

```
pos[0] = 0;
for(m = t_len - 1, nuc_pos = t_e - 1; m >= 0; --m, --nuc_pos) {
    if(nuc_pos < 0) {
        nuc_pos = template_length - 1;
    }
    D_ptr[q_len] = (0 < k) ? 0 : (W1 + (t_len - 1 - m) * U);
    Q_prev = (t_len + q_len) * (MM + U + W1);

    t_nuc = getNuc(template, nuc_pos);
    for(n = q_len - 1; n >= 0; --n) {
        E_ptr[n] = 0;

        /* update Q and P, gap openings */
        Q = D_ptr[n + 1] + W1;
        P_ptr[n] = D_prev[n] + W1;
        if(Q < P_ptr[n]) {
            D_ptr[n] = P_ptr[n];
            e = 4;
        } else {
            D_ptr[n] = Q;
            e = 2;
        }

        /* update Q and P, gap extensions */
        /* mark bit 4 and 5 as possible gap-opennings, if necessary */
        thisScore = Q_prev + U;
        if(Q < thisScore) {
            Q = thisScore;
            if(e == 4) {
                D_ptr[n] |= 1;
                e = 3;
            }
        } else {
            E_ptr[n] |= 16;
        }
        thisScore = P_prev[n] + U;
        if(P_ptr[n] < thisScore) {
            P_ptr[n] = thisScore;
            if(D_ptr[n] < thisScore) {
                D_ptr[n] = thisScore;
                e = 5;
            }
        } else {
            E_ptr[n] |= 32;
        }

        /* Update D, match */
        thisScore = D_prev[n + 1] + d[t_nuc][query[n]];
        if(D_ptr[n] < thisScore) {
            D_ptr[n] = thisScore;
            E_ptr[n] |= 1;
        } else {
            E_ptr[n] |= e;
        }
    }
    Q_prev = Q;
}

E_ptr -= (q_len + 1);

if(k < 0 && Stat.score <= *D_ptr) {
    Stat.score = *D_ptr;
    pos[0] = m;
}

tmp = D_ptr;
D_ptr = D_prev;
D_prev = tmp;

tmp = P_ptr;
P_ptr = P_prev;
P_prev = tmp;
}
E_ptr = E;
```

Alignment and Mapping

Philip T.L.C. Clausen

SHOGUN

T/I/O

A
S
Z
O
C

T/O

I/O

I Sequences

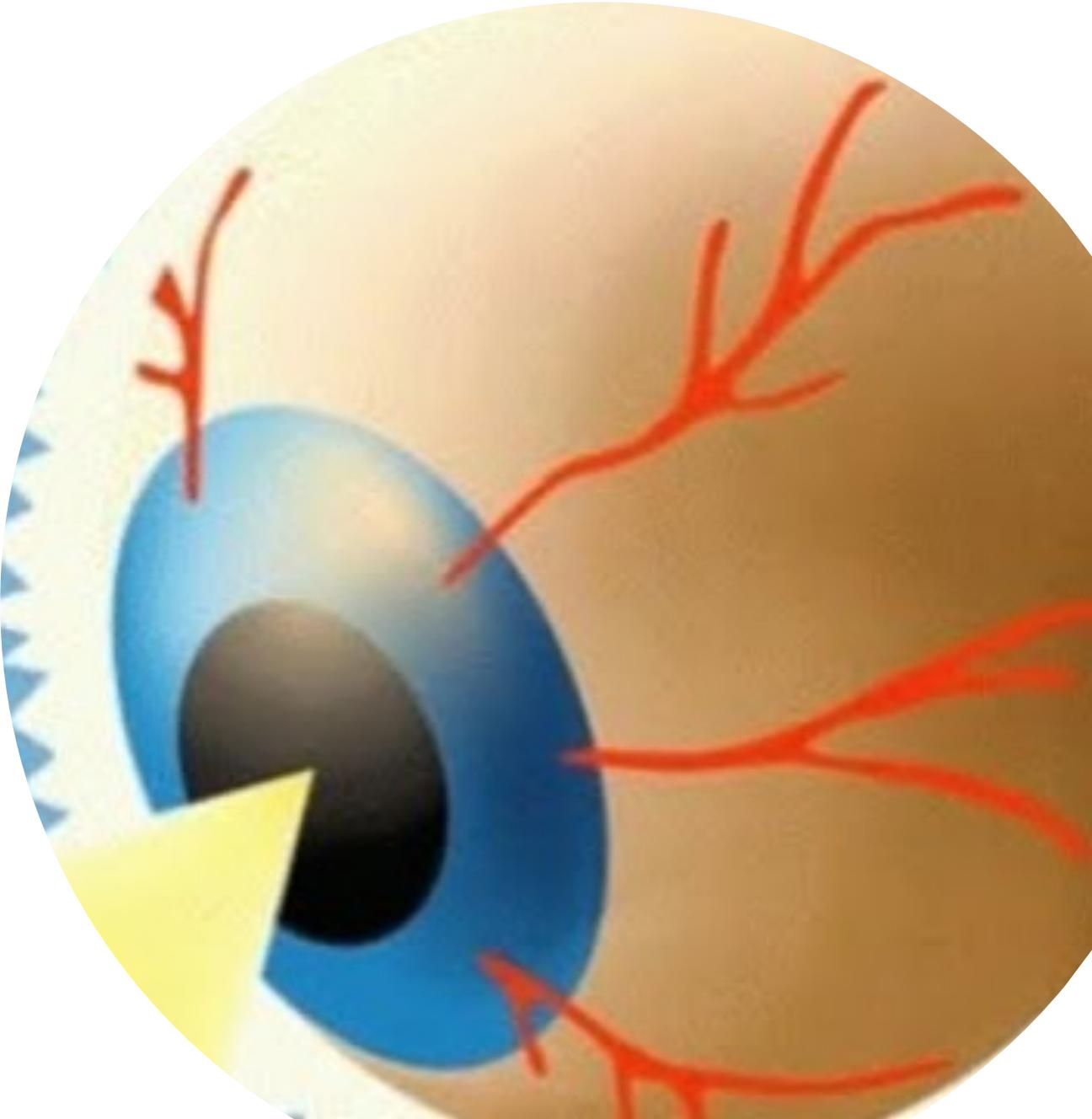
/ Mapping and Alignment

O (Intermediate) Result

/

- What is alignment
- What is mapping
- Redundant sequences

1951





1966

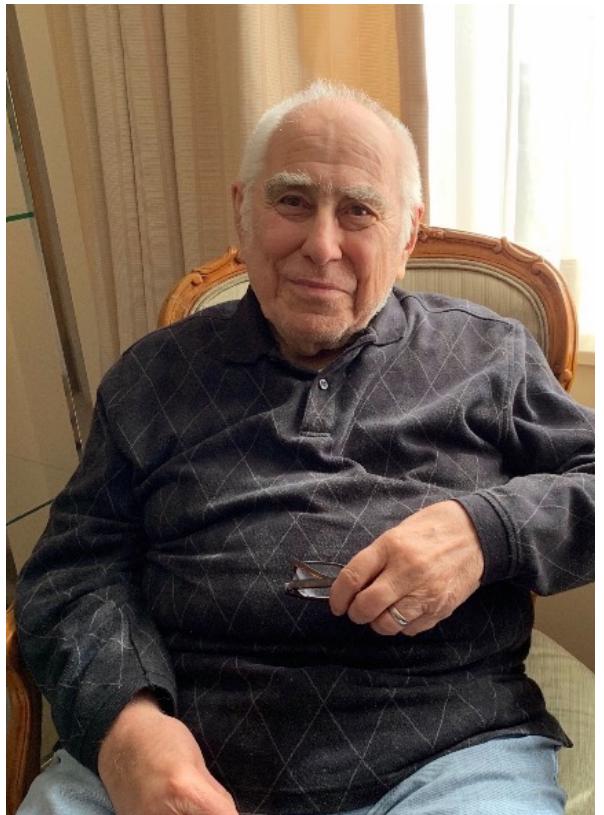
Number of possible alignments

$$\binom{m+n}{n} = \frac{(m+n)!}{m! n!}$$



150 X 150

1970



Saul B. Needleman

$$D(i, j) = \max \begin{cases} D(i + 1, j + 1) + S(T_i, Q_j) \\ P(i, j) \\ Q(i, j) \end{cases}$$

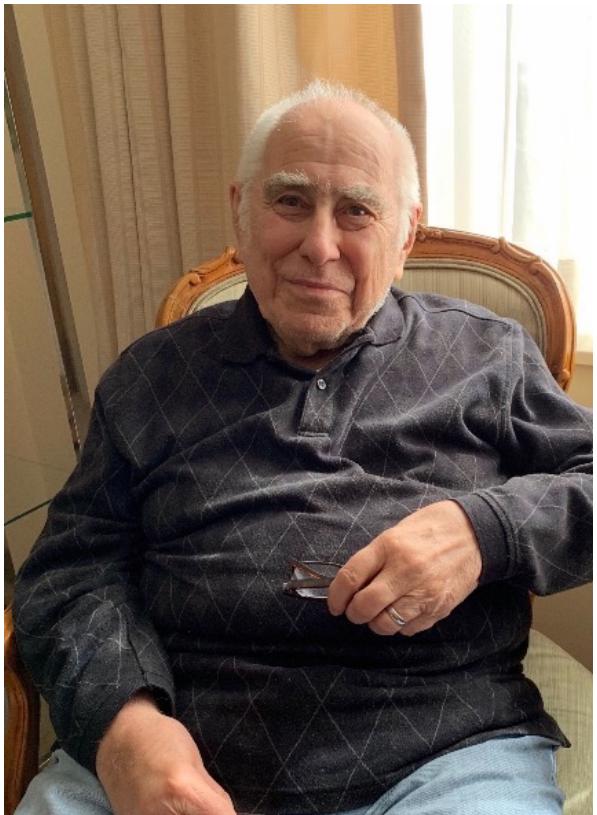
$$P(i, j) = \max \begin{cases} D(i + 1, j) + w_1 \\ P(i + 1, j) + U \end{cases}$$

$$Q(i, j) = \max \begin{cases} D(i, j + 1) + w_1 \\ Q(i, j + 1) + U \end{cases}$$



Christian D. Wunsch

Simplified NW



Saul B. Needleman

$$D(i, j) = \max \begin{cases} D(i + 1, j + 1) + S(T_i, Q_j) \\ D(i + 1, j) + W \\ D(i, j + 1) + W \end{cases}$$



Christian D. Wunsch

What if we made the two sequences form a matrix?

$$D(i, j) = \max \begin{cases} D(i + 1, j + 1) + S(T_i, Q_j) \\ D(i + 1, j) + W \\ D(i, j + 1) + W \end{cases}$$
$$S(x, y) = \begin{cases} M & x = y \\ P & x \neq y \end{cases}$$
$$M = 1, P = -1, W = -3$$

Q\T	A	C	G	T	-
A					
G					
T					
T					
-					

What if we made the two sequences form a matrix?

$$D(i, j) = \max \begin{cases} D(i + 1, j + 1) + S(T_i, Q_j) \\ D(i + 1, j) + W \\ D(i, j + 1) + W \end{cases}$$

$$S(x, y) = \begin{cases} M & x = y \\ P & x \neq y \end{cases}$$

$$M = 1, P = -1, W = -3$$

T: ACGT
|__|
Q: AGTT

Q\T	A	C	G	T	-
A	0	-2	-4	-8	-12
G	-4	-1	-1	-5	-9
T	-6	-3	0	-2	-6
T	-8	-5	-2	1	-3
-	-12	-9	-6	-3	0

Drastic improvement on computation speed

$$\frac{(m + n)!}{m! \ n!} \gg m \ n$$

Sequence lengths	Needleman-Wunsch	Exhaustive search	Number of e ⁻ in the universe
150 X 150	22 500	$\sim 10^{89}$	$\sim 10^{80}$

And now you...

$$D(i, j) = \max \begin{cases} D(i + 1, j + 1) + S(T_i, Q_j) \\ D(i + 1, j) + W \\ D(i, j + 1) + W \end{cases}$$

$$S(x, y) = \begin{cases} M & x = y \\ P & x \neq y \end{cases}$$

$$M = 1, P = -1, W = -3$$

T:	ACGT
	__
Q:	AGTT

Q\T	A	C	G	T	-
A	0	-2	-4	-8	-12
G	-4	-1	-1	-5	-9
T	-6	-3	0	-2	-6
T	-8	-5	-2	1	-3
-	-12	-9	-6	-3	0

$$D_{i,j} = \max \begin{cases} D_{i+1,j+1} + S(T_i, Q_j) \\ D_{i+1,j} + W \\ D_{i,j+1} + W \end{cases}, \quad S(x,y) = \begin{cases} M & x = y \\ P & x \neq y \end{cases}$$

M: 1 P: -2 W: -3

<i>Q\T</i>	C	G	T	A	A	C	G	T	A	T	-
G	0	3									
T			2								
A				1							
T					0						
C						2					
G							1				
T								0			
T									3	-1	
A										2	
T											1
-											0

NW score: 0

T: CGTAACGT-AT
||||||_|||
Q: -GTATCGTTAT

Score = 8 * M + 1 * P + 2 * W = 8 - 2 - 6 = 0

T: CGTAACG-TAT
||||||_|||
Q: -GTATCGTTAT

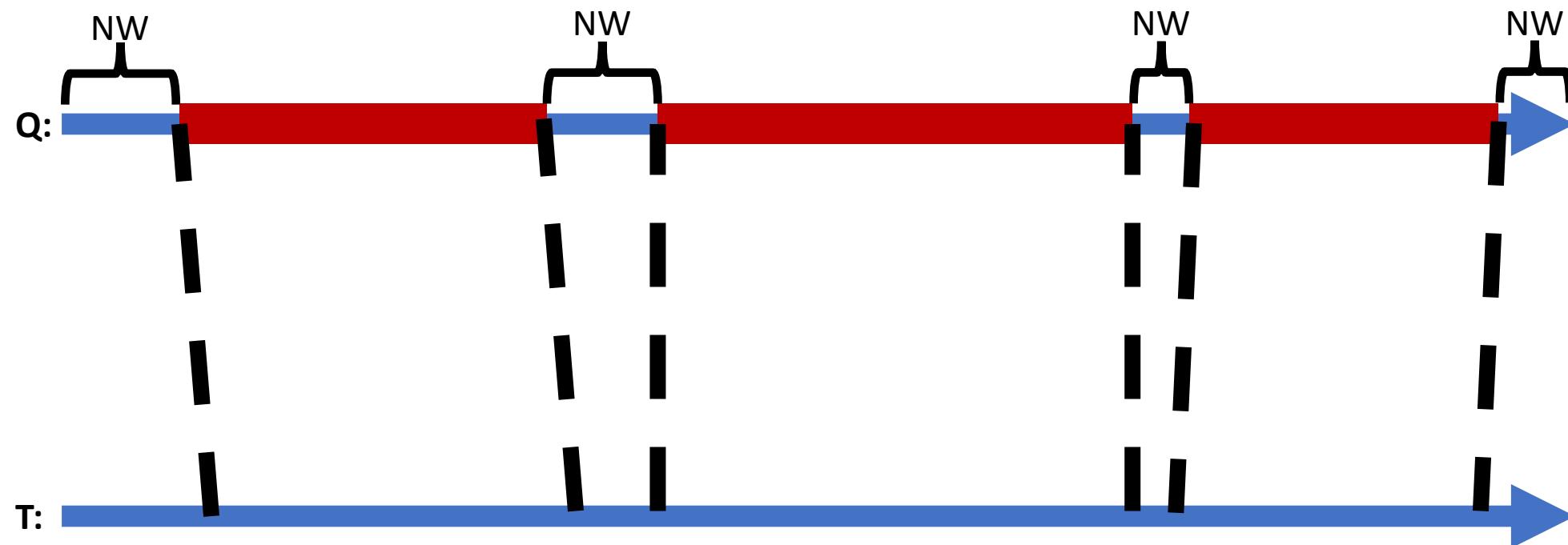
Is a run time of “m n” good enough?



1983

David J. Lipman

Fasta, 1985



Seed and extend

Simple example, k = 5:

Q: AAGACTCCGACTGGGACTTTGATGTTCGAAAGA

T: GACTGGGACTTTGATG

Seed and extend

Simple example, $k = 5$:

Q: AAGACTCCGACTGGGACTTTGATGTTCGAAAGA

T: GACTGGGACTTTGATG

Seed and extend

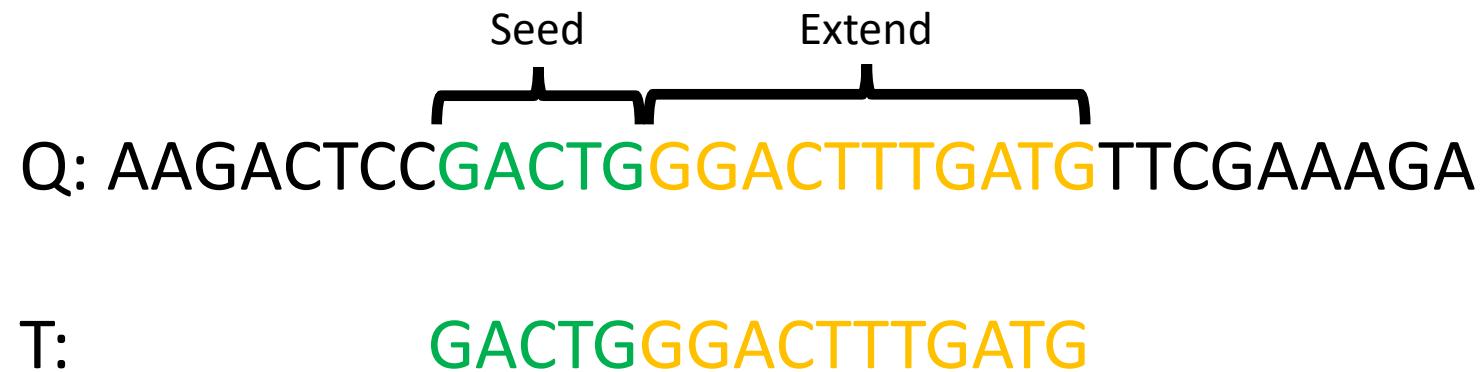
Simple example, k = 5:

Q: AAGACTCC**GACTG**GGACTTGTGATGTTCGAAAGA

T: **GACTG**GGACTTGTGATG

Seed and extend

Simple example, $k = 5$:



Seed and extend, impact on computation

$$m \cdot n \gg m + n$$

And now you...

$K = 10$, *prefix : T*

K = 10, prefix : T

@read 1					
GC GCGG TATTATCCCGTGTGACGCCGGCAAGAGTAACCGTCGCCGCATACTACTATTCTCAGAATGACTTG	TAAAAGATGC	91	TCTAGCTTCC	593	TGTTTTGCT 59
+	TAAAGTTCTG	209	TCTCAACAGC	143	TTAAAGTTCT 208
IIIHIIHHHHHIHIGIHHHIHIIHHGHIGEIIIEGFHGIHIFHGFIGIFGFIHEFIFGGIIFIGFIDIIG	TAAAGTTGCA	635	TCTCAGAACATG	285	TTAACTGGCG 573
>blaTEM-1a	TAATCTGGA	692	TCTCGCGGTA	717	TTAACGATTG 847
0 ATGAGTATTCAACATTCCGTGTCGCCCTT	TAACACTGCG	386	TCTGACAACG	407	TTAATAGACT 612
30 ATTCCCTTTTGC GGCACTTGCCTTCCT	TAACCAGAG	373	TCTGCGCTCG	653	TTACATCGAA 128
60 GTTTTGCTCACCCAGAAACGCTGGTGAAA	TAACCGCTT	436	TCTGCTATGT	215	TTACGGATGG 331
90 GTAAAAGATGCTGAAGATCAGTTGGGTGCA	TAACTCGCCT	469	TCTGGAGCCG	696	TTACTCTAGC 589
120 CGAGTGGTTACATCGAACTGGATCTAAC	TAACTGGCGA	574	TCTTACGGAT	329	TTACTTCTGA 402
150 AGCGGTAAAGATCCTTGAGAGTTTCGCC	TAAGAGAATT	349	TGAAAGTAAA	85	TTATCCCGTG 234
180 GAAGAACGTTTCCAATGATGAGCACTTT	TAAGATCCTT	155	TGAACGAAAT	806	TTATCTACAC 772
210 AAAGTTCTGCTATGTGGCGCGTATTATCC	TAAGCATTGG	848	TGAAGATCAG	101	TTATGCAGTG 357
240 CGTGTGACGCCGGCAAGAGCAACTCGGT	TAAGCCCTCC	752	TGAAGCCATA	503	TTATCCCTT 28
270 CGCCGCATAACATTCTCAGAATGACTTG	TAATAGACTG	613	TGAATGAAGC	499	TTATTGCTGA 682
300 GTTGAGTACTCACCAAGTCACAGAAAAGCAT	TACACGACGG	777	TGACAACGAT	409	TTCAACATTT 7
330 CTTACGGATGGCATGACAGTAAGAGAATTA	TACACTATT	277	TGACACCACG	527	TTCCAATGAT 190
360 TGCAGTGCTGCCATAACCATGAGTGATAAC	TACATCGAAC	129	TGACAGTAAG	343	TTCCCGGCAA 599
390 ACTGCGGCCAACTTACTTCTGACAACGATC	TACCAAACGA	511	TGACGCCGG	245	TTCCCTTTTT 31
420 GGAGGACCGAAGGAGCTAACCGCTTTTG	TACGGATGGC	332	TGACTTGGTT	293	TTCCGGCTGG 667
450 CACAACATGGGGATCATGTAACCGCCCTT	TACTCACCAG	306	TGAGAGTTT	164	TTCCGTGTG 15
480 GATCGTTGGGAAACCGGAGCTGAATGAAGCC	TACTCTAGCT	590	TGAGATAGGT	827	TTCTGTGTTT 55
510 ATACCAAACGACGAGCGTGACACCACGATG	TACTTACTCT	586	TGAGCACTT	199	TTCGCCCCGA 172
540 CCTGCAGCAATGGCAACACGTTGCGCAA	TACTTCTGAC	403	TGAGCGTGG	707	TTCTCAGAAT 284
570 CTATTAACCTGGCGAACTACTTACTCTAGCT	TAGACAGATC	815	TGAGTACTCA	302	TTCTGACAAAC 406
600 TCCC GGCAACAATTAA TAGACTGGATGGAG	TAGACTGGAT	616	TGAGTATTCA	1	TTCTGCGCTC 652
630 GCGGATAAAGTTGCAGGACCACCTCTGCGC	TAGCTCCCG	595	TGAGTGATAA	379	TTCTGCTATG 214
660 TCGGCCCTCCGGCTGGCTGGTTATTGCT	TAGGTGCTC	832	TGATAAAATCT	689	TTGACGCCGG 244
690 GATAAAATCTGGAGCCGGTGAGCGTGGGTCT	TAGTTATCTA	769	TGATAACACT	383	TTGAGAGTTT 163
720 CGCGGTATCATTGCAAGCACTGGGGCCAGAT	TATCATTGCA	725	TGATCGTTGG	479	TTGAGTACTC 301
750 GGTAAAGCCCTCCGTATCGTAGTTATCTAC	TATCCCGTGT	235	TGATGAGCAC	196	TTGATCGTTG 478
780 ACGACGGGGAGTCAGGCAACTATGGATGAA	TATCGTAGTT	764	TGATTAAGCA	844	TTGCACAACA 447
810 CGAAATAGACAGATCGCTGAGATAGGTGCC	TATCTACACG	773	TGCAACAACT	448	TTGCAGCACT 730
840 TCACTGATTAAGCATTGGTAA	TATGCAGTGC	358	TGCACGAGTG	116	TTGCAGGACC 640
	TATGGATGAA	800	TGCAGCAATG	542	TTGCCCTCCT 50
	TATGTGGCGC	220	TGCAGCACTG	731	TTGCGAAC 561
	TATTAACTGG	571	TGCAGGACCA	641	TTGCGGCATT 40
	TATTATCCCG	232	TGCAGTGCTG	360	TTGCTCACCC 64
	TATTCAACAT	5	TGCCTAAACC	368	TTGCTGATAA 685
	TATTCCTT	29	TGCCTCACTG	836	TTGGGAACCG 485
	TATTCTCAGA	282	TGCCTGCAGC	538	TTGGGTGCAC 111
	TATTGCTGAT	683	TGCCTTCCGT	51	TTGGTTGAGT 297
	TCAACAGCGG	145	TGCGCAAACCT	562	TTAAAGTTC 207
	TCAACATTTC	8	TGCGCTCGC	655	TTTATTGCTG 681
	TCACAGAAA	316	TGCGGCATT	41	TTTCAATGA 189
	TCACCAGTC	309	TGCGGCCAAC	392	TTTCCGTGTC 14
	TCACCCAGAA	68	TGCTATGTGG	217	TTTCGCCCCG 171
	TCACTGATTA	840	TGCTCACCCA	65	TTTGACAAAC 446
	TCGGTCGCCG	265	TGCTGAAGAT	98	TTTGCTTCC 49

K = 10, prefix : T

@read 1

GCGCGG TATTATCCCGTGTGACGCCGGCAAGAGTAACCGTCGCCGCATAACTATTCTCAGAATGACTTG

+

IIIHIIHHHHHIHIGIHHHIHIIHHGHIGEIIIEGFHGIHIFHGFIGIGFHIHEFIFGGIIFIGFIDIIG

>blaTEM-1a

0	ATGAGTATTCAACATTCCGTGTCGCCCTT
30	ATTCCCTTTTGC GG CATTTCGCCTTCCT
60	GTTTTGCTCACCCAGAAACGCTGGTGAAA
90	GTAAAAGATGCTGAAGATCAGTTGGGTGCA
120	CGAGTGGTTACATCGAACTGGATCTAAC
150	AGCGGTAAAGATCCTTGAGAGTTTCGCC
180	GAAGAACGTTTCCAATGATGAGCACTTT
210	AAAGTTCTGCTATGTGGCGGGTATTATCC
240	CGTGTGACGCCGGCAAGAGCAACTCGGT
270	CGCCGCATAACATTCTCAGAATGACTTG
300	GTGAGTACTCACCAAGTCACAGAAAAGCAT
330	CTTACGGATGGCATGACAGTAAGAGAATTA
360	TGCAGTGCTGCCATAACCAGTAGTGATAAC
390	ACTGCGGCCAACTTACTTCTGACAACGATC
420	GGAGGACCGAAGGAGCTAACCGCTTTTG
450	CACAAACATGGGGATCATGTAACCGCCCTT
480	GATCGTTGGGACCGGAGCTGAATGAAGCC
510	ATACCAAACGACGAGCGTGACACCACGATG
540	CCTGCAGCAATGGCAACACGTTGCGCAA
570	CTATTAACGGCGAACTACTTACTCTAGCT
600	TCCC GG CAACAATTAA TAGACTGGATGGAG
630	GCGGATAAAGTTGCAGGACCACTTCTGCGC
660	TCGGCCCTCCGGCTGGCTGGTTATTGCT
690	GATAAAATCTGGAGCCGGTGAGCGTGGGTCT
720	CGCGGTATCATTGCAAGCACTGGGCCAGAT
750	GGTAAGCCCTCCGTATCGTAGTTATCTAC
780	ACGACGGGGAGTCAGGCAACTATGGATGAA
810	CGAAATAGACAGATCGCTGAGATAGGTGCC
840	TCACTGATTAAGCATTGGTAA

TAAAAGATGC	91	TCTAGCTTCC	593	TGTTTTGCT	59
TAAAGTTCTG	209	TCTCAACAGC	143	TTAAAGTTCT	208
TAAAGTTGCA	635	TCTCAGAATG	285	TTAACTGGCG	573
TAAATCTGGA	692	TCTCGCGGT	717	TTAACGATTG	847
TAACACTGCG	386	TCTGACAACG	407	TTAATAGACT	612
TAACCAGAG	373	TCTGCGCTG	653	TTACATCGAA	128
TAACCGCTT	436	TCTGCTATGT	215	TTACGGATGG	331
TAACTCGC	469	TCTGGAGCCG	696	TTACTCTAGC	589
TAACGGCGA	574	TCTTACGGAT	329	TTACTCTGA	402
TAAGAGAATT	349	TGAAAGTAAA	85	TTATCCCGT	234
TAAGATCCTT	155	TGAACGAAAT	806	TTATCTACAC	772
TAAGCATTG	848	TGAAGATCAG	101	TTATGCAGTG	357
TAAGCCCTC	752	TGAAGCCATA	503	TTATCCCTT	28
TAATAGACTG	613	TGAATGAAGC	499	TTATTGCTGA	682
TACACGACGG	777	TGACAACGAT	409	TTCAACATTT	7
TACACTATT	277	TGACACCACG	527	TTCCAATGAT	190
TACATCGAAC	129	TGACAGTAAG	343	TTCCGGCAA	599
TACCAAACGA	511	TGACGCCGG	245	TTCCCTTTT	31
TACGGATGGC	332	TGACTTGGTT	293	TTCCGGCTGG	667
TACTCACCAG	306	TGAGAGTTT	164	TTCCGTGTCG	15
TACTCTAGCT	590	TGAGATAGGT	827	TTCTGTTTT	55
TACTTACTCT	586	TGAGCACTT	199	TTCGCCCCGA	172
TACTTCTGAC	403	TGAGCGTGGG	707	TTCTCAGAAT	284
TAGACAGATC	815	TGAGTACTCA	302	TTCTGACAC	406
TAGACTGGAT	616	TGAGTATTCA	1	TTCTGCGCTC	652
TAGCTCCCG	595	TGAGTGATAA	379	TTCTGCTATG	214
TAGGTGCTC	832	TGATAAAATCT	689	TTGACGCCG	244
TAGTTATCTA	769	TGATAACACT	383	TTGAGAGTT	163
TATCATTGCA	725	TGATCGTTGG	479	TTGAGTACTC	301
TATCCCGTGT	235	TGATGAGCAC	196	TTGATCGTTG	478
TATCGTAGTT	764	TGATTAAGCA	844	TTGCACAACA	447
TATCTACACG	773	TGCACAAACAT	448	TTGCAGCACT	730
TATGCAGTGC	358	TGCACGAGTG	116	TTGCAGGACC	640
TATGGATGAA	800	TGCAGCAATG	542	TTGCCTTCCT	50
TATGTGGCGC	220	TGCAGCACTG	731	TTGCGAAC	561
TATTAACGG	571	TGCAGGACCA	641	TTGCGGATT	40
TATTATCCCG	232	TGCAGTGCTG	360	TTGCTCACCC	64
TATTCAACAT	5	TGCCTAAACC	368	TTGCTGATAA	685
TATTCCCTT	29	TGCCTCACTG	836	TTGGGAACCG	485
TATTCTCAGA	282	TGCCTGCAGC	538	TTGGGTGCAC	111
TATTGCTGAT	683	TGCCTTCCTG	51	TTGGTTGAGT	297
TCAACAGCGG	145	TGCGCAAAC	562	TTAAAGTTC	207
TCAACATTTC	8	TGCGCTCGC	655	TTTATTGCTG	681
TCACAGAAA	316	TGCGGCATT	41	TTTCAATGA	189
TCACCAGTC	309	TGCGGCCAAC	392	TTTCCGTGTC	14
TCACCCAGAA	68	TGCTATGTGG	217	TTTCGCCCCG	171
TCACTGATTA	840	TGCTCACCA	65	TTTGACAAAC	446
TCGGTCGCCG	265	TGCTGAAGAT	98	TTTGCTTCC	49

K = 10, prefix : T

@read 1

GCGCGG TATTATCCCGTGTGACGCCGGCAAGAGTAACCGTCGCCGCATAACTATTCTCAGAATGACTTG

+

IIIHIIHHHHHIHIGIHHHIHIIHHGHIGEIIIEGFHGIHIFHGFIGIGFHIHEFIFGGIIFIGFIDIIG

>blaTEM-1a

0 ATGAGTATTCAACATTCCGTGTCGCCCTT

30 ATTCCCTTTTGC GG CATTTCGCCTTCCT

60 GTTTTGCTCACCCAGAACGCTGGTGAAA

90 GTAAAAGATGCTGAAGATCAGTTGGGTGCA

120 CGAGTGGTTACATCGAACTGGATCTAAC

150 AGCGGTAAGATCCTTGAGAGTTTCGCC

180 GAAGAACGTTTCCAATGATGAGCACTTT

210 AAAGTTCTGCTATGTGGCGCG TATTATCC

240 CGTGTGACGCCGGCAAGAGCAACTCGGT

270 CGCCGCATAACACTATTCTCAGAATGACTTG

300 GTTGAGTACTCACCAAGTCACAGAAAAGCAT

330 CTTACGGATGGCATGACAGTAAGAGAATTA

360 TGCAGTGCTGCCATAACCATGAGTGATAAC

390 ACTGCGGCCAACTTACTTCTGACAACGATC

420 GGAGGACCGAAGGAGCTAACCGCTTTTG

450 CACAACATGGGGATCATGTAACCGCCCTT

480 GATCGTTGGAACCGGAGCTGAATGAAGCC

510 ATACCAAACGACGAGCGTGACACCACGATG

540 CCTGCAGCAATGGCAACACGTTGCGCAA

570 CTATTAACCTGGCGAACTACTTACTCTAGCT

600 TCCC GG CAACAATTAA TAGACTGGATGGAG

630 GCGGATAAAGTTGCAGGACCACTTCTGCGC

660 TCGGCCCTCCGGCTGGCTGGTTATTGCT

690 GATAAACTGGAGCCGGTGAGCGTGGGTCT

720 CGCGGTATCATTGCA GCACTGGGCCAGAT

750 GGTAAAGCCCTCCGTATCGTAGTTATCTAC

780 ACGACGGGGAGTCAGGCAACTATGGATGAA

810 CGAAATAGACAGATCGCTGAGATAGGTGCC

840 TCACTGATTAAGCATTGGTAA

TAAAAGATGC	91	TCTAGCTTCC	593	TGTTTTGCT	59
TAAAGTTCTG	209	TCTCAACAGC	143	TTAAAGTTCT	208
TAAAGTTGCA	635	TCTCAGAATG	285	TTAACTGGCG	573
TAAATCTGGA	692	TCTCGCGGTA	717	TTAACGATTG	847
TAACACTGCG	386	TCTGACAACG	407	TTAATAGACT	612
TAACCAGAG	373	TCTGCGCTG	653	TTACATCGAA	128
TAACCGCTT	436	TCTGCTATGT	215	TTACGGATGG	331
TAACTCGCCT	469	TCTGGAGCCG	696	TTACTCTAGC	589
TAACTGGCGA	574	TCTTACGGAT	329	TTACTTCTGA	402
TAAGAGAATT	349	TGAAAGTAAA	85	TTATCCCGTG	234
TAAGATCCTT	155	TGAACGAAAT	806	TTATCTACAC	772
TAAGCATTGG	848	TGAAGATCAG	101	TTATGCAGTG	357
TAAGCCCTCC	752	TGAAGCCATA	503	TTATCCCTT	28
TAATAGACTG	613	TGAATGAAGC	499	TTATTGCTGA	682
TACACGACGG	777	TGACAACGAT	409	TTCAACATTT	7
TACACTATT	277	TGACACCA CG	527	TTCCAATGAT	190
TACATCGAAC	129	TGACAGTAAG	343	TTCCGGCAA	599
TACCAAACGA	511	TGACGCCGG	245	TTCCCTTTTT	31
TACGGATGGC	332	TGACTTGGTT	293	TTCCGGCTGG	667
TACTCACCAG	306	TGAGAGTTT	164	TTCCGTGTG	15
TACTCTAGCT	590	TGAGATAGGT	827	TTCTGTTTT	55
TACTTACTCT	586	TGAGCACTT	199	TTCGCCCCGA	172
TACTTCTGAC	403	TGAGCGTGGG	707	TTCTCAGAAT	284
TAGACAGATC	815	TGAGTACTCA	302	TTCTGACAC	406
TAGACTGGAT	616	TGAGTATTCA	1	TTCTGCGCTC	652
TAGCTCCCG	595	TGAGTGATAA	379	TTCTGCTATG	214
TAGGTGCTC	832	TGATAAAATCT	689	TTGACGCCG	244
TAGTTATCTA	769	TGATAACACT	383	TTGAGAGTTT	163
TATCATTGCA	725	TGATCGTTGG	479	TTGAGTACTC	301
TATCCCGTGT	235	TGATGAGCAC	196	TTGATCGTTG	478
TATCGTAGTT	764	TGATTAAGCA	844	TTGCACAACA	447
TATCTACACG	773	TGCACAAACAT	448	TTGCAGCACT	730
TATGCAGTGC	358	TGCACGAGTG	116	TTGCAGGACC	640
TATGGATGAA	800	TGCAGCAATG	542	TTGCCCTCCT	50
TATGTGGCGC	220	TGCAGCACTG	731	TTGCGAAC	561
TATTAACCTGG	571	TGCAGGACCA	641	TTGCGGATT	40
TATTATCCCG	232	TGCAGTGCTG	360	TTGCTCACCC	64
TATTCAACAT	5	TGCATAAAC	368	TTGCTGATAA	685
TATTCCCTT	29	TGCCTCACTG	836	TTGGGAACCG	485
TATTCTCAGA	282	TGCCTGCAGC	538	TTGGGTGCAC	111
TATTGCTGAT	683	TGCCTTCCTG	51	TTGGTTGAGT	297
TCAACAGCGG	145	TGCGCAAAC	562	TTAAAGTTC	207
TCAACATTTC	8	TGCGCTCGC	655	TTTATTGCTG	681
TCACAGAAA	316	TGCGGCATT	41	TTTCAATGA	189
TCACCAGTC	309	TGCGGCCAAC	392	TTTCCGTGTC	14
TCACCCAGAA	68	TGCTATGTGG	217	TTTCGCCCCG	171
TCACTGATTA	840	TGCTCACCA	65	TTTGACAAAC	446
TCGGTCGCCG	265	TGCTGAAGAT	98	TTTGCTTCC	49

K = 10, prefix : T

@read 1

GC	GCGGTATTATCCCGT	TGTCACGCCGGCAAGAG	TAACTCGTCGCCGCATA	ACTATTCTCAGAATGACTTG	TAAAGATGC	91	TCTAGCTTCC	593	TGTTTTGCT	59
+					TAAAGTTCTG	209	TCTCAACAGC	143	TTAACATTCT	208
IIIHIIHHHHHIHIGIHHHIHIIHHHGHI	GEIIIEGFHGI	IHHFGFIGIFGFI	HIFIHEFIFGGI	IFIGFIDIIG	TAAAGTTGCA	635	TCTCAGAATG	285	TTAACTGGCG	573
>blaTEM-1a					TAATCTGGA	692	TCTCGCGGT	717	TTAACGATTG	847
0	ATGAGTATTCAACATTCCGTGTC	CCCTT			TAACACTGCG	386	TCTGACAACG	407	TTAATAGACT	612
30	ATTCCCTTTTGCGGCAT	TTTGCCTT			TAACCAGAG	373	TCTGCGCTG	653	TTACATCGAA	128
60	GTTTTGCTCACCCAGAA	ACGCTGGT	GAAA		TAACCGCTT	436	TCTGCTATGT	215	TTACGGATGG	331
90	GTAAAAGATGCTGAAGAT	CAGTTGGT	GCA		TAAC	469	TCTGGAGCCG	696	TTACTCTAGC	589
120	CGAGTGGTTACATCGA	ACTGGATCT	CAAC		TAAGAGAATT	574	TCTTACGGAT	329	TTACTCTGA	402
150	AGCGGTAA	GATCCTGAGAG	TTTCGCCCC		TAAGATCCTT	349	TGAAAGTAAA	85	TTATCCCCTG	234
180	GAAGAACGTTTCCAAT	GATGAG	CACT	TTT	TAAGCATTG	155	TGAACGAAAT	806	TTATCTACAC	772
210	AAAGTTCTGCTATGTG	GGCGCG	TATTATCC		TAAGCCCTC	848	TGAAGATCAG	101	TTATGCAGTG	357
240	CGTGTGACGCCGGCAAGAG	CAACTCGGT			TAATAGACTG	752	TGAAGCCATA	503	TTATCCCTT	28
270	CGCCGCATA	ACACTATTCT	CAGAATGACTTG		TACACGACGG	613	TGAATGAAGC	499	TTATTGCTGA	682
300	GTTGAGTACTCACCA	GTCACAGAAA	AGCAT		TACACTATT	777	TGACAACGAT	409	TTCAACATT	7
330	CTTACGGATGGCAT	GACAGTAAGAGA	ATT		TACATCGAAC	277	TGACACCACG	527	TTCCAATGAT	190
360	TGCAGTGCTGCC	CATAACCATGAGT	GATAAC		TACCAAACGA	129	TGACAGTAAG	343	TTCCCGGCAA	599
390	ACTGCGGCCAA	CTTACTTCT	GACAACGATC		TACGGATGGC	511	TGACGCCGG	245	TTCCCTTTT	31
420	GGAGGACCGAAGGAG	CTAACCGCT	TTTTG		TACTCACCAG	332	TGACTTGGTT	293	TTCCGGCTGG	667
450	CACAA	CATGGGGATCATG	TAAC	TGCC	TACTCTAGCT	306	TGAGAGTTT	164	TTCCGTGTG	15
480	GATCGTTGGGAA	CCGGAGCTGA	ATGAAGCC		TACTTACTCT	590	TGAGATAGGT	827	TTCTGTTTT	55
510	ATACCAAACGAC	GAGCGTGACACCACGATG			TACTTCTGAC	586	TGAGCACTT	199	TTCGCCCCGA	172
540	CCTGCAGCAATGG	CAACACGTT	CGCGCAA		TAGACAGATC	403	TGAGCGTGGG	707	TTCTCAGAAT	284
570	CTATTAACTGGCGA	ACTACTTACT	CTAGCT		TAGACTGGAT	815	TGAGTACTCA	302	TTCTGACAC	406
600	TCCC	GGCAACAA	TTAATAGACTGG	ATGGAG	TAGCTTCCG	616	TGAGTATTCA	1	TTCTGCGCTC	652
630	GGGATAAAGTTG	CAGGACCA	TTCTGCGC		TAGCTCCG	595	TGAGTGATAA	379	TTCTGCTATG	214
660	TCGGCCCTCCGG	CTGGCTGG	TTATTGCT		TAGGTGCTC	832	TGATAAAATCT	689	TTGACGCCG	244
690	GATAAA	ATCTGGGAG	CCGGT	GAGCGTGGT	TAGTTATCTA	769	TGATAAACACT	383	TTGAGAGTTT	163
720	CGCGGTATCATTG	CAGCA	CTGGGGCCAGAT		TATCATTGCA	725	TGATCGTTGG	479	TTGAGTACTC	301
750	GGTAAGCCCTCCG	TATCGTAGTT	TATCTAC		TATCCCGTGT	235	TGATGAGCAC	196	TTGATCGTTG	478
780	ACGACGGGGAG	TCA	GCGACTATGGATGAA		TATCGTAGTT	764	TGATTAAGCA	844	TTGCACAACA	447
810	CGAAATAGACAGA	TGCTGAGA	TAGGTGCC		TATCTACACG	773	TGCAACACAT	448	TTGCAGCACT	730
840	TCACTGATTAA	AGCATTGGTAA			TATGCAGTGC	358	TGCACGAGTG	116	TTGCAGGACC	640
					TATGGATGAA	800	TGCAGCAATG	542	TTGCCTTCC	50
					TATGTGGCGC	220	TGCAGCACTG	731	TTGCGAAC	561
					TATTA	571	TGCAGGACCA	641	TTGCGGATT	40
					ACTCCG	232	TGCAGTGCTG	360	TTGCTACCC	64
					TATTAACAT	5	TGCCTAAACC	368	TTGCTGATAA	685
					TATTCCCTT	29	TGCCTCACTG	836	TTGGGACCG	485
					TATTCTCAGA	282	TGCCTGCAGC	538	TTGGGTGCAC	111
					TATTGCTGAT	683	TGCCTTCC	51	TTGGTTGAGT	297
					TCAACAGCGG	145	TGCGCAAAC	562	TTAAAGTTC	207
					TCAACATTTC	8	TGCGCTCGC	655	TTTATTGCTG	681
					TCACAGAAA	316	TGCGGCATT	41	TTTCAATGA	189
					TCACCAGTC	309	TGCGGCCAAC	392	TTTCCGTGTC	14
					TCACCCAGAA	68	TGCTATGTGG	217	TTTCGCCCCG	171
					TCACTGATTA	840	TGCTCACCA	65	TTTGACAAAC	446
					TCGGTCGCCG	265	TGCTGAAGAT	98	TTTGCTTCC	49

K = 10, prefix : T

@read 1

GCGCGGTATTATCCCGT	T	TAAAGATGC	91	TCTAGCTTCC	593	TGTTTTGCT	59
+ IIIHIIHHHHHIHIGIHHHIHIIHHHGHI	EIIIIEGFHGI	TAAAGTTCTG	209	TCTCAACAGC	143	TTAACATTCT	208
HIGEIIIIEGF	HFIGFIGFI	TAAAGTTGCA	635	TCTCAGAATG	285	TTAACTGGCG	573
HFIGFI	FFIGFIGFIDIIG	TAAATCTGGA	692	TCTCGCGGT	717	TTAACGATTG	847
>blaTEM-1a		TAACACTGCG	386	TCTGACAACG	407	TTAATAGACT	612
0 ATGAGTATTCAACATTCCGTGTCGCC	TT	TAACCAGAG	373	TCTGCGCTG	653	TTACATCGAA	128
30 ATTCCCTTTTGCGGCATTTGCCTTC	TT	TAACCGCTT	436	TCTGCTATGT	215	TTACGGATGG	331
60 GTTTTGCTCACCCAGAAACGCTGGT	GAAA	TAACTCGC	469	TCTGGAGCCG	696	TTACTCTAGC	589
90 GTAAAAGATGCTGAAGATCAGTTGG	TGCA	TAACGGCGA	574	TCTTACGGAT	329	TTACTCTGA	402
120 CGAGTGGTTACATCGAACTGGATCT	CAAC	TAAGAGAATT	349	TGAAAGTAAA	85	TTATCCCGTG	234
150 AGCGGTAAAGATCCTTGAGAGTTT	CGCCCC	TAAGATCCTT	155	TGAACGAAAT	806	TTATCTACAC	772
180 GAAGAACGTTTCCAATGATGAGCA	CTTT	TAAGCATTG	848	TGAAGATCAG	101	TTATGCAGTG	357
210 AAAGTTCTGCTATGTGCGCGG	TATTATCC	TAAGCCCTC	752	TGAAGCCATA	503	TTATCCCTT	28
240 CGTGTGACGCCGGCAAGAGCAACT	CGGT	TAATAGACTG	613	TGAATGAAGC	499	TTATTGCTGA	682
270 CGCCGCATAACACTATTCTCAGA	ATGACTT	TACACGACGG	777	TGACAACGAT	409	TTCAACATTT	7
300 GTT GAGTACTCACCA	GTCAG	TACACTATT	277	TGACACCACG	527	TTCCAATGAT	190
330 CTTACGGATGGCATGACAGTAAG	GAGAATT	TACATCGAAC	129	TGACAGTAAG	343	TTCCGGCAA	599
360 TGCAGTGCTGCCATAACCATGAG	TGATAAC	TACCAAACGA	511	TGACGCCGG	245	TTCCCTTTTT	31
390 ACTGCGGCCAACTTACTTCTGAC	AACGATC	TACGGATGGC	332	TGACTTGGTT	293	TTCCGGCTGG	667
420 GGAGGACCGAAGGAGCTAACCG	CTTTTG	TACTCACCAG	306	TGAGAGTTT	164	TTCCGTGTG	15
450 CACAA CATGGGGATCATGTAA	CTCGCCTT	TACTCTAGCT	590	TGAGATAGGT	827	TTCTGTGTTT	55
480 GATCGTTGGAACCGGAGCTGA	ATGAAGCC	TACTTACTCT	586	TGAGCACTT	199	TTCGCCCCGA	172
510 ATACCAAAACGACGAGCGTGA	CACACCAGATG	TACTTCTGAC	403	TGAGCGTGGG	707	TTCTCAGAAT	284
540 CCTGCAGCAATGGCAACACGTT	CGC	TAGACAGATC	815	TGAGTACTCA	302	TTCTGACAAAC	406
570 CTATTAAC TGCGAAACTTACT	TACTCTAGCT	TAGACTGGAT	616	TGAGTATTCA	1	TTCTGCGCTC	652
600 TCCC GGCAACATTAA	TAGACTGG	TAGCTCCCG	595	TGAGTGATAA	379	TTCTGCTATG	214
630 GCGGATAAAAGTTGCAGGACC	TTCTGCGC	TAGGTGCTC	832	TGATAAAATCT	689	TTGACGCCG	244
660 TC GGCCCTTCCGGCTGGCTGG	TTATTGCT	TAGTTATCTA	769	TGATAAACACT	383	TTGAGAGTTT	163
690 GATAAAATCTGGAGCCGGTGA	GCGTGGGTCT	TATCATTGCA	725	TGATCGTTGG	479	TTGAGTACTC	301
720 CGCGGTATATTGCA	GACTGGGCCAGAT	TATCCCGTGT	235	TGATGAGCAC	196	TTGATCGTTG	478
750 GGTAAAGCCCTCCGTATCGT	AGTTATCTAC	TATCGTAGTT	764	TGATTAAGCA	844	TTGCACAACA	447
780 ACGACGGGGAGTCAGGCA	ACTATGGATGAA	TATCTACACG	773	TGCACAAACAT	448	TTGCAGCACT	730
810 CGAAATAGACAGATCGCTGAG	ATAGGTGCC	TATGCAGTGC	358	TGCACGAGTG	116	TTGCAGGACC	640
840 TCACTGATTAAGCATTGGTAA	AA	TATGGCGC	220	TGCAGCAATG	542	TTGCCTTCCT	50
		TATTAACTGG	571	TGCAGCACTG	731	TTGCGAACAC	561
		TATTATCCCG	232	TGCAGGACCA	641	TTGCGGATT	40
		TATTCAACAT	5	TGCAGTGC	360	TTGCTCACCC	64
		TATTCCCTT	29	TGCCTAAACC	368	TTGCTGATAA	685
		TATTCTCAGA	282	TGCCTCACTG	836	TTGGGAAACCG	485
		TATTGCTAT	683	TGCCTGCAGC	538	TTGGGTGCAC	111
		TCAACAGCG	145	TGCCTTCCTG	51	TTGGTTGAGT	297
		TCAACATTTC	8	TGCGCAAAC	562	TTAAAGTTC	207
		TCACAGAAAA	316	TGCGCTCGC	655	TTTATTGCTG	681
		TCACCAGTC	309	TGCGGCATT	41	TTTCAATGA	189
		TCACCCAGAA	68	TGCGGCCAAC	392	TTTCCGTGTC	14
		TCACTGATTA	840	TGCTATGTGG	217	TTTCGCCCCG	171
		TCGGTCGCCG	265	TGCTCACCCA	65	TTTGACAAAC	446
				TGCTGAAGAT	98	TTTGCTTC	49

K = 10, prefix : T

@read 1

GCGCGGTATTATCCGTGTTGACGCCGGCAAGAGTAACTCGGTCGCCGCATAACTATTCTCAGAATGACTTG	TAAAAGATGC	91	TCTAGCTTCC	593	TGTTTTGCT	59
+	TAAAGTTCTG	209	TCTCAACAGC	143	TTAACATTCT	208
IIIHIIHHHHHIHIGIHHHIHIIHHGHIGEIIIEGFHGIIHIFGFIGIGFHIHEFIFGGIIFIGFIDIIG	TAAAGTTGCA	635	TCTCAGAAATG	285	TTAACTGGCG	573
>blaTEM-1a	TAAATCTGGA	692	TCTCGCGGTA	717	TTAACGATTG	847
0 ATGAGTATTCAACATTCCGTGTCGCCCTT	TAACACTGCG	386	TCTGACAACG	407	TTAAATAGACT	612
30 ATTCCCTTTTGCAGCATTTGCCTTCCT	TAACCATGAG	373	TCTGCGCTCG	653	TTACATCGAA	128
60 GTTTTGCTCACCCAGAAACGCTGGTGAAA	TAACCGCTT	436	TCTGCTATGT	215	TTACGGATGG	331
90 GTAAAAGATGCTGAAGATCAGTTGGGTGCA	TAACTCGCCT	469	TCTGGAGCCG	696	TTACTCTAGC	589
120 CGAGTGGTTACATCGAACTGGATCTAAC	TAACTGGCGA	574	TCTTACGGAT	329	TTACTTCTGA	402
150 AGCGGTAAAGATCCTTGAGAGTTTCGCC	TAAGAGAATT	349	TGAAAGTAAA	85	TTATCCCGTG	234
180 GAAGAACGTTTCCAATGATGAGCACTTT	TAAGATCCTT	155	TGAACGAAAT	806	TTATCTACAC	772
210 AAAGTTCTGCTATGTGGCGCGTATTATCC	TAAGCATTGG	848	TGAAGATCAG	101	TTATGCAGTG	357
240 CGTGTGACGCCGGCAAGAGCAACTCGGT	TAAGCCCTCC	752	TGAAGCCATA	503	TTATCCCTT	28
270 CGCCGCATAACACTATTCTCAGAATGACTTG	TAATAGACTG	613	TGAATGAAGC	499	TTATTGCTGA	682
300 GTTGAGTACTCACCAAGTCACAGAAAAGCAT	TACACGACGG	777	TGACAACGAT	409	TTCAACATTT	7
330 CTTACGGATGGCATGACAGTAAGAGAATTA	TACACTATT	277	TGACACCACG	527	TTCCAATGAT	190
360 TGCAGTGCTGCCATAACCATGAGTGATAAC	TACATCGAAC	129	TGACAGTAAG	343	TTCCGGCAA	599
390 ACTGCGGCCAACTTACTTCTGACAACGATC	TACCAAACGA	511	TGACGCCGG	245	TTCCCTTTTT	31
420 GGAGGACCGAAGGAGCTAACCGCTTTTG	TACGGATGGC	332	TGACTTGGTT	293	TTCCGGCTGG	667
450 CACAACATGGGGATCATGTAACTGCCCTT	TACTCACCAG	306	TGAGAGTTT	164	TTCCGTGTCG	15
480 GATCGTTGGAACCGGAGCTGAATGAAGCC	TACTCTAGCT	590	TGAGATAGGT	827	TTCTGTTTT	55
510 ATACCAAACGACGAGCGTACACACCGATG	TACTTACTCT	586	TGAGCACTT	199	TTCGCCCCGA	172
540 CCTGCAGCAATGGCAACACGTTGCGCAA	TACTTCTGAC	403	TGAGCGTGGG	707	TTCTCAGAAT	284
570 CTATTAACCTGGCGAACTACTTACTCTAGCT	TAGACAGATC	815	TGAGTACTCA	302	TTCTGACAC	406
600 TCCCAGCAACAATTAAAGACTGGATGGAG	TAGACTGGAT	616	TGAGTATTCA	1	TTCTGCGCTC	652
630 GCGGATAAAAGTTGCAGGACCACTCTGCGC	TAGCTCCCG	595	TGAGTGATAA	379	TTCTGCTATG	214
660 TCAGGCCCTCCGGCTGGCTGGTTATTGCT	TAGGTGCTC	832	TGATAAAATCT	689	TTGACGCCG	244
690 GATAAAATCTGGAGCCGGTGAGCGTGGGTCT	TAGTTATCTA	769	TGATAACACT	383	TTGAGAGTTT	163
720 CGCGGTATCATTGCACTGGGCCAGAT	TATCATTGCA	725	TGATCGTTGG	479	TTGAGTACTC	301
750 GGTAAAGCCCTCCGTATCGTAGTTATCTAC	TATCCCGTGT	235	TGATGAGCAC	196	TTGATCGTTG	478
780 ACGACGGGGAGTCAGGCAACTATGGATGAA	TATCGTAGTT	764	TGATTAAGCA	844	TTGCACAACA	447
810 CGAAATAGACAGATCGCTGAGATAGGTGCC	TATCTACACG	773	TGCAACACAT	448	TTGCAGCACT	730
840 TCACTGATTAAGCATTGGTAA	TATGCAGTGC	358	TGCACGAGTG	116	TTGCAGGACC	640
	TATGGATGAA	800	TGCAGCAATG	542	TTGCCTTCCT	50
	TATGTGGCGC	220	TGCAGCACTG	731	TTGCGAAC	561
	TATTAACTGG	571	TGCAGGACCA	641	TTGCGGCATT	40
	TATTATCCCG	232	TGCAGTGCTG	360	TTGCTACCC	64
	TATTCAACAT	5	TGCCTAAACC	368	TTGCTGATAA	685
	TATTCCTTT	29	TGCCTCACTG	836	TTGGGAACCG	485
	TATTCTCAGA	282	TGCCTGCAGC	538	TTGGGTGCAC	111
	TATTGCTGAT	683	TGCCTTCCTG	51	TTGGTTGAGT	297
	TCAACAGCGG	145	TGCGCAAAC	562	TTAAAGTTC	207
	TCAACATTTC	8	TGCGCTCGC	655	TTTATTGCTG	681
	TCACAGAAA	316	TGCGGCATT	41	TTTCAATGA	189
	TCACCACTCA	309	TGCGGCCAAC	392	TTTCCGTGTC	14
	TCACCCAGAA	68	TGCTATGTGG	217	TTTCGCCCCG	171
	TCACTGATTA	840	TGCTCACCCA	65	TTTGACAAAC	446
	TCGGTCGCCG	265	TGCTGAAGAT	98	TTTGCTTCC	49

K = 10, prefix : T

@read 1

GCGCGGTATTATCCGTGTTGACGCCGGCAAGAGTAACTCGGTCGCCGCATAACACTATTCTCAGAATGACTTG

+

IIIHIIHHHHHIHIGIHHHIHIIHHGHIGEIIIEGFHGIIHIFGFIGIGFHIHEFIFGGIIFIGFIDIIG

>blaTEM-1a

0	ATGAGTATTCAACATTCCGTGTCGCCCTT
30	ATTCCCTTTTGC GG CATTTCGCCTTCCT
60	GTTTTGCTCACCCAGAACGCTGGTGAAA
90	GTAAAAGATGCTGAAGATCAGTTGGGTGCA
120	CGAGTGGTTACATCGAACTGGATCTCAAC
150	AGCGGTAAAGATCCTTGAGAGTTTCGCC
180	GAAGAACGTTTCCAATGATGAGCACTTT
210	AAAGTTCTGCTATGTGGCGGGTATTATCC
240	CGTGTGACGCCGGCAAGAGCAACTCGGT
270	CGCCGCATAACACTATTCTCAGAATGACTTG
300	GTTGAGTACTCACCAAGTCACAGAAAAGCAT
330	CTTACGGATGGCATGACAGTAAGAGAATTA
360	TGCAGTGCTGCCATAACCATGAGTGATAAC
390	ACTGCGGCCAACTTACTTCTGACAACGATC
420	GGAGGACCGAAGGAGCTAACCGCTTTTG
450	CACAAACATGGGGATCATGTAACTCGCC
480	GATCGTTGGGACCGGAGCTGAATGAAGCC
510	ATACCAAACGACGAGCGTACACCGATG
540	CCTGCAGCAATGGCAACACGTTGCGCAA
570	CTATTAACCTGGCGAACTACTTACTCTAGCT
600	TCCC GG CAACAATTAA TAGACTGGATGGAG
630	GCGGATAAAGTTGCAGGACCACTTCTGCGC
660	TCGGCCCTCCGGCTGGCTGGTTATTGCT
690	GATAAAATCTGGAGCCGGTGAGCGTGGTCT
720	CGCGGTATCATTGCA GCACTGGGCCAGAT
750	GGTAAGCCCTCCGTATCGTAGTTATCTAC
780	ACGACGGGGAGTCAGGCAACTATGGATGAA
810	CGAAATAGACAGATCGCTGAGATAGGTGCC
840	TCACTGATTAAGCATTGGTAA

TAAAAGATGC	91	TCTAGCTTCC	593	TGTTTTGCT	59
TAAAGTTCTG	209	TCTCAACAGC	143	TTAAAGTTCT	208
TAAAGTTGCA	635	TCTCAGAATG	285	TTAACTGGCG	573
TAAATCTGGA	692	TCTCGCGGT	717	TTAACGATTG	847
TAACACTGCG	386	TCTGACAACG	407	TTAATAGACT	612
TAACCAGAG	373	TCTGCGCTG	653	TTACATCGAA	128
TAACCGCTT	436	TCTGCTATGT	215	TTACGGATGG	331
TAACTCGC	469	TCTGGAGCCG	696	TTACTCTAGC	589
TAACGGCGA	574	TCTTACGGAT	329	TTACTCTGA	402
TAAGAGAATT	349	TGAAAGTAAA	85	TTATCCCGT	234
TAAGATCCTT	155	TGAACGAAAT	806	TTATCTACAC	772
TAAGCATTG	848	TGAAGATCAG	101	TTATGCAGTG	357
TAAGCCCTCC	752	TGAAGCCATA	503	TTATCCCTT	28
TAATAGACTG	613	TGAATGAAGC	499	TTATTGCTGA	682
TACACGACGG	777	TGACAACGAT	409	TTCAACATTT	7
TACACTATT	277	TGACACCACG	527	TTCCAATGAT	190
TACATCGAAC	129	TGACAGTAAG	343	TTCCGGCAA	599
TACCAAACGA	511	TGACGCCGG	245	TTCCCTTTT	31
TACGGATGGC	332	TGACTTGGTT	293	TTCCGGCTGG	667
TACTCACCAG	306	TGAGAGTTT	164	TTCCGTGTCG	15
TACTCTAGCT	590	TGAGATAGGT	827	TTCTGTTTT	55
TACTTACTCT	586	TGAGCACTT	199	TTGCCCGA	172
TACTTCTGAC	403	TGAGCGTGGG	707	TTCTCAGAAT	284
TAGACAGATC	815	TGAGTACTCA	302	TTCTGACAC	406
TAGACTGGAT	616	TGAGTATTCA	1	TTCTGCGCTC	652
TAGCTCCCG	595	TGAGTGATAA	379	TTCTGCTATG	214
TAGGTGCCTC	832	TGATAAAATCT	689	TTGACGCCG	244
TAGTTATCTA	769	TGATAACACT	383	TTGAGAGTT	163
TATCATTGCA	725	TGATCGTTGG	479	TTGAGTACTC	301
TATCCCGTGT	235	TGATGAGCAC	196	TTGATCGTTG	478
TATCGTAGTT	764	TGATTAAGCA	844	TTGCACAACA	447
TATCTACACG	773	TGCACAAACAT	448	TTGCAGCACT	730
TATGCAGTGC	358	TGCACGAGTG	116	TTGCAGGACC	640
TATGGATGAA	800	TGCAGCAATG	542	TTGCCTTCCT	50
TATGTGGCGC	220	TGCAGCACTG	731	TTGCGAAC	561
TATTAACTGG	571	TGCAGGACCA	641	TTGCGGCATT	40
TATTATCCCG	232	TGCAGTGCTG	360	TTGCTCACCC	64
TATTCAACAT	5	TGCCTAAACC	368	TTGCTGATAA	685
TATTCCTT	29	TGCCTCACTG	836	TTGGGAACCG	485
TATTCTCAGA	282	TGCCTGCAGC	538	TTGGGTGCAC	111
TATTGCTGAT	683	TGCCTTCCTG	51	TTGGTTGAGT	297
TCAACAGCGG	145	TGCGCAAAC	562	TTAAAGTTC	207
TCAACATTTC	8	TGCGCTCGC	655	TTTATTGCTG	681
TCACAGAAA	316	TGCGGCATT	41	TTTCAATGA	189
TCACCAGTC	309	TGCGGCCAAC	392	TTTCCGTGTC	14
TCACCCAGAA	68	TGCTATGTGG	217	TTTCGCCCCG	171
TCACTGATTA	840	TGCTCACCA	65	TTTGACAAAC	446
TCGGTCGCCG	265	TGCTGAAGAT	98	TTTGCTTCC	49

@read 1
GCCGGTATTATCCGTGTTACGCCGGCAAGAGTAACTCGTCGCCCATACACTATTCTCAGAATGACTTG
+
IIIIHIIHHHHHIHIGIHHHIHIIHHGHIGEIIIIEFHGIHIFGFIGIGFHIHEFIFGGIIFIGFIDIIG

@read 2
GAATGACTTGCCTTGAGTACTCACCAAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCA
+
IGIIH=IGGIDIIAIHIIIIIGEIIGIIHIAIIIHIIID@ICE@IIC?FIB>AA=DEIIDBIII?D@IFG

@read 3
TAACACTGCCAACCTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTGCACAACATTG
+
IGIIIHIGGIHIIHHHFHFHGGFDIIHIFHIDIDIFIHHGGCFIIFIIAHGAIIIDIIHFIIGGFIEI

@read 4
CCACTTCTGCCTCGGCCATCCGGCTGGCTGGTTATTGCTGATAATCTGGAGCCGGTGAGCGTGGTCTCG
+
GIEHGHGGFIIFGIIIEIHIGAIDIIF@FDFIEGIHIIIIHIIIIHBADHIBHIFIFII@AFFIEHI>F

@read 5
TCCTTGAGAGTTTCGCCCGAAGAACGTTCCAATGATGAGCACTTTAAAGTTCTGCTATGTGGCGCGTA
+
GEIIBEEIIHEEGGIIFHCEFHIIHIGIIFIIIFHIIIDEIIHEIHIIDDBHI=IBIHGIBIIIBD<HGI@IE

@read 6
GGCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTGCACAACATGGGAATCATG
+
IIIIHIIICDEIEIFIHIIICIHFIEIIIDIIIGDCEIIIEGIBIHEEBIIGFGGEIIAIGEIIIAII=DHGH

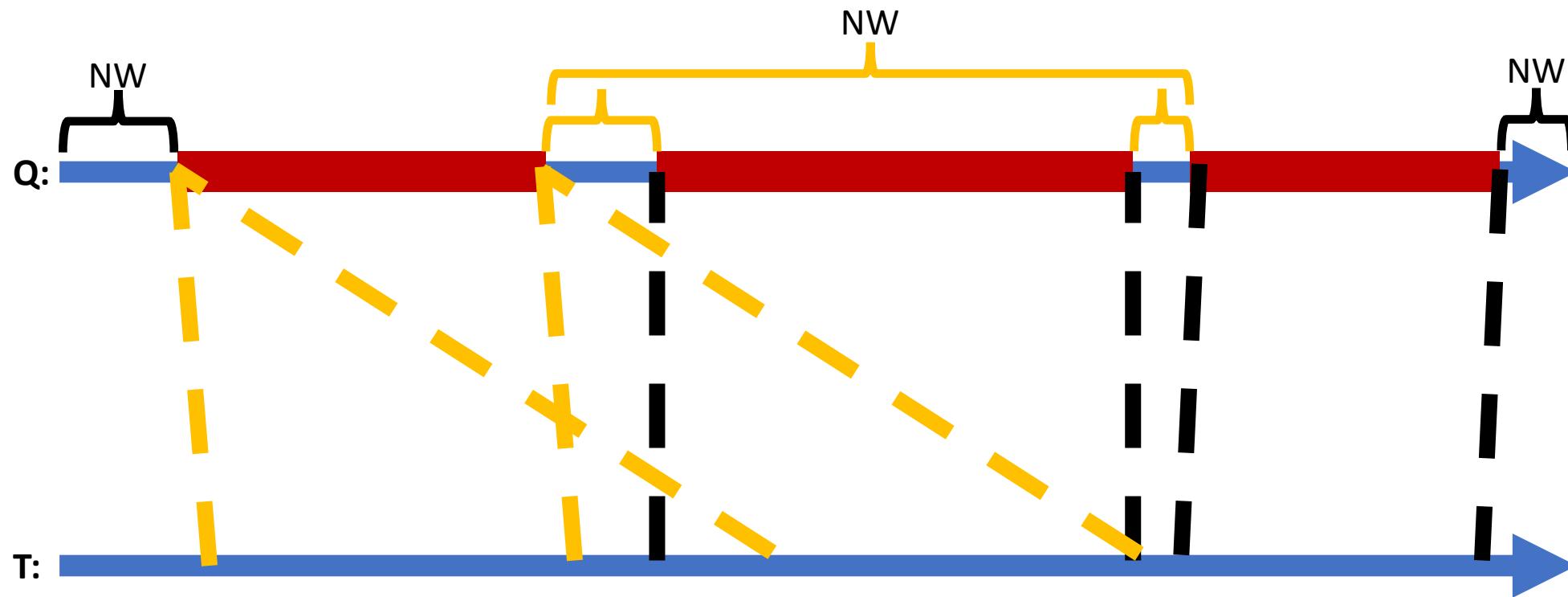
@read 7
ACTCTAGCTCCGGCAACAATTAATAGACTGGATGGAGGCGGATTAAGTTGCAGGACCACTCTCGCTCGC
+
IIIDICDHFIHHIIIIICECIEDIAEIEHIIGBDIIGFIIGIIIBIEIAFG!IEDBHIIGBIII?CI

@read 8
CCATCCGGCTGGCTGGTTATTGCTGATAATCTGGAGCCGGTGAGCGTGGTCTCGCGTATTCATTGAGCAC
+
FGIFIGIIFIIGDHIBIIHIEIGFFIIIIIGFHIH@IACIIHIIII?HIHIDFIBCIFIHIIIBH?III

>blaTEM-1a

0 ATGAGTATTCAACATTCCGTGTCGCCCTT
30 ATTCCCTTTTGCGGCATTTGCCCTCCT
60 GTTTTGCTCACCCAGAAACGCTGGTAAA
90 GTAAAAGATGCTGAAGATCAGTTGGTGCA
120 CGAGTGGGTTACATCGAACTGGATCTCAAC
150 AGCGGTAAGATCCTTGAGAGTTTCGCC
180 GAAGAACGTTTCCAATGATGAGCACTTT
210 AAAGTTCTGCTATGTGGCGCGTATTATCC
240 CGTGGTACGCCGGCAAGAGTAACTCGGT
270 CGCCGCATAACTATTCTCAGAATGACTTG
300 CTTGAGTACTCACCAAGTCACAGAAAAGCAT
330 CTTACGGATGGCATGACAGTAAGAGAATTA
360 TGCAGTGCTGCCATAACCAGTGAATGATAAC
390 ACTCGGCCAACCTACTTCTGACAACGATC
420 GGAGGACCGAAGGAGCTAACCGCTTTTG
450 CACAACATGGGATCATGTAACTCGCCTT
480 GATCGTGGGAACCGGAGCTGAATGAAGCC
510 ATACCAAACGACGAGCGTGACACCACGATG
540 CCTGCAGCAATGGCAACAACGTTGCGAAA
570 CTATTAACTGGCGAACTACTTACTCTAGCT
600 TCCCGGCAACAATTAATAGACTGGATGGAG
630 GCAGGATTAGTTGCAGGACCACTCTGCG
660 TCAGGCCATCCGGCTGGCTGGTTATTGCT
690 GATAAATCTGGAGCCGGTGAGCGTGGTCT
720 CGCGGTATCATTGCAGCACTGGGGCCAGAT
750 GGTAAGCCCTCCGTATCGTAGTTATCTAC
780 ACGACGGGGAGTCAGGCAACTATGGATGAA
810 CGAAATAGACAGATCGCTGAGATAGGTGCC
840 TCACTGATTAAGCATTGGTAA

There is some more to it...



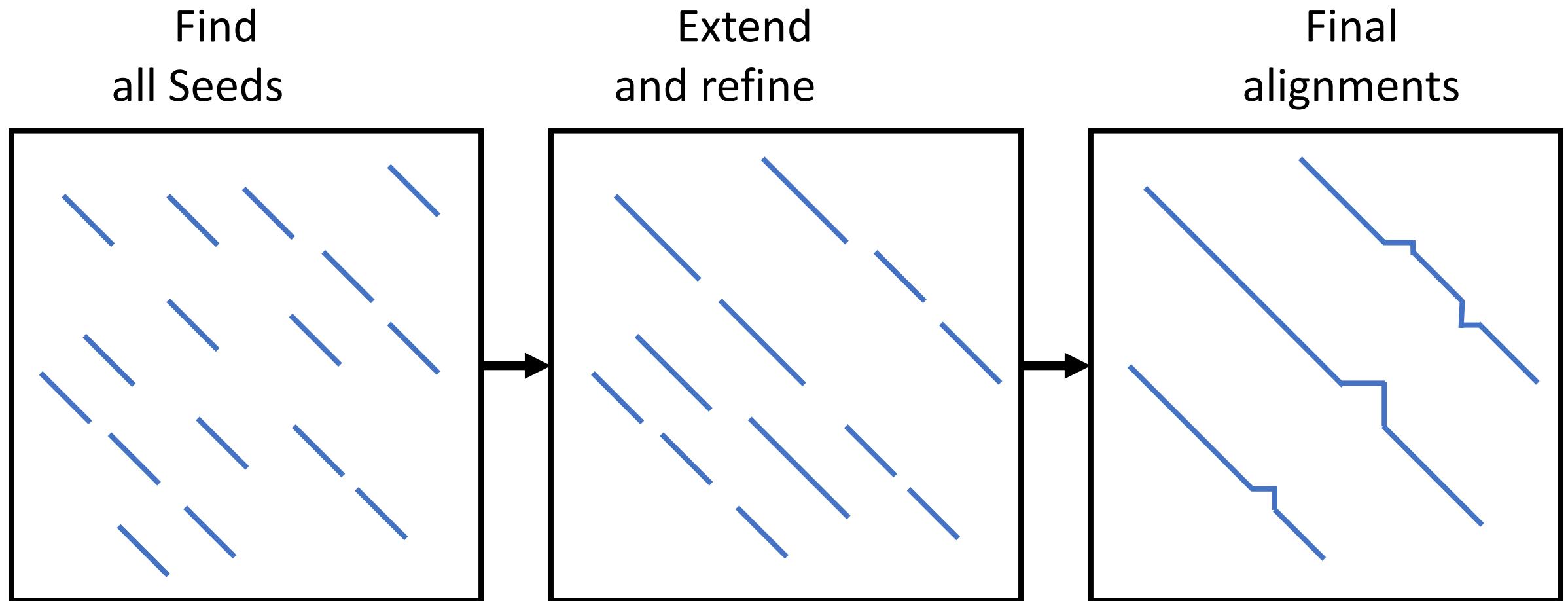
Chaining, 90's

$$f(i) = \max \left\{ \max_{i > j \geq 1} \{ f(j) + \alpha(j, i) - \beta(j, i) \}, w_i \right\}$$

$$\alpha(j, i) = \min \{ \min \{ y_i - y_j, x_i - x_j \}, w_i \}$$

$$\beta(j, i) = \gamma_c ((y_i - y_j) - (x_i - x_j))$$

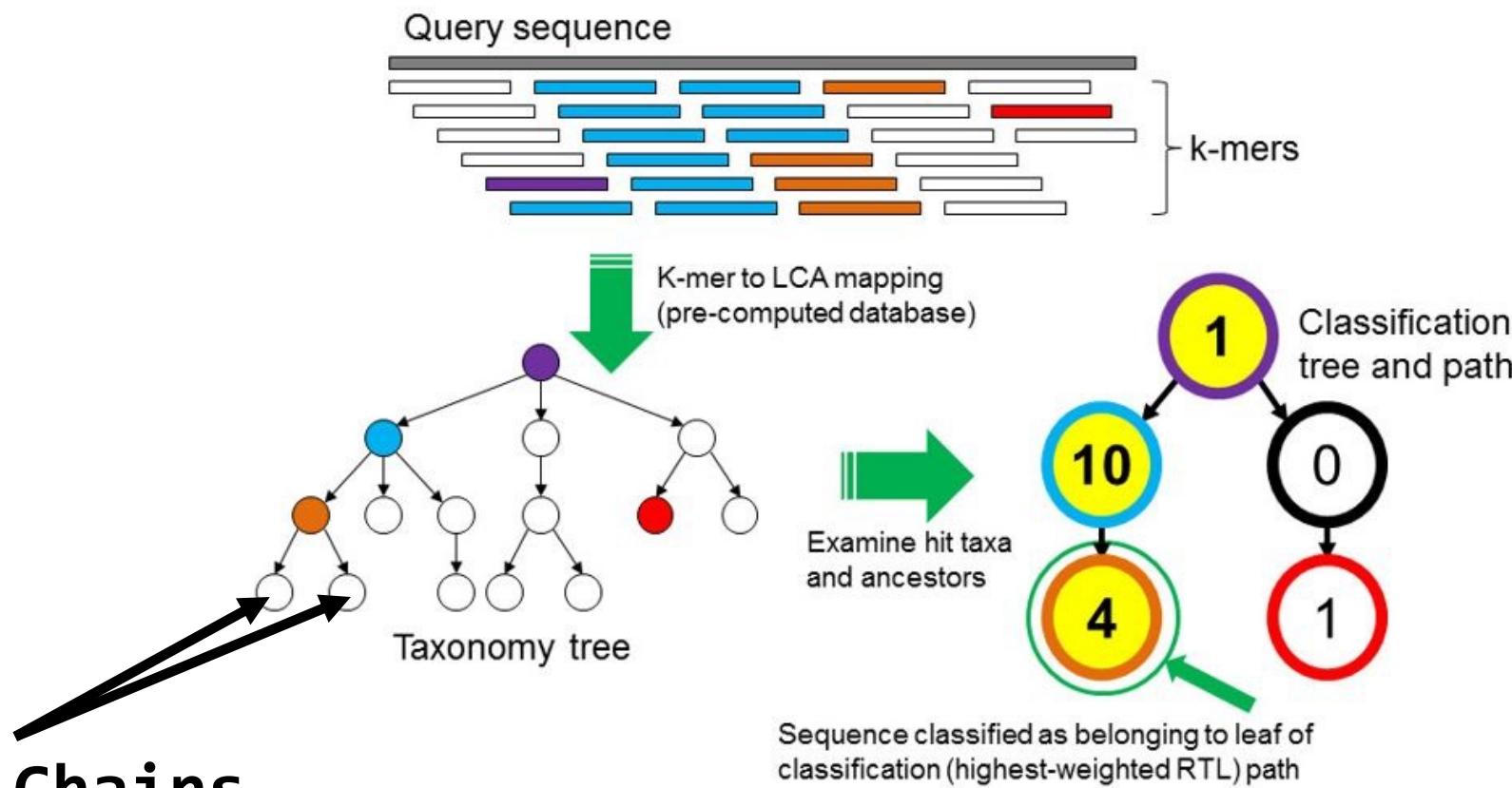
Mapping and Alignment



The redundant sequences

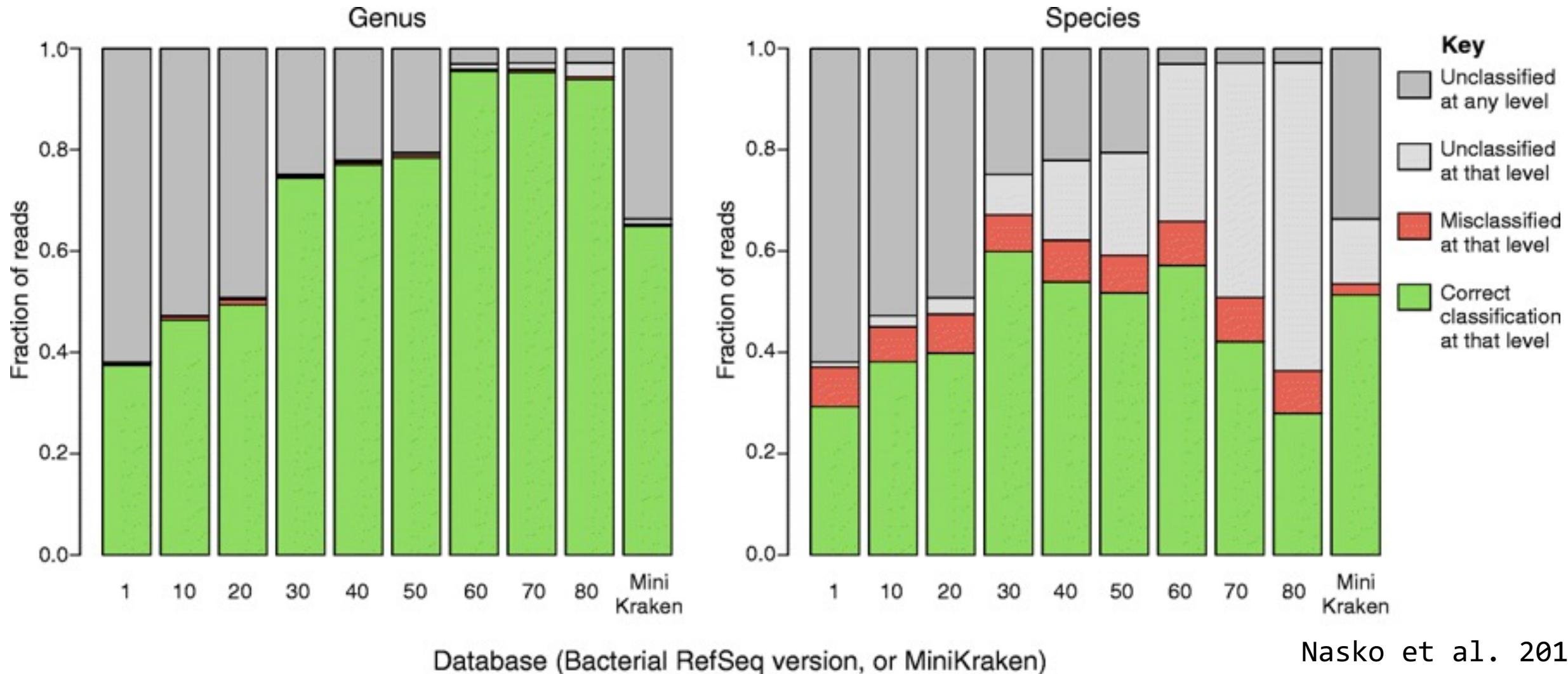
- LCA.
- Use the mapping and alignment methods just introduced.
- Something else?

LCA, 2014

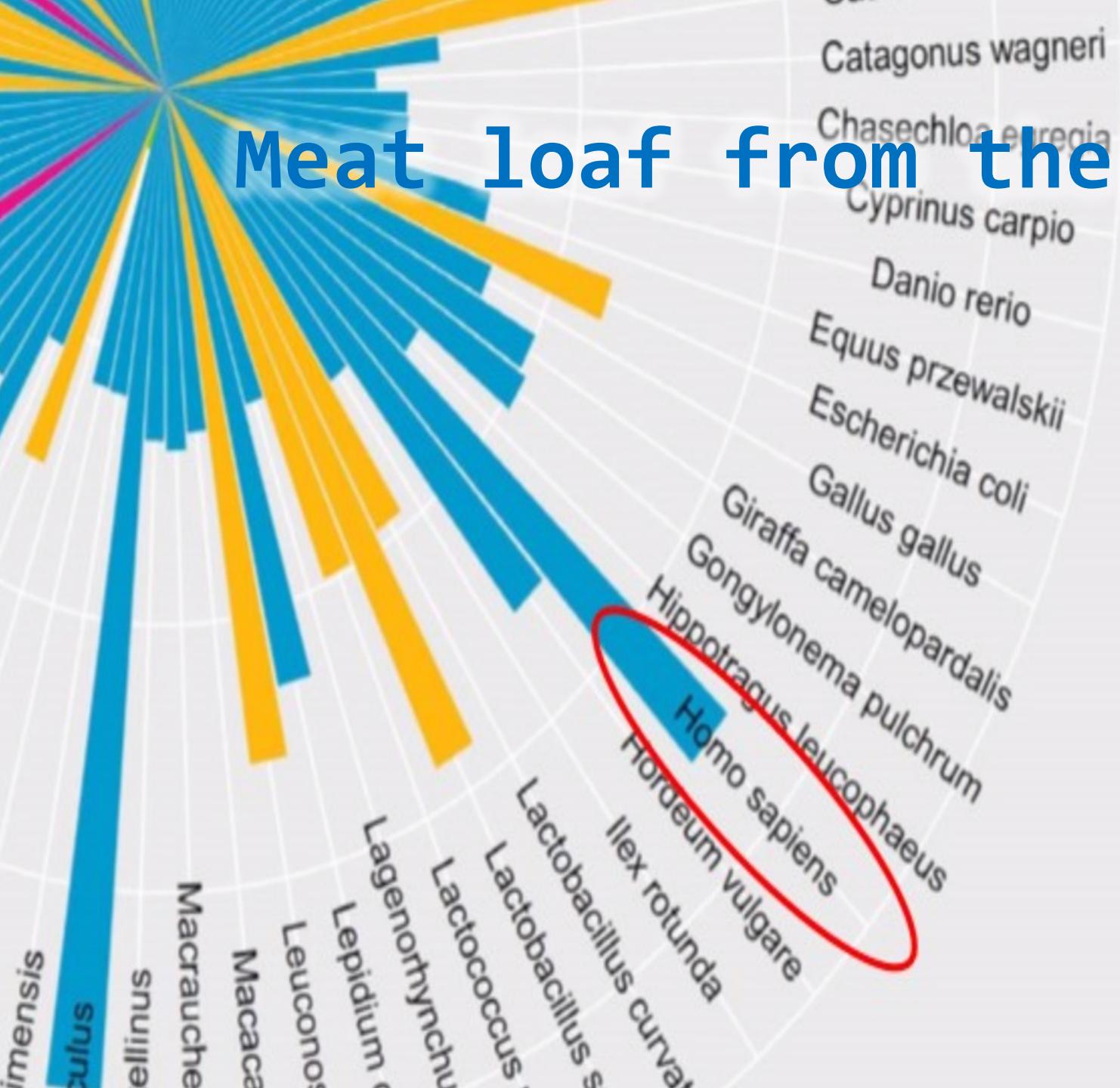


Equal Chains

RefSeq database growth and the LCA-algorithm, 2018



Meat loaf from the FLI canteen



- 183 blaTEM variants
 - Max distance of 38 SNPs
 - 150 bp subsequence from the middle of blaTEM-1A:
GTCCTGCAACTTATCCGCCTCCATCCAGTCTATTAAATTGTTGCCGGGAAGCTAGAGTAAGTAGTCGCCAGTTA
ATAGTTGCGAACGTTGCCATTGCTGCAGGCATCGTGGTGTACGCTCGTCGGTATGGCTTCATTCA
 - 131 perfect matches.

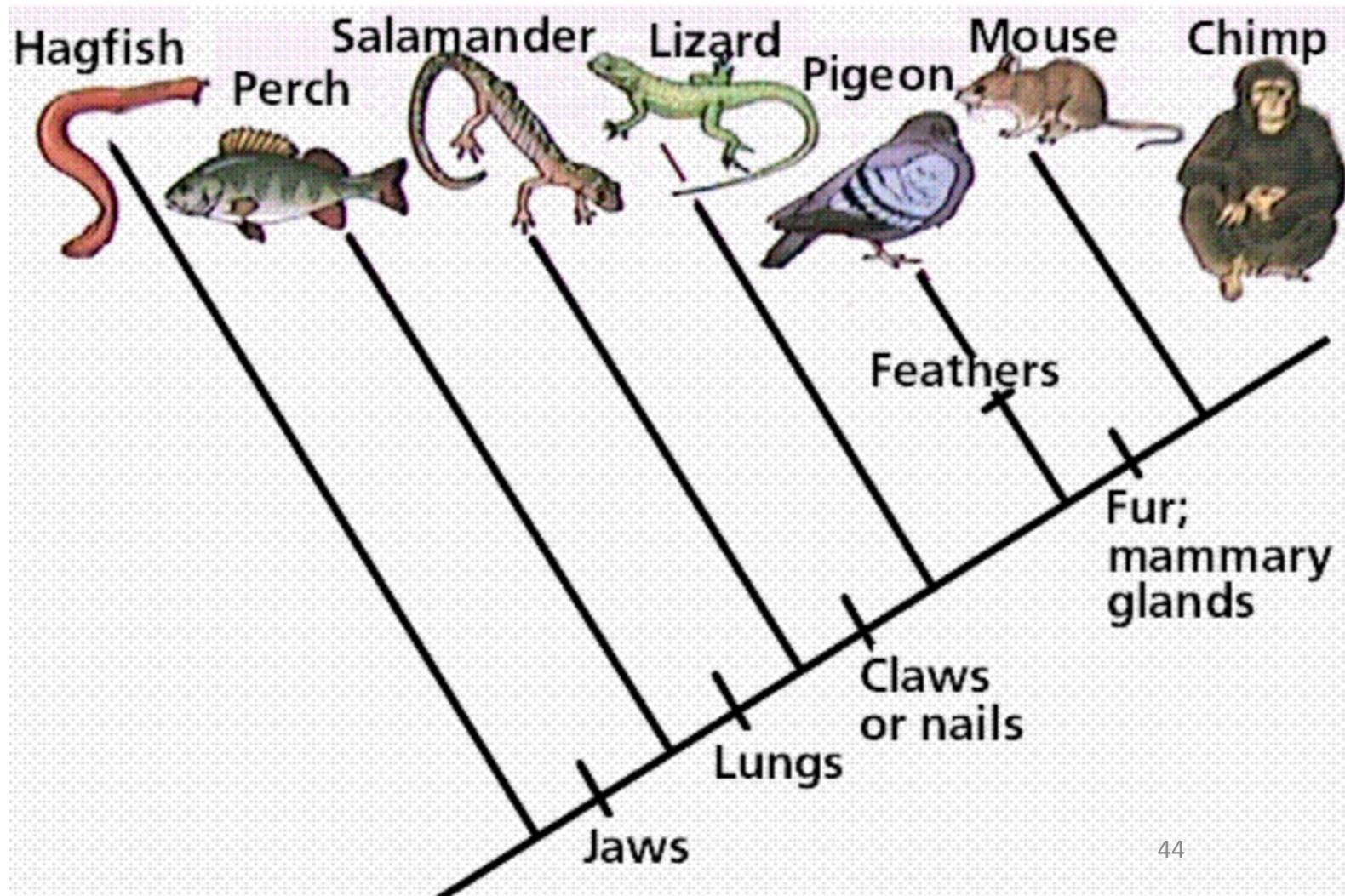
ConClave, 2018



Find best candidates

$$T_m \in \operatorname{argmax}_{i \in T} \{ f(t_i) \}$$

- Feathers
 - $T_m \in \{\text{Pigeon}\}$
- Jaws
 - $T_m \in \{\text{Perch, Salamander, Lizard, Pigeon, Mouse, Chimp}\}$
- Lungs
 - $T_m \in \{\text{Salamander, Lizard, Pigeon, Mouse, Chimp}\}$



Calculate the total score of all reads against the best candidates

$$C(t) = \sum_{k \in K} \max \begin{cases} f(t_k) & \tau \leq f(t_k) \\ 0 & \text{else} \end{cases}$$

- Feathers
 - $T_m \in \{\text{Pigeon}\}$
- Jaws
 - $T_m \in \{\text{Perch, Salamander, Lizard, Pigeon, Mouse, Chimp}\}$
- Lungs
 - $T_m \in \{\text{Salamander, Lizard, Pigeon, Mouse, Chimp}\}$

t	C(t)
Hagfish	0
Perch	1
Salamander	2
Lizard	2
Pigeon	3
Mouse	2
Chimp	2

Calculate the total score of all reads against the best candidates

$$S_k \in \operatorname{argmax}_{i \in T_m} \{ C(t_i) \}$$

- Feathers
 - $T_m \in \{\text{Pigeon}\}$
 - $S_k \in \{\text{Pigeon}\}$
- Jaws
 - $T_m \in \{\text{Perch, Salamander, Lizard, Pigeon, Mouse, Chimp}\}$
 - $S_k \in \{\text{Pigeon}\}$
- Lungs
 - $T_m \in \{\text{Salamander, Lizard, Pigeon, Mouse, Chimp}\}$
 - $S_k \in \{\text{Pigeon}\}$

t	C(t)
Hagfish	0
Perch	1
Salamander	2
Lizard	2
Pigeon	3
Mouse	2
Chimp	2

And now you...

$$T_m \in \operatorname{argmax}_{i \in T} \{ f(t_i) \} \quad (1)$$

$$C(t) = \sum_{k \in K} \max \begin{cases} f(t_k) & \tau \leq f(t_k) \\ 0 & \text{else} \end{cases} \quad (2)$$

$$S_k \in \operatorname{argmax}_{i \in T_m} \{ C(t_i) \} \quad (3)$$

Query sequence	$f(t)$	T_m	S_k
GCGCGGTATTATCCCGTGGACGCCGGCAAGAGTAACTCGGTCGCCGCATAACTATTCTCAGAATGACTTG	64	2,3,12	
GAATGACTTGCTTGAGTACTCACCAAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCA	63	1,2,3,7,8,9	
TAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTGACAAACATTG	72	1,2,3,4,5,6,8,9,11,12	
CCACTTCTGCGCTCGGCCATCCGGCTGGCTGGTTATTGCTGATAAAATCTGGAGCCGGTAGACGTGGTCTCG	66	3,5,8	
TCCTTGAGAGTTTCGCCCCGAAGAACGTTTCCAATGATGAGCACTTAAAGTTCTGCTATGTGGCGCGGT	74	2,5,12	
GGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTGACAAACATTGGAAATCATG	63	1,2,3,4,5,6,8,9,11,12	
ACTCTAGCTTCCC GGCAACAATTAATAGACTGGATGGAGGCGGATTAAGTTGCAGGACCCTCTGCGCTCGC	39	1,2,4,5,6,7,8,9,10,11	
CCATCCGGCTGGCTGGTTATTGCTGATAAACTGGAGCCGGTAGACGTGGTCTCGGGTATCATTGCAGCAC	71	2,5,12	

Templates:

- 01 blaTEM-1A_4_HM749966
- 02 blaTEM-2_1_X54606
- 03 blaTEM-2_3_AJ251946
- 04 blaTEM-3_1_X64523
- 05 blaTEM-6_1_X57972
- 06 blaTEM-8_1_X65252
- 07 blaTEM-10_1_AF093512
- 08 blaTEM-11_1_AY874537
- 09 blaTEM-12_1_M88143
- 10 blaTEM-15_1_AM849805
- 11 blaTEM-16_1_X65254
- 12 blaTEM-17_1_Y14574

ConClave Scores:

$$T_m \in \operatorname{argmax}_{i \in T} \{ f(t_i) \} \quad (1)$$

$$C(t) = \sum_{k \in K} \max \begin{cases} f(t_k) & \tau \leq f(t_k) \\ 0 & \text{else} \end{cases} \quad (2)$$

$$S_k \in \operatorname{argmax}_{i \in T_m} \{ C(t_i) \} \quad (3)$$

Predicted Template(s): _____

Query sequence	f(t)	T _m	S _k
GCGCGGTATTATCCCGTGGACGCCGGCAAGAGTAACTCGGTCGCCGCATAACTATTCTCAGAATGACTTG	64	2,3,12	
GAATGACTTCCTTCACTTACCTCACCAAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCA	63	1,2,3,7,8,9	
TAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGGAGCTAACCGCTTTTGACAAACATTG	72	1,2,3,4,5,6,8,9,11,12	
CCACTTCTGCGCTCGGCCATCCGGCTGGCTGGTTATTGCTGATAAAATCTGGAGCCGGTAGACGTGGCTCG	66	3,5,8	
TCCTTGAGAGTTTCGCCCCGAAGAACGTTTCCAATGATGAGCAGCTTTAAAGTTCTGCTATGTGGCGCGGT	74	2,5,12	
GGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGGAGCTAACCGCTTTTGACAAACATTGGGAATCATG	63	1,2,3,4,5,6,8,9,11,12	
ACTCTAGCTTCCC GGCAACAATTAAAGACTGGATGGAGGCGGATTAAGTTGCAGGACCCTCTGCGCTCGC	39	1,2,4,5,6,7,8,9,10,11	
CCATCCGGCTGGCTGGTTATTGCTGATAAACTGGAGCCGGTGAGCGTGGGTCTCGGGTATCATTGCAGCAC	71	2,5,12	

Templates:

- 01 blaTEM-1A_4_HM749966
- 02 blaTEM-2_1_X54606
- 03 blaTEM-2_3_AJ251946
- 04 blaTEM-3_1_X64523
- 05 blaTEM-6_1_X57972
- 06 blaTEM-8_1_X65252
- 07 blaTEM-10_1_AF093512
- 08 blaTEM-11_1_AY874537
- 09 blaTEM-12_1_M88143
- 10 blaTEM-15_1_AM849805
- 11 blaTEM-16_1_X65254
- 12 blaTEM-17_1_Y14574

ConClave Scores:

- _____
- 64
- 64
- _____
- _____
- _____
- _____
- _____
- _____
- _____
- _____
- 64

$$T_m \in \operatorname{argmax}_{i \in T} \{ f(t_i) \} \quad (1)$$

$$C(t) = \sum_{k \in K} \max \begin{cases} f(t_k) & \tau \leq f(t_k) \\ 0 & \text{else} \end{cases} \quad (2)$$

$$S_k \in \operatorname{argmax}_{i \in T_m} \{ C(t_i) \} \quad (3)$$

Predicted Template(s): _____

Query sequence	$f(t)$	T_m	S_k
GCGCGCTATTATCCCGTGTGACGCCGGCAAGAGTAACTCGGTCGCCGCATAACTATTCTCAGAACTGACTTG	64	2,3,12	
GAATGACTTGCTTGAGTACTCACAGACTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCA	63	1,2,3,7,8,9	
TAACACTGCCCGAACTTACTTCTGACAACGATCGGAGGACCGAAGGGAGCTAACCGCTTTTGACAAACATTG	72	1,2,3,4,5,6,8,9,11,12	
CCACTTCTGCGCTCGGCCCATCCGGCTGGCTGGTTATTGCTGATAAAATCTGGAGCCGGTGAGCGTGGGTCTG	66	3,5,8	
TCCTTGAGAGTTTCGCCCCGAAGAACGTTTCCAATGATGAGCAGCTTTAAAGTTCTGCTATGTGGCGCGGT	74	2,5,12	
GGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGGAGCTAACCGCTTTTGACAAACATTGGGAATCATG	63	1,2,3,4,5,6,8,9,11,12	
ACTCTAGCTTCCGGCAACAATTAAATAGACTGGATGGAGGCGGATTAAGTTGCAGGACCCTCTGCGCTCGC	39	1,2,4,5,6,7,8,9,10,11	
CCATCCGGCTGGCTGGTTATTGCTGATAAACTGGAGCCGGTGAGCGTGGGTCTCGGGTATCATTGCAGCAC	71	2,5,12	

Templates:

- 01 blaTEM-1A_4_HM749966
- 02 blaTEM-2_1_X54606
- 03 blaTEM-2_3_AJ251946
- 04 blaTEM-3_1_X64523
- 05 blaTEM-6_1_X57972
- 06 blaTEM-8_1_X65252
- 07 blaTEM-10_1_AF093512
- 08 blaTEM-11_1_AY874537
- 09 blaTEM-12_1_M88143
- 10 blaTEM-15_1_AM849805
- 11 blaTEM-16_1_X65254
- 12 blaTEM-17_1_Y14574

ConClave Scores:

- 63
- 64+63
- 64+63
-
-
-
- 63
- 63
- 63
-
-
- 64

$$T_m \in \operatorname{argmax}_{i \in T} \{ f(t_i) \} \quad (1)$$

$$C(t) = \sum_{k \in K} \max \begin{cases} f(t_k) & \tau \leq f(t_k) \\ 0 & \text{else} \end{cases} \quad (2)$$

$$S_k \in \operatorname{argmax}_{i \in T_m} \{ C(t_i) \} \quad (3)$$

Predicted Template(s): _____

Query sequence

	$f(t)$	T_m	S_k
GCGCGGTATTATCCCGTGGACGCCGGCAAGAGTAACCTGGTCGCCGCATAACTATTCTCAGAATGACTTG	64	2,3,12	
GAATGACTTGCTTGAGTACTCACCAAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCA	63	1,2,3,7,8,9	
TAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGGAGCTAACCGCTTTTGACAAACATTG	72	1,2,3,4,5,6,8,9,11,12	
CCACTTCTGCGCTCGGCCATCCGGCTGGCTGGTTATTGCTGATAAATCTGGAGCCGGTAGACGTGGTCTCG	66	3,5,8	
TCCTTGAGAGTTTCGCCCCGAAGAACGTTTCCAATGATGAGCAGTTAAAGTTCTGCTATGTGGCGCGGT	74	2,5,12	
GGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGGAGCTAACCGCTTTTGACAAACATTGGAAATCATG	63	1,2,3,4,5,6,8,9,11,12	
ACTCTAGCTTCCC GGCAACAATTAAAGACTGGATGGAGGCGGATTAAGTTGCAGGACCCTCTGCGCTCGC	39	1,2,4,5,6,7,8,9,10,11	
CCATCCGGCTGGCTGGTTATTGCTGATAAATCTGGAGCCGGTAGACGTGGTCTCGGGTATCATTGCAGCAC	71	2,5,12	

Templates:

- 01 blaTEM-1A_4_HM749966
 02 blaTEM-2_1_X54606
 03 blaTEM-2_3_AJ251946
 04 blaTEM-3_1_X64523
 05 blaTEM-6_1_X57972
 06 blaTEM-8_1_X65252
 07 blaTEM-10_1_AF093512
 08 blaTEM-11_1_AY874537
 09 blaTEM-12_1_M88143
 10 blaTEM-15_1_AM849805
 11 blaTEM-16_1_X65254
 12 blaTEM-17_1_Y14574

ConClave Scores:

- 63

 64+63

 64+63

 63

 63

 63

 64

$$T_m \in \operatorname{argmax}_{i \in T} \{ f(t_i) \} \quad (1)$$

$$C(t) = \sum_{k \in K} \max \begin{cases} f(t_k) & \tau \leq f(t_k) \\ 0 & \text{else} \end{cases} \quad (2)$$

$$S_k \in \operatorname{argmax}_{i \in T_m} \{ C(t_i) \} \quad (3)$$

Predicted Template(s): _____

Query sequence	f(t)	T _m	S _k
GCGCGGTATTATCCCGTGTGACGCCGGCAAGAGTAACTCGGTCGCCGCATAACTATTCTCAGAATGACTTG	64	2,3,12	2
GAATGACTTGCTTGAGTACTCACCAAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCA	63	1,2,3,7,8,9	2
TAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGGAGCTAACCGCTTTTGACAAACATTG	72	1,2,3,4,5,6,8,9,11,12	2
CCACTTCTCGCCTCGGCCATCCGGCTGGCTGGTTATTGCTGATAAAATCTGGAGCCGGTGAGCGTGGTCTCG	66	3,5,8	5
TCCTTGAGAGTTTCGCCCCGAAGAACGTTTCCAATGATGAGCACTTTAAAGTTCTGCTATGTGGCGCGTA	74	2,5,12	2
GGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGGAGCTAACCGCTTTTGACAAACATTGGGAATCATG	63	1,2,3,4,5,6,8,9,11,12	2
ACTCTAGCTTCCC GGCAACAAATTAAATAGACTGGATGGAGGCGGATTAAGTTGCAGGACCCTCTCGCCTCGC	39	1,2,4,5,6,7,8,9,10,11	2
CCATCCGGCTGGCTGGTTATTGCTGATAAAATCTGGAGCCGGTGAGCGTGGTCTCGCGGTATTCAGCAC	71	2,5,12	2

Templates:

01	blaTEM-1A_4_HM749966
02	blaTEM-2_1_X54606
03	blaTEM-2_3_AJ251946
04	blaTEM-3_1_X64523
05	blaTEM-6_1_X57972
06	blaTEM-8_1_X65252
07	blaTEM-10_1_AF093512
08	blaTEM-11_1_AY874537
09	blaTEM-12_1_M88143
10	blaTEM-15_1_AM849805
11	blaTEM-16_1_X65254
12	blaTEM-17_1_Y14574

ConClave Scores:

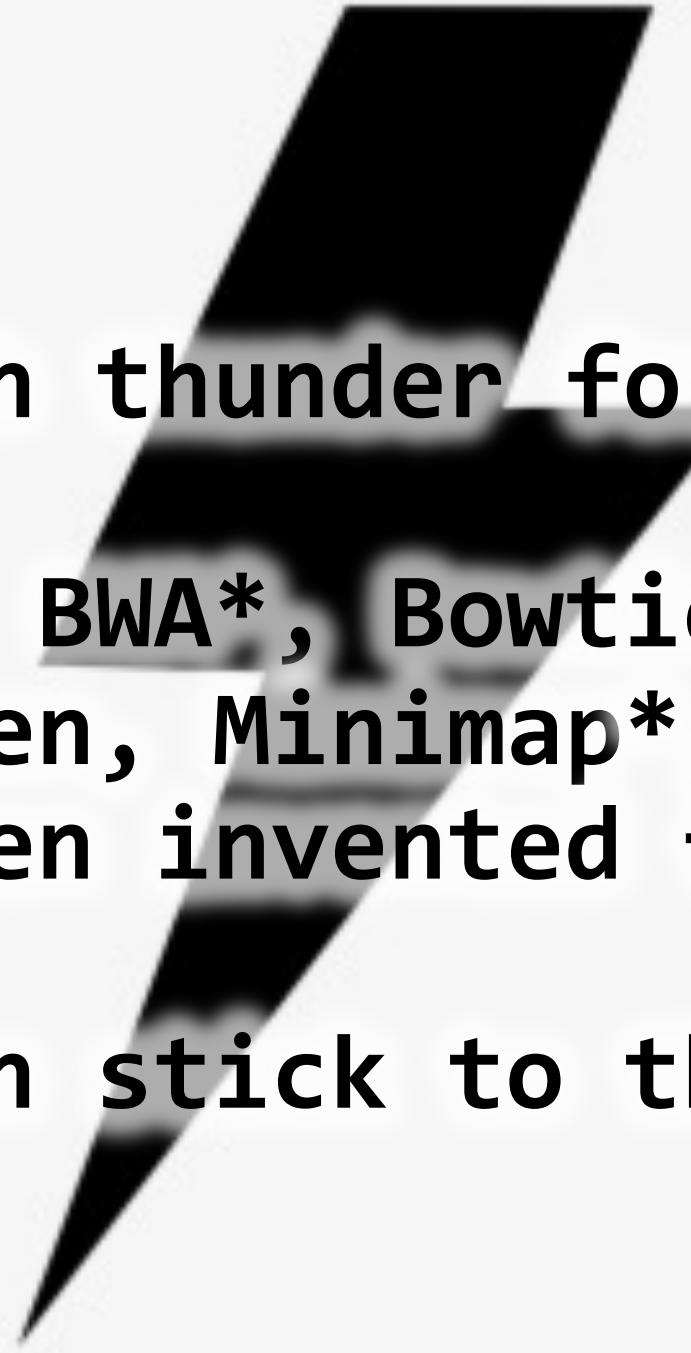
237
446
328
174
385
174
102
303
311
39
174
270

$$T_m \in \operatorname{argmax}_{i \in T} \{ f(t_i) \} \quad (1)$$

$$C(t) = \sum_{k \in K} \max \begin{cases} f(t_k) & \tau \leq f(t_k) \\ 0 & \text{else} \end{cases} \quad (2)$$

$$S_k \in \operatorname{argmax}_{i \in T_m} \{ C(t_i) \} \quad (3)$$

Predicted Template(s): "blaTEM-2_1_X54606" and "blaTEM-6_1_X57972"



Enough thunder for now:
KMA, BLAST, BWA*, Bowtie*, Kraken*,
Bracken, Minimap*, etc.
have been invented for you.

So you can stick to the basics.

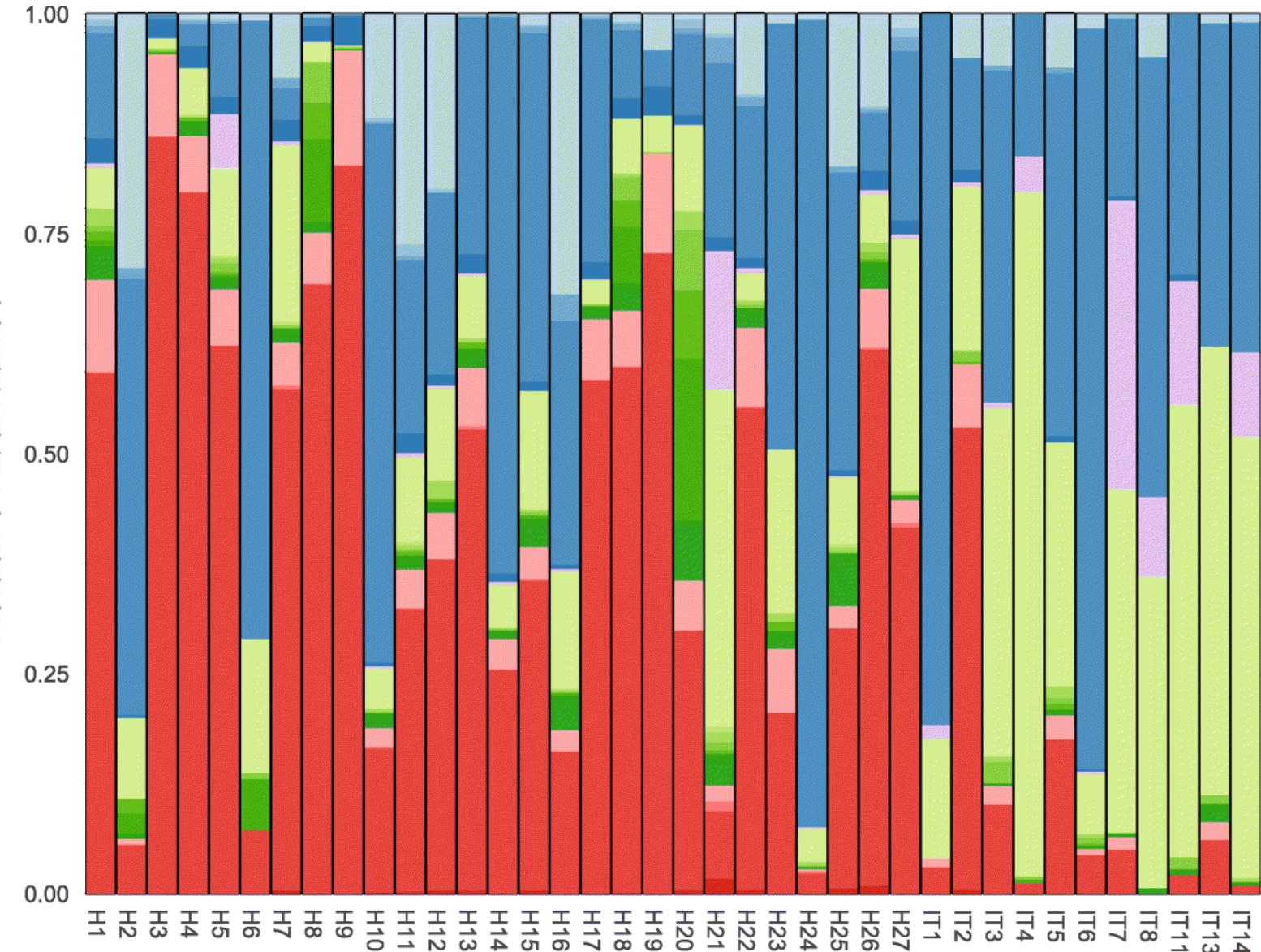
I/O

I Sequences

/ Mapping and Alignment

O (Intermediate) Result

It's not butterflies but reptiles and amphibians



Frogs

- Bufonidae
- Dicroglossidae
- Pipidae
- Pyxicephalidae

Crocodiles

- Crocodylidae

Snakes and Lizards

- Colubridae
- Dactyloidae
- Elapidae
- Gekkonidae
- Pythonidae
- Varanidae
- Viperidae

Turtles and tortoises

- Cheloniidae
- Dermochelyidae
- Geoemydidae
- Testudinidae

I and O through /
on a board



Questions and Exercises

Papers:

minimap2

Chaining

Methods

kma

ConClave

Methods & Results

ccmetagen

LCA & ConClave

Methods & Results

nasko

LCA pitfalls

Results & Discussion

marcelino

DB pitfalls

Results & Discussion