	G	G	C	T	A	T	A	C	G	G	T	T
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1
C	2	2	2	2	2	1	1	0	1	2	2	2	1	1	1	1	2
G	3	3	3	3	3	2	1	1	1	2	3	3	2	1	1	2	2
T	4	3	4	3	3	3	2	2	1	2	2	3	3	2	2	1	2
A	5	4	3	4	4	4	3	3	2	1	2	2	3	3	3	2	2
T	6	5	4	3	4	5	4	4	3	2	1	2	3	4	4	3	2
G	7	6	5	4	4	4	5	5	4	3	2	2	3	3	4	4	3
C	8	7	6	5	5	5	5	5	5	4	3	3	2	3	4	5	4

T

		ϵ	T	A	T	T	G	G	C	T	A	T	A	C	G	G	T	T
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1
	C	2	2	2	2	2	1	1	0	1	2	2	2	1	1	1	1	2
	G	3	3	3	3	3	2	1	1	1	2	3	3	2	1	1	2	2
P	T	4	3	4	3	3	3	2	2	1	2	2	3	3	2	2	1	2
	A	5	4	3	4	4	4	3	3	2	1	2	2	3	3	3	2	2
	T	6	5	4	3	4	5	4	4	3	2	1	2	3	4	4	3	2
	G	7	6	5	4	4	4	5	5	4	3	2	2	3	3	4	4	3
	C	8	7	6	5	5	5	5	5	5	4	3	3	2	3	4	5	4

T

		ϵ	T	A	T	T	G	G	C	T	A	T	A	C	G	G	T	T
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1
	C	2	2	2	2	2	1	1	0	1	2	2	2	1	1	1	1	2
	G	3	3	3	3	3	2	1	1	1	2	3	3	2	1	1	2	2
P	T	4	3	4	3	3	3	2	2	1	2	2	3	3	2	2	1	2
	A	5	4	3	4	4	4	3	3	2	1	2	2	3	3	3	2	2
	T	6	5	4	3	4	5	4	4	3	2	1	2	3	4	4	3	2
	G	7	6	5	4	4	4	5	5	4	3	2	2	3	3	4	4	3
	C	8	7	6	5	5	5	5	5	5	4	3	3	2	3	4	5	4

T

		ϵ	T	A	T	T	G	G	C	T	A	T	A	C	G	G	T	T
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1
	C	2	2	2	2	2	1	1	0	1	2	2	2	1	1	1	1	2
P	G	3	3	3	3	3	2	1	1	1	2	3	3	2	1	1	2	2
	T	4	3	4	3	3	3	2	2	1	2	2	3	3	2	2	1	2
	A	5	4	3	4	4	4	3	3	2	1	2	2	3	3	3	2	2
	T	6	5	4	3	4	5	4	4	3	2	1	2	3	4	4	3	2
	G	7	6	5	4	4	4	5	5	4	3	2	2	3	3	4	4	3
	C	8	7	6	5	5	5	5	5	5	4	3	3	2	3	4	5	4

T

		ϵ	T	A	T	T	G	G	C	T	A	T	A	C	G	G	T	T
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1
	C	2	2	2	2	2	1	1	0	1	2	2	2	1	1	1	1	2
	G	3	3	3	3	3	2	1	1	1	2	3	3	2	1	1	2	2
P	T	4	3	4	3	3	3	2	2	1	2	2	3	3	2	2	1	2
	A	5	4	3	4	4	4	3	3	2	1	2	2	3	3	3	2	2
	T	6	5	4	3	4	5	4	4	3	2	1	2	3	4	4	3	2
	G	7	6	5	4	4	4	5	5	4	3	2	2	3	3	4	4	3
	C	8	7	6	5	5	5	5	5	5	4	3	3	2	3	4	5	4

T

		ϵ	T	A	T	T	G	G	C	T	A	T	A	C	G	G	T	T
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1
	C	2	2	2	2	2	1	1	0	1	2	2	2	1	1	1	1	2
	G	3	3	3	3	3	2	1	1	1	2	3	3	2	1	1	2	2
P	T	4	3	4	3	3	3	2	2	1	2	2	3	3	2	2	1	2
	A	5	4	3	4	4	4	3	3	2	1	2	2	3	3	3	2	2
	T	6	5	4	3	4	5	4	4	3	2	1	2	3	4	4	3	2
	G	7	6	5	4	4	4	5	5	4	3	2	2	3	3	4	4	3
	C	8	7	6	5	5	5	5	5	5	4	3	3	2	3	4	5	4

T

		ϵ	T	A	T	T	G	G	C	T	A	T	A	C	G	G	T	T
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1
	C	2	2	2	2	2	1	1	0	1	2	2	2	1	1	1	1	2
	G	3	3	3	3	3	2	1	1	1	2	3	3	2	1	1	2	2
P	T	4	3	4	3	3	3	2	2	1	2	2	3	3	2	2	1	2
	A	5	4	3	4	4	4	3	3	2	1	2	2	3	3	3	2	2
	T	6	5	4	3	4	5	4	4	3	2	1	2	3	4	4	3	2
	G	7	6	5	4	4	4	5	5	4	3	2	2	3	3	4	4	3
	C	8	7	6	5	5	5	5	5	5	4	3	3	2	3	4	5	4

Match occurs at offset 5

		ϵ	T	A	T	T	G	G	C	T	A	T	A	C	G	G	T	T
offset →		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1
	C	2	2	2	2	2	1	1	0	1	2	2	2	1	1	1	1	2
	G	3	3	3	3	3	2	1	1	1	2	3	3	2	1	1	2	2
P	T	4	3	4	3	3	3	2	2	1	2	2	3	3	2	2	1	2
	A	5	4	3	4	4	4	3	3	2	1	2	2	3	3	3	2	2
	T	6	5	4	3	4	5	4	4	3	2	1	2	3	4	4	3	2
	G	7	6	5	4	4	4	5	5	4	3	2	2	3	3	4	4	3
	C	8	7	6	5	5	5	5	5	5	4	3	3	2	3	4	5	4

T: G C - T A T A C

P: G C G T A T G C

T

P		ϵ	T	A	T	T	G	G	C	T	A	T	A	C	G	G	T	T
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1
	C	2	2	2	2	2	1	1	0	1	2	2	2	1	1	1	1	2
	G	3	3	3	3	3	2	1	1	1	2	3	3	2	1	1	2	2
	T	4	3	4	3	3	3	2	2	1	2	2	3	3	2	2	1	2
	A	5	4	3	4	4	4	3	3	2	1	2	2	3	3	3	2	2
	T	6	5	4	3	4	5	4	4	3	2	1	2	3	4	4	3	2
	G	7	6	5	4	4	4	5	5	4	3	2	2	3	3	4	4	3
	C	8	7	6	5	5	5	5	5	5	4	3	3	2	3	4	5	4