



Βιοπληροφορική

Ενότητα 7:

Στοίχιση ακολουθιών ανά
ζεύγη – Τεχνικές Στοίχισης
Ακολουθιών, (1/2) 1ΔΩ

Τμήμα: Βιοτεχνολογίας

Όνομα καθηγητή: Τ. Θηραίου



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης





Μαθησιακοί Στόχοι

- Παρουσίαση της μεθόδου κατασκευής και των εφαρμογών των διαγραμμάτων πινάκων σημείων.
- Επεξήγηση των αλγορίθμων δυναμικού προγραμματισμού για τη στοίχιση ακολουθιών ανά ζεύγη.
- Κατανόηση της διαφοράς μεταξύ της βέλτιστης και της βιολογικά σωστής στοίχισης.



Λέξεις Κλειδιά

- Λέξεις κλειδιά: Διαγράμματα Πινάκων Σημείων, Αλγόριθμοι Δυναμικού Προγραμματισμού.
- Key words: Dot Matrix Plots, Dynamic programming algorithms for pairwise sequence alignment, Needleman-Wunsch algorithm, Smith-Waterman algorithm.



Τεχνικές Στοίχισης Ακολουθιών

- Διαγράμματα Πινάκων Σημείων (Dot Matrix Plots).
- Αλγόριθμοι Δυναμικού Προγραμματισμού:
 - Needleman-Wunsch algorithm (ολική στοίχιση).
 - Smith-Waterman algorithm (τοπική στοίχιση).
- Ευριστικοί Αλγόριθμοι:
 - FASTA.
 - BLAST.



Dot Plots 1/5

	T	G	C	A	A	T	C	G	G
A									
A									
C									
T									
G									
A									
A									
T									
C									

A dot plot matrix showing the pairwise relationships between nucleotides T, G, C, A, and A. The diagonal elements are shaded gray. A red dotted line is drawn from the top-right cell (A, A) to the bottom-left cell (C, C).



Dot Plots 2/5

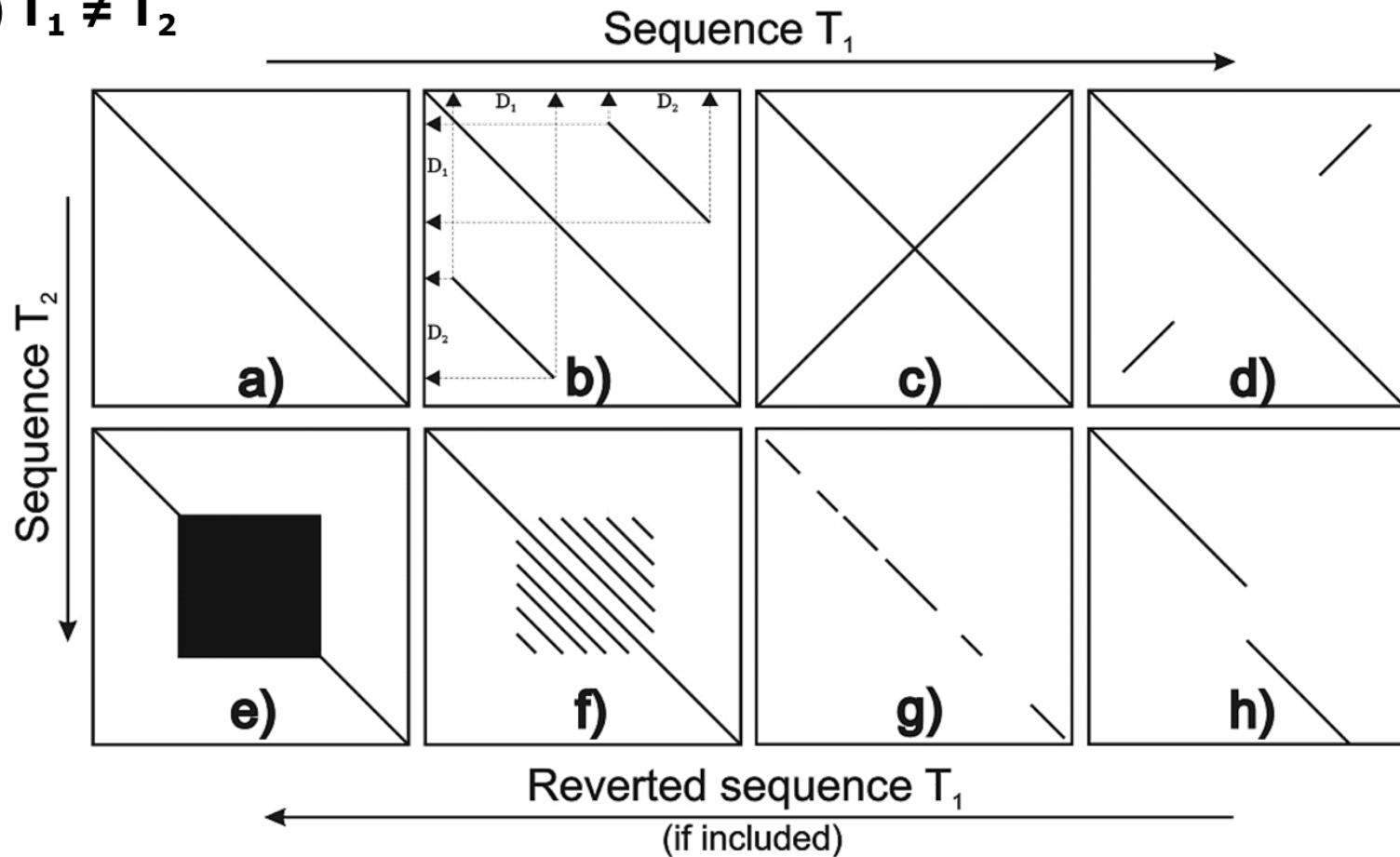
- κατασκευή δισδιάστατου ορθογώνιου διαγράμματος με τις υπό σύγκριση ακολουθίες στη θέση των αξόνων.
- χρωματισμός των κελιών που αντιστοιχούν σε ταυτόσημα κατάλοιπα των δύο ακολουθιών.
- περιοχές του πίνακα με χρωματισμένα κελιά **κατά μία διαγώνιο** υποδεικνύουν περιοχές ταύτισης των ακολουθιών.
- εφαρμογή τεχνικών φιλτραρίσματος για τη μείωση του θορύβου (μεμονωμένες ή μικρού μήκους στοιχίσεις).
- χρήση πινάκων αντικαταστάσεως για τη βαθμολόγηση ζευγών καταλοίπων.
- **υποκειμενική ερμηνεία** των αποτελεσμάτων.



Dot Plots 3/5

a-f) $T_1 = T_2$

g-h) $T_1 \neq T_2$





Dot Plots 5/5

- **άμεση οπτικοποίηση:**
 - επαναλήψεων.
 - παλίνδρομων ακολουθιών.
 - περιοχών χαμηλής πολυπλοκότητας.
 - ενθέσεων ή διαγραφών.
- μελέτη γονιδιωμάτων.
- δομή RNA.



Προγράμματα Στοίχισης Ακολουθιών

● Dot Plots:

- Dotlet.

<http://myhits.isb-sib.ch/cgi-bin/dotlet>

- Gepard.

<http://www.helmholtz-muenchen.de/en/mips/services/analysis-tools/gepard/index.html>



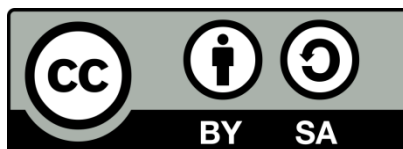
Βιβλιογραφία

- David Mount, "Bioinformatics: Sequence and Genome Analysis", Cold Spring Harbor Laboratory Press; 2nd edition (March 12, 2013).
- Jonathan Pevsner, "Bioinformatics and Functional Genomics", Wiley-Blackwell; 2nd edition (May 4, 2009).
- Andreas D. Baxevanis, B. F. Francis Ouellette, "Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins", Wiley-Interscience; 3rd edition (October 29, 2004).



Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδεια χρήσης, η άδεια χρήσης αναφέρεται ρητώς.





Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στο πλαίσιο του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα Γεωπονικού Πανεπιστημίου Αθηνών**» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ



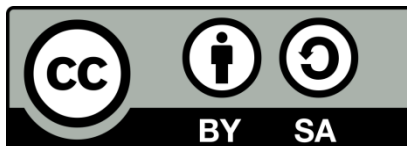
Σημείωμα Αναφοράς

Copyright Γεωπονικό Πανεπιστήμιο Αθηνών 2015. Τμήμα Βιοτεχνολογίας, Θηραίου Τριάς. «Βιοπληροφορική». Έκδοση: 1.0. Αθήνα 2015. Διαθέσιμο από τη δικτυακή διεύθυνση:
<https://mediasrv.aua.gr/eclass/courses/OCDB100/>



Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Παρόμοια Διανομή 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων, π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



Η άδεια αυτή ανήκει στις άδειες που ακολουθούν τις προδιαγραφές του Ορισμού Ανοικτής Γνώσης [2], είναι ανοικτό πολιτιστικό έργο [3] και για το λόγο αυτό αποτελεί ανοικτό περιεχόμενο [4].

[1] <http://creativecommons.org/licenses/by-sa/4.0/>

[2] <http://opendefinition.org/okd/ellinika/>

[3] <http://freedomdefined.org/Definition/EI>

[4] <http://opendefinition.org/buttons/>



Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
 - το Σημείωμα Αδειοδότησης
 - τη δήλωση Διατήρησης Σημειωμάτων
 - το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)
- μαζί με τους συνοδευόμενους υπερσυνδέσμους.