

CS342: Bioinformatics

Lecture 12

Dynamic programming matrix:

		j → (sequence y)								
		0	1	2	3	4	5	6	7	8 = N
			T	G	C	T	C	G	T	A
i ↓ (sequence x)	0	0	-6	-12	-18	-24	-30	-36	-42	-48
	1 T	-6	5	-1	-7	-13	-19	-25	-31	-37
	2 T	-12	-1	3	-3	-2	-8	-14	-20	-26
	3 C	-18	-7	-3	8	2	3	-3	-9	-15
	4 A	-24	-13	-9	2	6	0	1	-5	-4
	5 T	-30	-19	-15	-4	7	4	-2	6	0
M = 6 A	-36	-25	-21	-10	1	5	2	0	11	

Optimum alignment scores 11:

T	-	-	T	C	A	T	A
T	G	C	T	C	G	T	A
+5	-6	-6	+5	+5	-2	+5	+5

POLICY-ISH

As Made-To-Order DNA Gets Cheaper, Keeping It Out Of The Wrong Hands Gets Harder

September 24, 2019 · 7:13 AM ET

Heard on [Morning Edition](#)



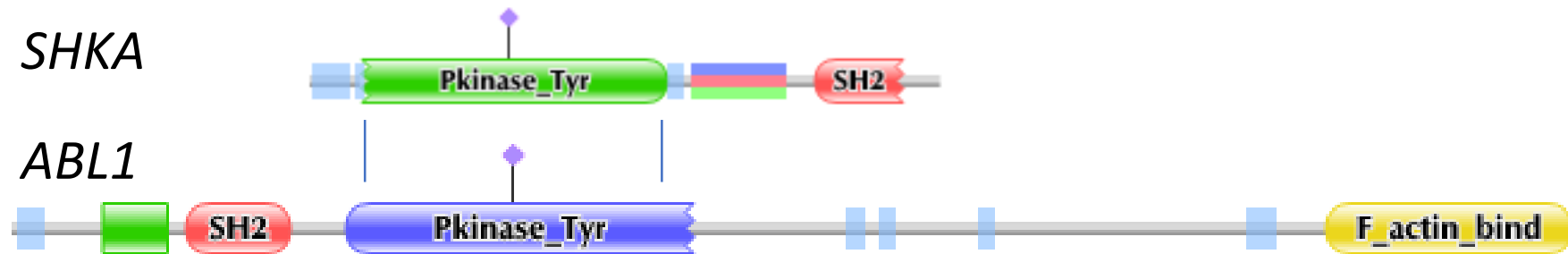
NELL GREENFIELDBOYCE



An employee of the Boston biotech company Ginkgo Bioworks runs a gene sequencing machine through its paces. The company synthesizes thousands of genes a month, which are then inserted into cells that become mini factories of useful products.

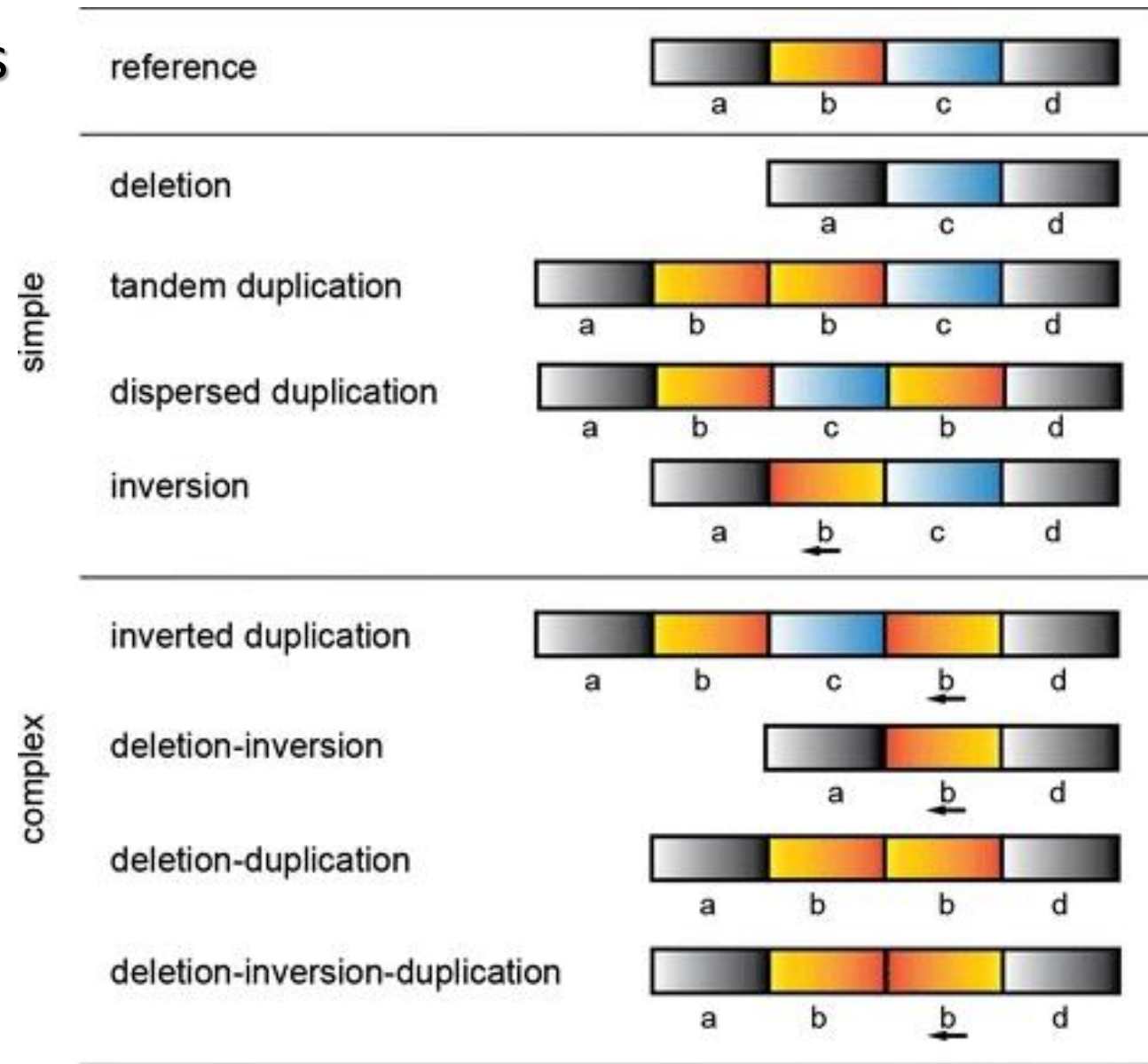
Protein Domains and Local Alignment

Example: Proteins often consist of multiple functional units, called *domains*. Number and order of domains varies in different proteins.



From Pfam database (<http://pfam.sanger.ac.uk/>)

Structural Variants



A naïve approach.

- Idea: Add a new set of edges to the edit graph, representing all possible gaps.

