

Genome 540 Discussion

Feb. 12th 2026



HW5 Review

Overview

1. Build an edit graph (WDAG) for the 3 amino acid sequences of insulin
(*Homo sapiens*, *Bubalus bubalis*, *Xenopus tropicalis*)
2. Save it (as a text file) in the same format as HW4
3. Run the highest path algorithm from HW4 on the edit graph

Building the edit graph

LRRMVNVDSVIVF

LLQYTPVESGLMLF

LNKYIRKIDSVILF



LLL
RLN
RQK
MYV
VTI
--R
--K
NPD
VVI
DED
SSS
-G-
VLV
IMI
VLL
FFF

Building the edit graph

LLL
RLN
RQK
MYY
VTI
--R
--K
NPD
VVI
DED
SSS
-G-
VLV
IMI
VLL
FFF

At some arbitrary position in the alignment, what are all the possibilities for the *next* alignment slot?

Building the edit graph

LLL
RLN
RQK
MYV
VTI
--R
--K
NPD
VVI
DED
SSS
-G-
VLV
IMI
VLL
FFF

LRRMVNVDSVIVF

 i=3

LNKYIRKIDSVILF

 j=5

LLQYTPVESGLMLF

 k=2

Building the edit graph

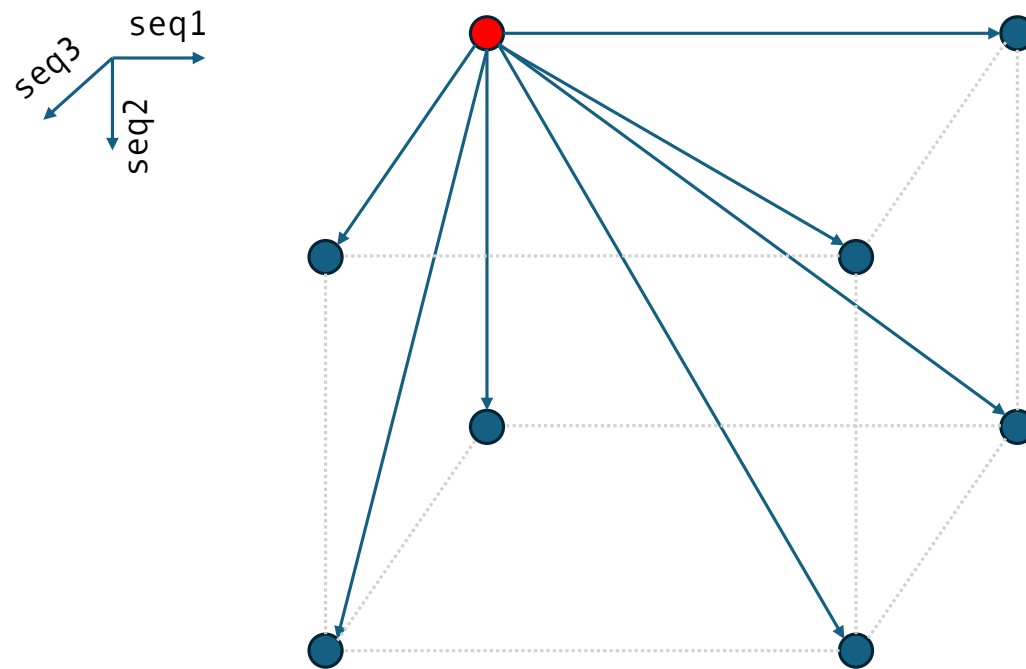
LLL
RLN
RQK
MYY
VTI
--R
--K
NPD
VVI
DED
SSS
-G-
VLV
IMI
VLL
FFF

for a given (i, j, k) :

$(i + 1, j + 1, k + 1)$	(consuming a base from each seq)
$(i + 1, j + 1, k)$	(gap in seq 3)
$(i + 1, j, k + 1)$	(gap in seq 2)
$(i, j + 1, k + 1)$	(gap in seq 1)
$(i + 1, j, k)$	(gap in seq 2 and 3)
$(i, j + 1, k)$	(gap in seq 1 and 3)
$(i, j, k + 1)$	(gap in seq 1 and 2)

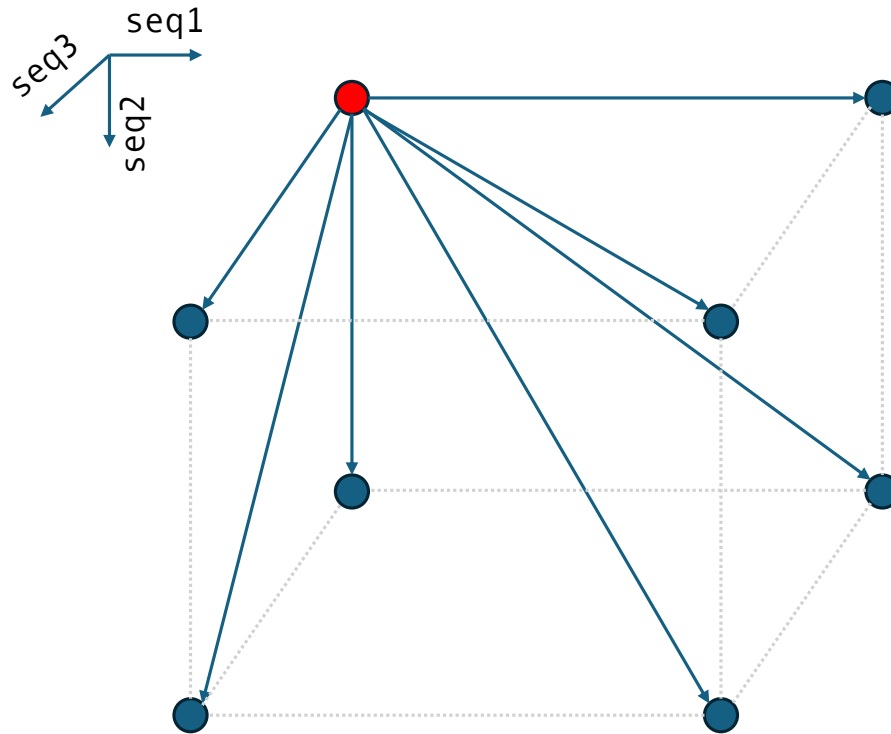
Building the edit graph

LLL
RLN
RQK
MYV
VTI
--R
--K
NPD
VVI
DED
SSS
-G-
VLV
IMI
VLL
FFF



Building the edit graph

LLL
RLN
RQK
MYV
VTI
--R
--K
NPD
VVI
DED
SSS
-G-
VLV
IMI
VLL
FFF



LRRMVNVDSVIVF

i=3

LNKYIRKDIDSVILF

j=5

LLQYTPVESGLMLF

k=2

Building the edit graph

LLL
 RLN
 RQK
 MYY
 VTI
 --R
 --K
 NPD
 VVI
 DED
 SSS
 -G-
 VLV
 IMI
 VLL
 FFF

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V	B	Z	X	*
A	4	-1	-2	-2	0	-1	-1	0	-2	-1	-1	-1	-1	-2	-1	1	0	-3	-2	0	-2	-1	0	-4
R	-1	5	0	-2	-3	1	0	-2	0	-3	-2	2	-1	-3	-2	-1	-1	-3	-2	-3	-1	0	-1	-4
N	-2	0	6	1	-3	0	0	0	1	-3	-3	0	-2	-3	-2	1	0	-4	-2	-3	3	0	-1	-4
D	-2	-2	1	6	-3	0	2	-1	-1	-3	-4	-1	-3	-3	-1	0	-1	-4	-3	-3	4	1	-1	-4
C	0	-3	-3	-3	9	-3	-4	-3	-3	-1	-1	-3	-1	-2	-3	-1	-1	-2	-2	-1	-3	-3	-2	-4
Q	-1	1	0	0	-3	5	2	-2	0	-3	-2	1	0	-3	-1	0	-1	-2	-1	-2	0	3	-1	-4
E	-1	0	0	2	-4	2	5	-2	0	-3	-3	1	-2	-3	-1	0	-1	-3	-2	-2	1	4	-1	-4
G	0	-2	0	-1	-3	-2	-2	6	-2	-4	-4	-2	-3	-3	-2	0	-2	-2	-3	-3	-1	-2	-1	-4
H	-2	0	1	-1	-3	0	0	-2	8	-3	-3	-1	-2	-1	-2	-1	-2	-2	2	-3	0	0	-1	-4
I	-1	-3	-3	-3	-1	-3	-3	-4	-3	4	2	-3	1	0	-3	-2	-1	-3	-1	3	-3	-3	-1	-4
L	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4	-2	2	0	-3	-2	-1	-2	-1	1	-4	-3	-1	-4
K	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5	-1	-3	-1	0	-1	-3	-2	-2	0	1	-1	-4
M	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5	0	-2	-1	-1	-1	-1	1	-3	-1	-1	-4
F	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6	-4	-2	-2	1	3	-1	-3	-3	-1	-4
P	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	7	-1	-1	-4	-3	-2	-2	-1	-2	-4
S	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4	1	-3	-2	-2	0	0	0	-4
T	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	1	5	-2	-2	0	-1	-1	0	-4
W	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	1	-4	-3	-2	11	2	-3	-4	-3	-2	-4
Y	-2	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7	-1	-3	-2	-1	-4
V	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	-2	0	-3	-1	4	-3	-2	-1	-4
B	-2	-1	3	4	-3	0	1	-1	0	-3	-4	0	-3	-3	-2	0	-1	-4	-3	-3	4	1	-1	-4
Z	-1	0	0	1	-3	3	4	-2	0	-3	-3	1	-1	-3	-1	0	-1	-3	-2	-2	1	4	-1	-4
X	0	-1	-1	-1	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-2	0	0	-2	-1	-1	-1	-1	-1	-4
*	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	1

Gap penalty: -6

LRRMVNVDSVIVF

i=3

LNKYIRKDIDSVILF

j=5

LLQYTPVESGLMLF

k=2

What to turn in:

- The highest path score
- All possible edge weights (sorted alphabetically)
- Histogram of edge counts (from the full edit graph)
- Highest scoring alignment (the highest path, edge names!)

Score: 82.0

Edge weights:

--A = -12
--C = -12
--D = -12
--E = -12
--F = -12

.
.
.
list all edge weights in alphabetical order
(only first/last 5 shown here)

.
.
.
YYS = 3
YYT = 3
YYV = 5
YYW = 11
YYY = 21

Edge counts:

--A = 8832
--C = 17664
--D = 52992
--E = 70656
--F = 44160

.
.
.
list all the edge counts in alphabetical order
(only first/last 5 shown here)

.
.
.
YYS = 48
YYT = 24
YYV = 72
YYW = 24
YYY = 60

Local alignment:

KKK
DLK
YWY
G--
LFL
KVN
REH
IPI

Joe Min

Highest weight path

Local Alignment
(What does that
mean for the WDAG
program you run?)

Score: 82.0

Local alignment:

KKK
DLK
YWY
G--
LFL
KVN
REH
IPI

Edit graph stats

- NOT all possible edge weights, only those which appear *in* the graph. (How many distinct edges could there be, theoretically?)
- Each entry in the edge weight table should have an entry in the edge count histogram.

Edge weights:

--A = -12
--C = -12
--D = -12
--E = -12
--F = -12

.

.

.

list all edge weights in alphabetical order
(only first/last 5 shown here)

.

.

.

YYS = 3

YYT = 3

YYV = 5

YYW = 11

YYY = 21

Edge counts:

--A = 8832
--C = 17664
--D = 52992
--E = 70656
--F = 44160

.

.

.

list all the edge counts in alphabetical order
(only first/last 5 shown here)

.

.

.

YYS = 48

YYT = 24

YYV = 72

YYW = 24

YYY = 60