



ΔΗΜΟΚΡΙΤΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΡΑΚΗΣ

DEMOCRITUS
UNIVERSITY
OF THRACE

Βιοπληροφορική και σύγχρονες ερευνητικές προσεγγίσεις

Πέτρος Κολοβός



ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΡΑΚΗΣ

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ

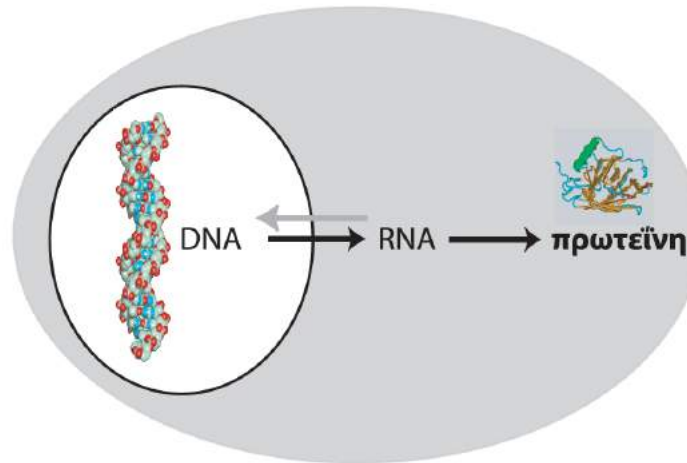
ΤΜΗΜΑ ΜΟΡΙΑΚΗΣ ΒΙΟΛΟΓΙΑΣ ΚΑΙ ΓΕΝΕΤΙΚΗΣ

Κεντρικό δόγμα

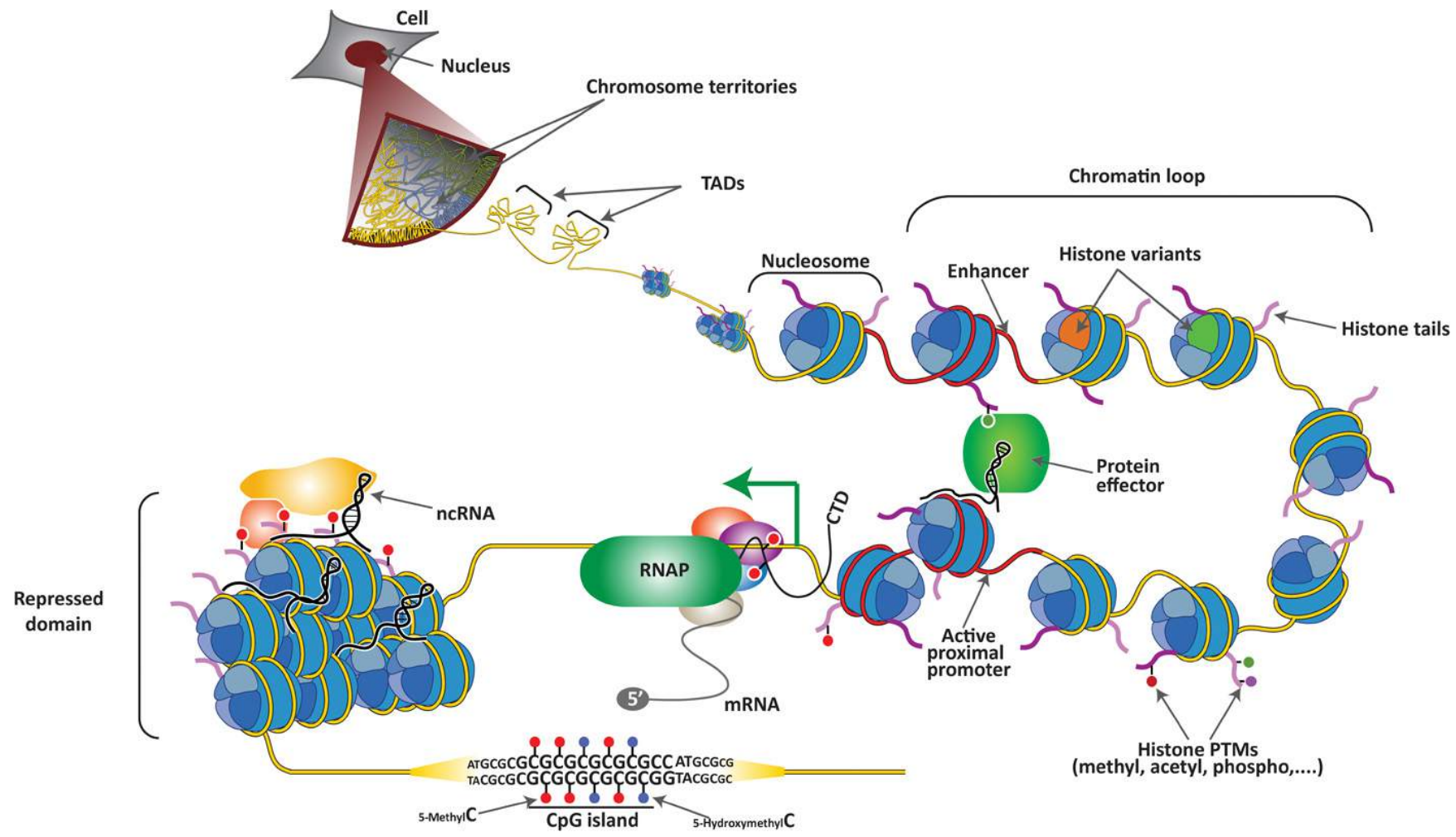
Κεντρικό δόγμα της μοριακής βιολογίας



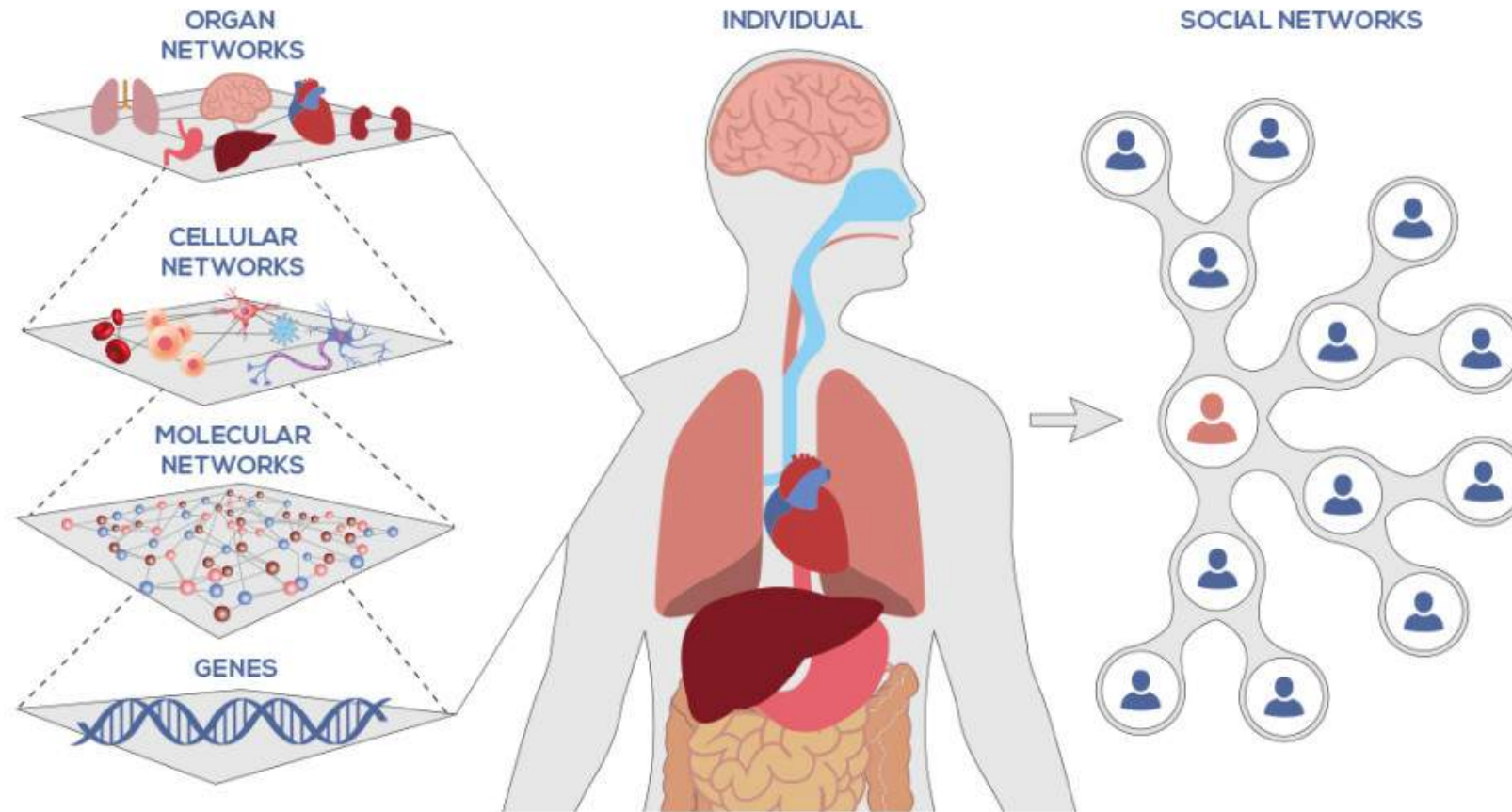
Κεντρικό δόγμα της γονιδιωματικής



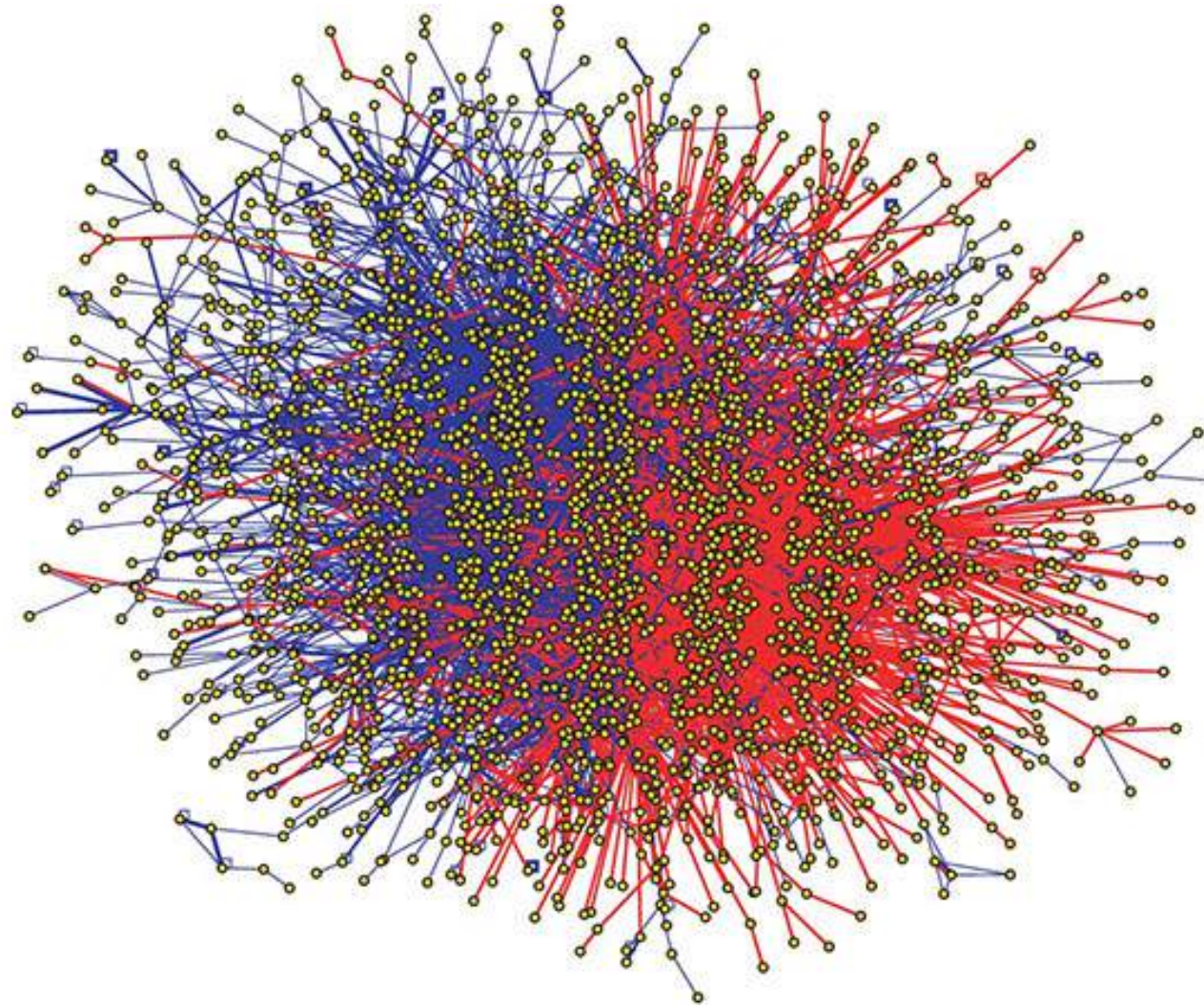
Οργάνωση της χρωματίνης



Ιεραρχικά δίκτυα βιολογικών συστημάτων



Δίκτυα πρωτεϊνικών αλληλεπιδράσεων



Τι είναι βιολογία συστημάτων

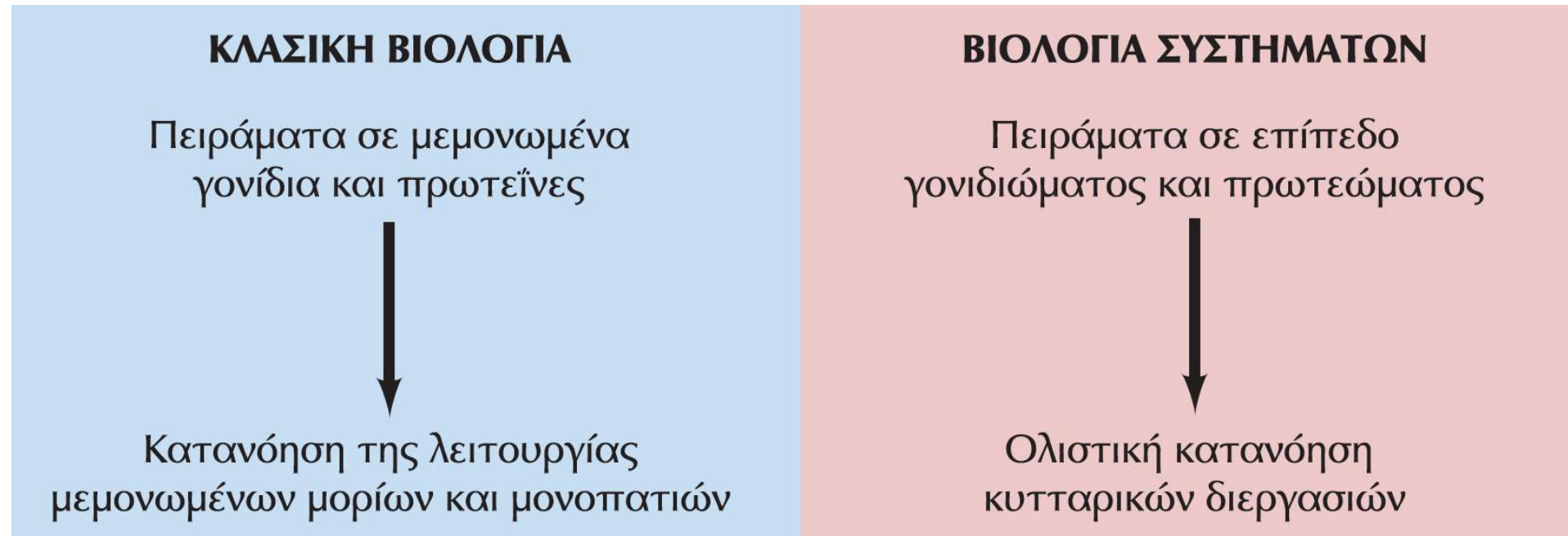


telegraph.co.uk credit westend61



visitgreece.gr

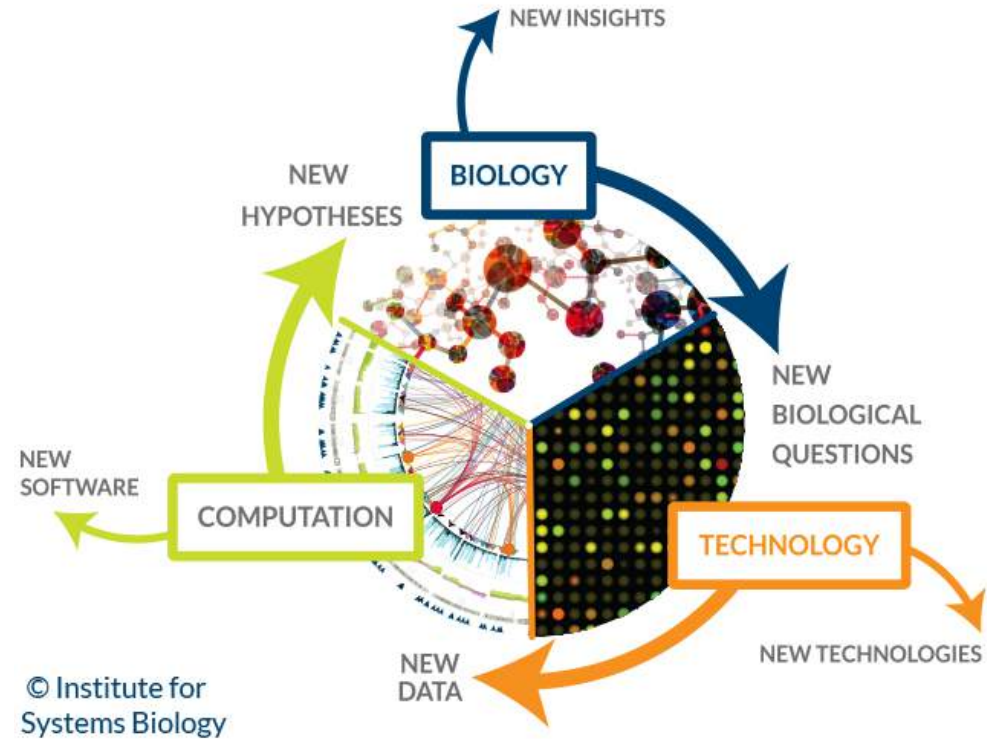
Τι είναι βιολογία συστημάτων



ΕΙΚΟΝΑ 5.15 Βιολογία συστημάτων. Η κλασική πειραματική βιολογία εστιάζεται στη μελέτη επιμέρους μορίων και μονοπατιών. Η βιολογία συστημάτων προσεγγίζει σφαιρικά τα βιολογικά συστήματα, αναλύοντας μεγάλους όγκους δεδομένων και εξάγοντας ποσοτικά μοντέλα που αποσκοπούν στην ολοκληρωμένη κατανόηση πολύπλοκων συστημάτων



Τι είναι βιολογία συστημάτων



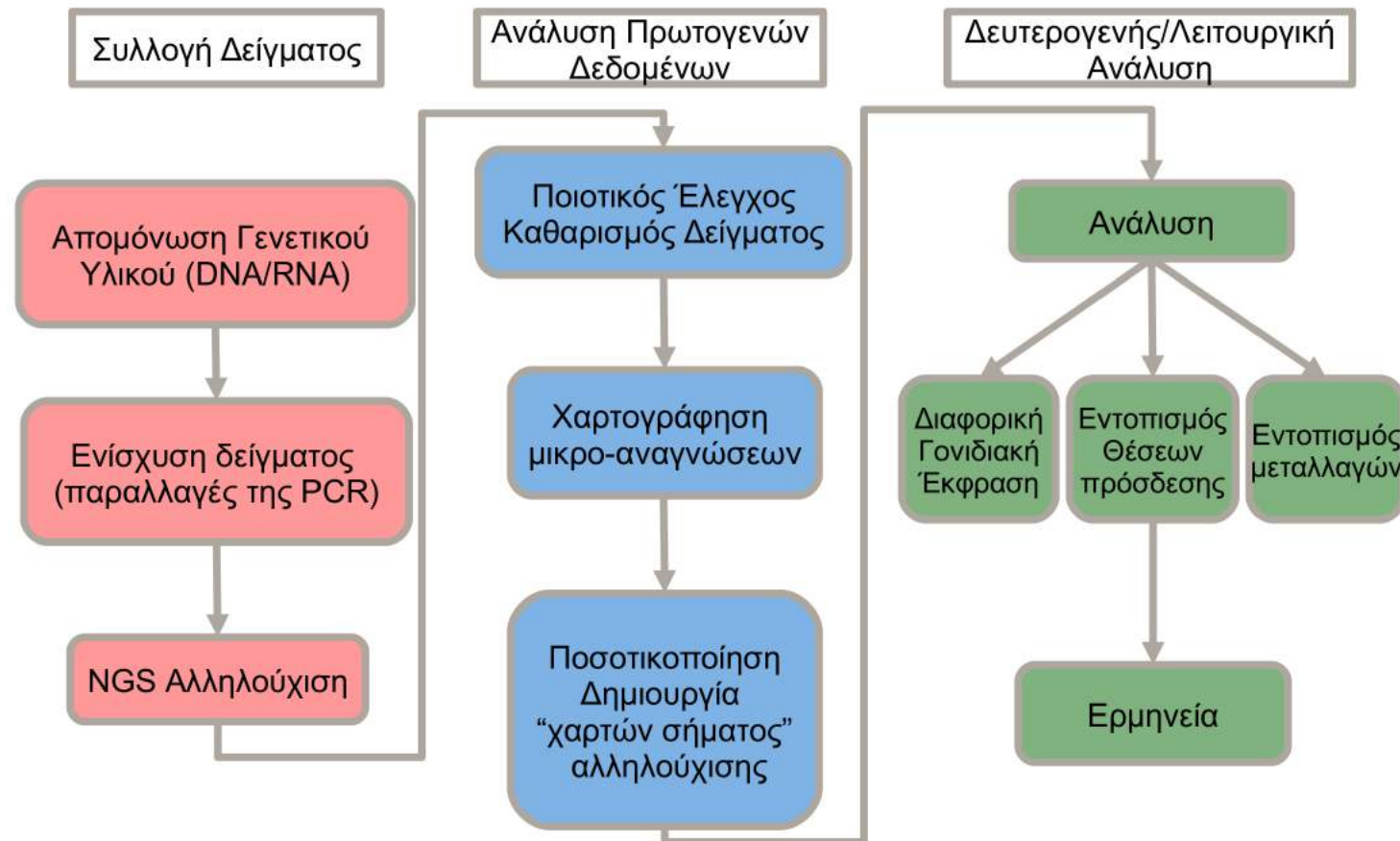
Η βιολογία συστημάτων επιδιώκει να εξηγήσει τις **ιδιότητες** και τη **συμπεριφορά** των σύνθετων βιολογικών συστημάτων από την άποψη των **συστατικών** τους και των **αλληλεπιδράσεων** τους.

α) τις ολιστικές προσεγγίσεις βιολογικών συστημάτων για την αποκρυπτογράφηση της πολυπλοκότητας των βιολογικών συστημάτων

β) τη μοντελοποίηση των δικτύων σε γονιδιακό–πρωτεϊνικό–λειτουργικό επίπεδο

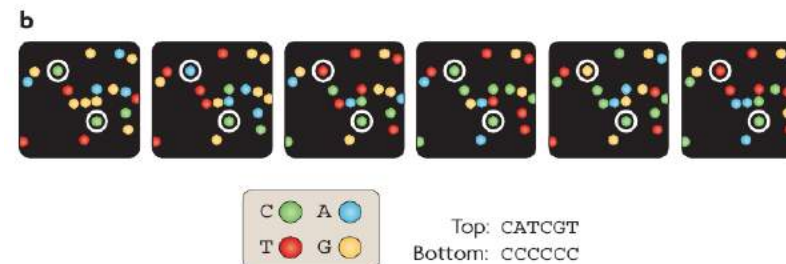
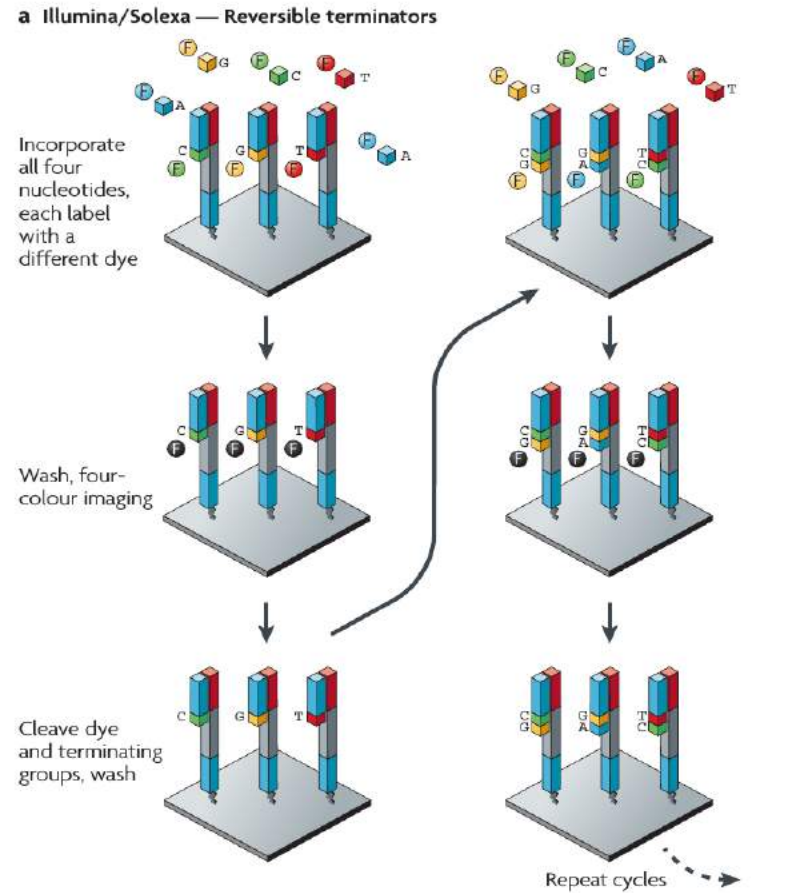
γ) τη βιολογία και την βιοπληροφορική

Γονιδιωματική μεγάλης κλίμακας

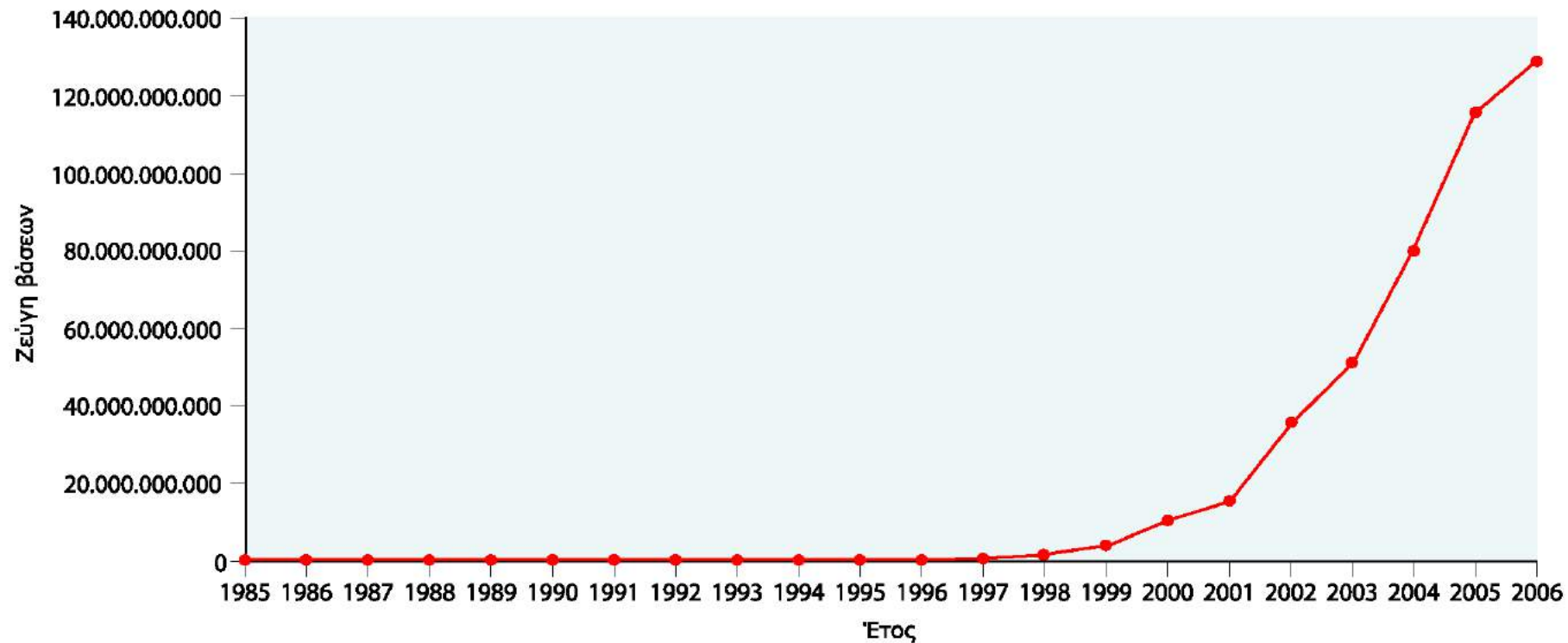


Εικόνα 11.1: Διάγραμμα ροής μιας γονιδιωματικής ανάλυσης μεγάλης κλίμακας με τη χρήση NGS μεθοδολογιών, από τη συλλογή και αλληλούχιση του δείγματος, στην ανάλυση των πρωτογενών δεδομένων αλληλουχίας ως τη λειτουργική ανάλυση που διαφέρει ανάλογα με την εφαρμογή.

Η μέθοδος της Illumina: πραγματοποίηση κύκλων αναστρέψιμου τερματισμού



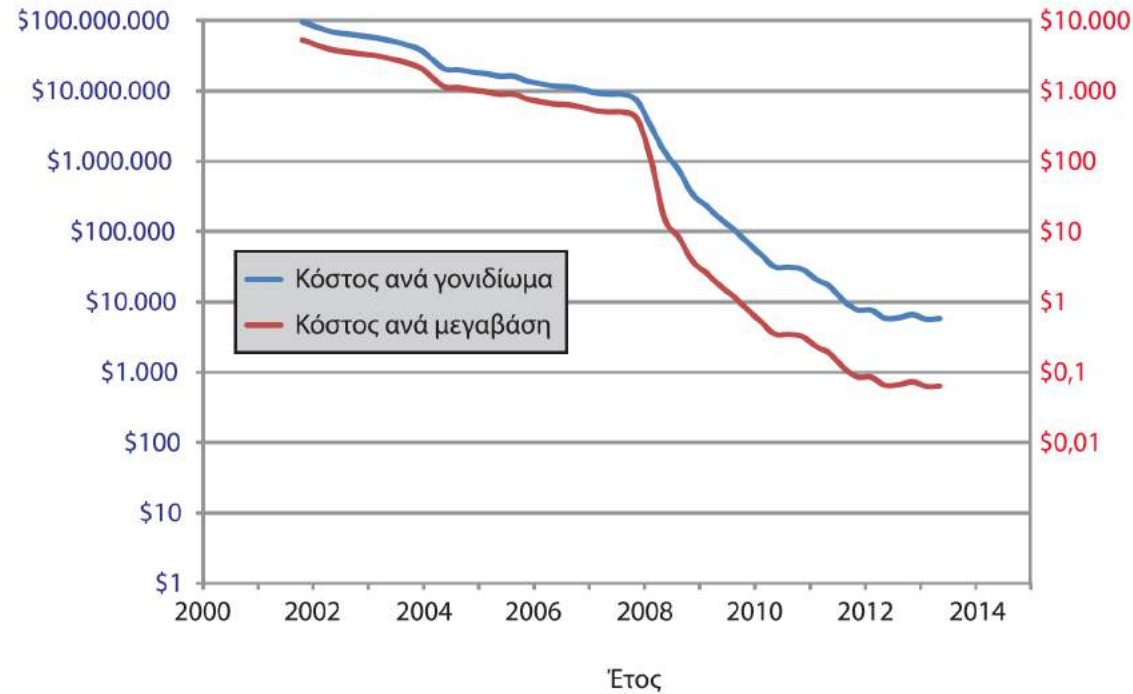
Νουκλεοτιδικές αλληλουχίες στις βάσεις δεδομένων



ΕΙΚΟΝΑ 10.1: Γραφική παράσταση της αύξησης των νουκλεοτιδικών αλληλουχιών στις δημόσιες βάσεις δεδομένων (GenBank, EBI, DDBJ) από το 1985 ως το 2006.



Αλληλούχιση επόμενης γενιάς



Εικόνα 9.3 Η μείωση του κόστους αλληλούχισης του DNA. Παρουσιάζεται το κόστος ανά μεγαβάση (εκατομμύρια ζεύγη βάσεων) για αλληλούχιση με ελάχιστη βαθμολογία ποιότητας Q20 (ή PHRED20). Το κόστος ανά γονιδίωμα αναφέρεται στο ανθρώπινο γονιδίωμα. Ο άξονας γ είναι λογαριθμικός.



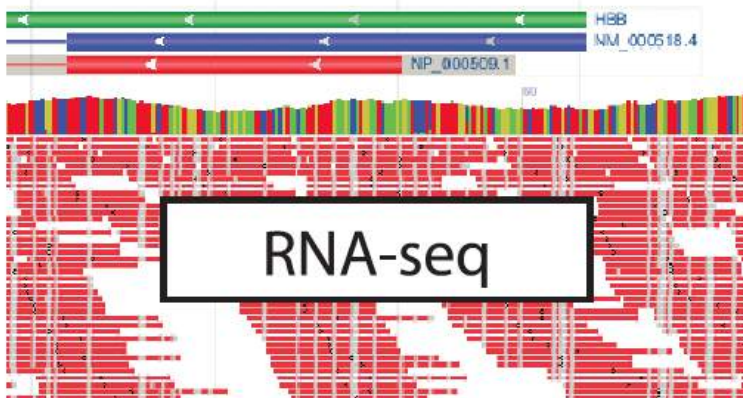
Πολύ-επίπεδη Πολυπλοκότητα

- Μελέτη έκφρασης γονιδίων
 - RNA-seq
- Εντοπισμός θέσεων πρόσδεσης μεταγραφικών παραγόντων και ευχρωματίνης
 - ChIP-seq
 - FAIRE-seq
 - DNaseI-seq
- Μελέτη επιγενετικών τροποποιήσεων σε γονιδιωματική κλίμακα
 - MeCapSeq
 - MeDIPSeq
 - bisulfite sequencing, BSseq
 - Μετα-μεταφραστικές τροποποιήσεις ιστονών (ChIP-seq)
- Αρχιτεκτονική της χρωματίνης (3D chromatin architecture)
 - xC, T2C, Hi-C

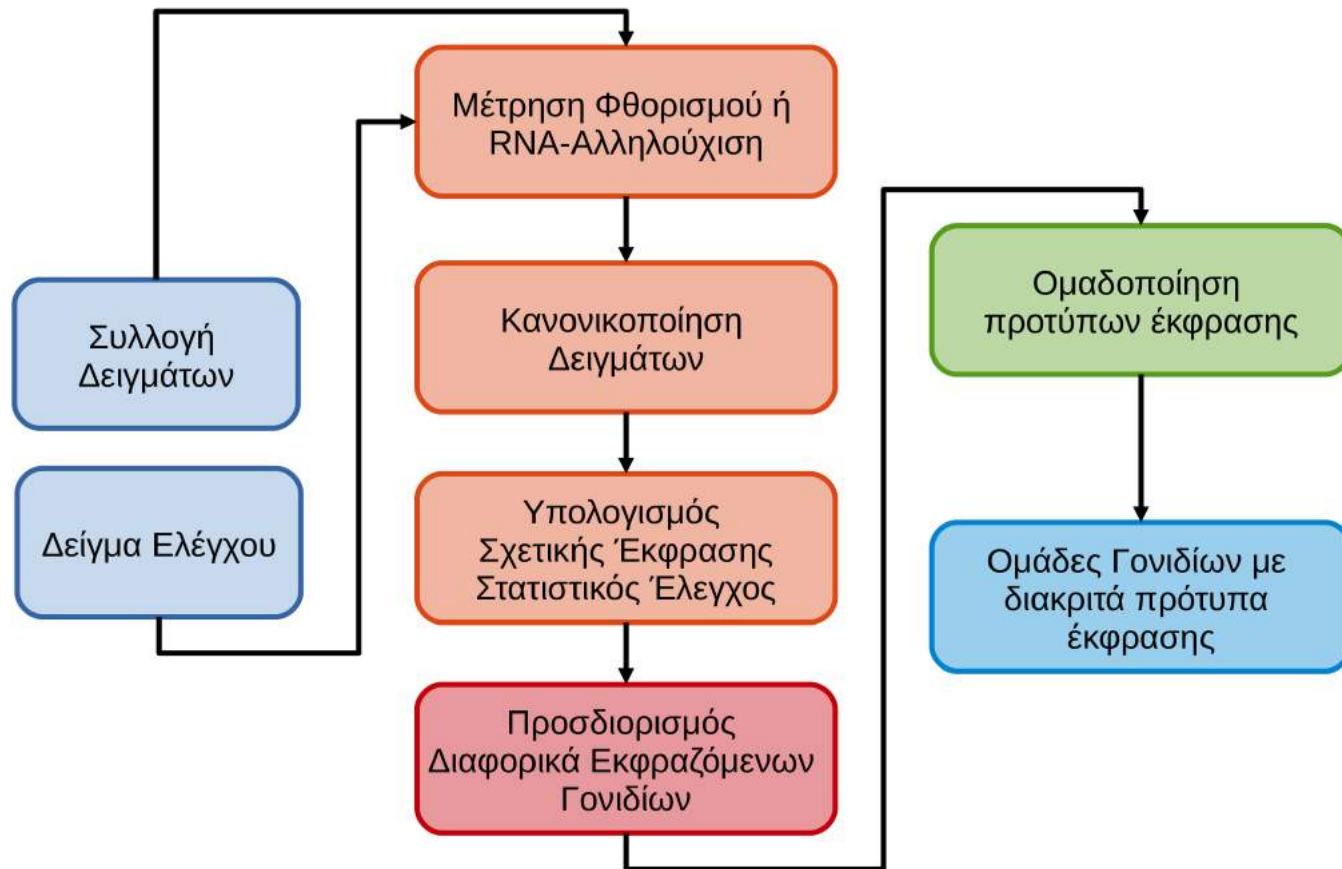
Μέθοδοι για τη μελέτη της γονιδιακής έκφρασης

Αλληλούχιση RNA (RNA Sequencing)

- ✓ Η υψηλή ζήτηση για αλληλούχιση χαμηλού κόστους έχει οδηγήσει την τελευταία δεκαετία στην ανάπτυξη της **αλληλούχισης υψηλής απόδοσης** (ή **αλληλούχισης επόμενης γενιάς**, next generation sequencing, **NGS**).
- ✓ Η **επεκτασιμότητα**, η **ταχύτητα** αλλά κυρίως η **σχέση κόστους-απόδοσης** των NGS εφαρμογών επιτρέπουν στους ερευνητές να μελετήσουν τα βιολογικά συστήματα σε επίπεδο που δεν ήταν δυνατό μέχρι πρότινος.
- ✓ Η NGS ποσοτικοποίηση της γονιδιακής έκφρασης γίνεται με τη μαζική αλληλούχιση mRNA που αρχικά απομονώνεται από το δείγμα και στη συνέχεια μετατρέπεται σε cDNA όπως στην περίπτωση των μικροσυστοιχιών.



Στάδια ανάλυσης ενός πειράματος γονιδιακής έκφρασης



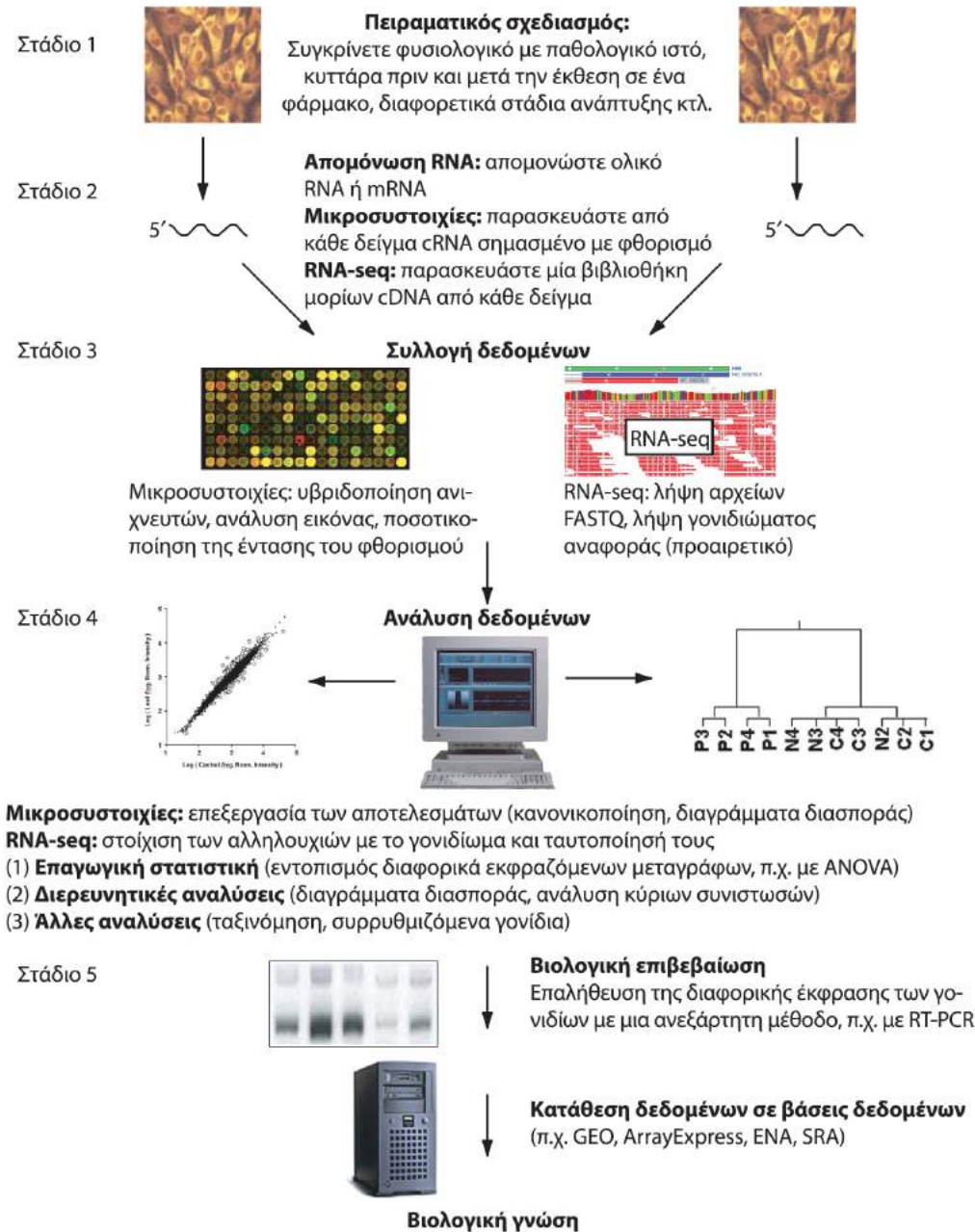
Εικόνα 7.2: Σχηματική αναπαράσταση των σταδίων ενός πειράματος γονιδιακής έκφρασης από την αποκομιδή των πρωτογενών δεδομένων ως τη δημιουργία ομάδων γονιδίων με χαρακτηριστικά πρότυπα έκφρασης.

Αλληλούχιση RNA (RNA Sequencing)

Τα στάδια της ανάλυσης περιλαμβάνουν:

- την απομόνωση του mRNA
- την ποσοτικοποίηση του mRNA
- τον προσδιορισμό των διαφορικά εκφραζόμενων γονιδίων
- την ομαδοποίηση γονιδίων ανάλογα με τα πρότυπα έκφρασής τους

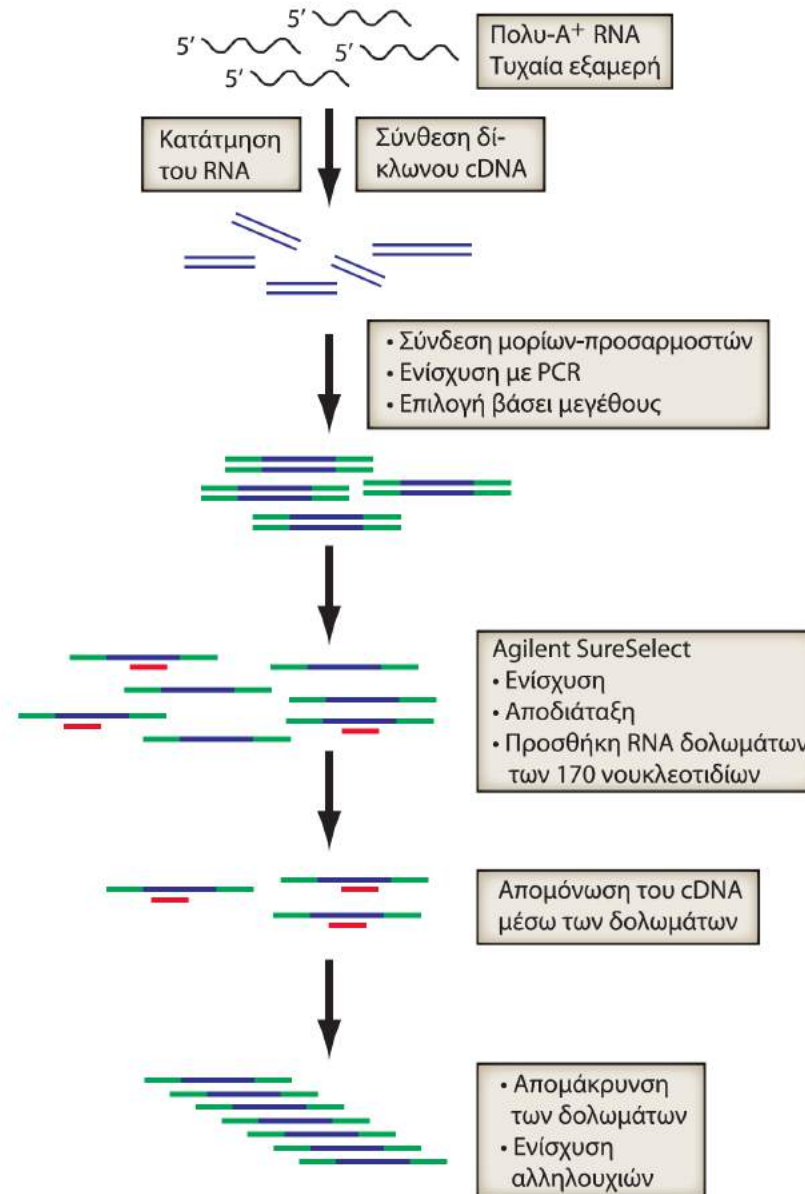
Επισκόπηση μικροσυστοιχιών και RNA-seq



Εικόνα 10.13 Επισκόπηση των πειραμάτων ανάλυσης της γονιδιακής έκφρασης με μικροσυστοιχίες και με RNA-seq.



Στοχευμένο RNA-seq



Εικόνα 11.22 Στοχευμένη μέθοδος RNA-seq. Η ροή εργασίας SureSelect της Agilent.



Διαφορική έκφραση γονιδίων

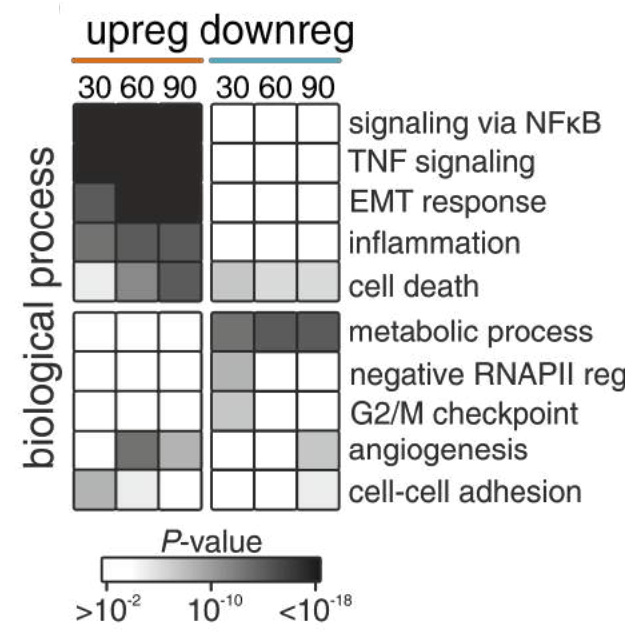
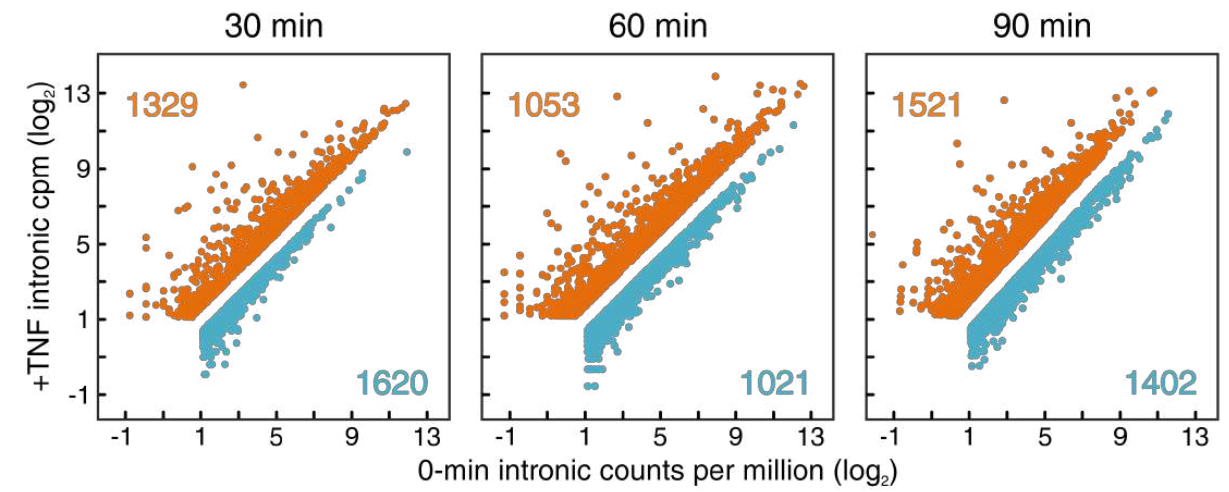
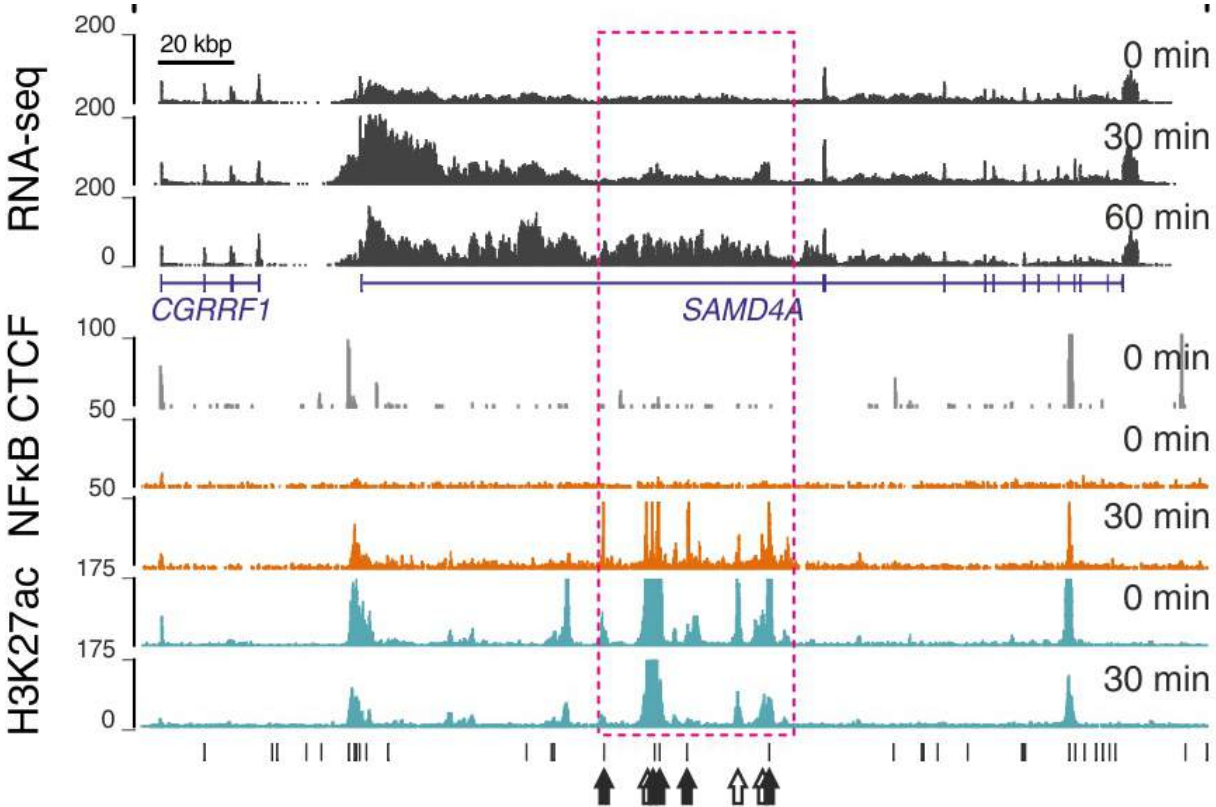
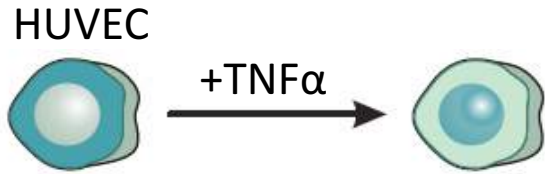
Gene	Normal1	Normal2	Normal3	Normal4	Mean (Normal)	Arthritic1	Arthritic2	Arthritic3	Arthritic 4	Mean (Arthritic)	Log (Arthritic/Normal)	P-value
Vcam	4.984	3.831	4.325	3.564	4.176	2.027	2.805	2.933	2.260	2.506	-0.737	0.005
Vegfa	8.773	6.616	8.199	6.134	7.431	4.325	6.308	6.769	5.372	5.694	-0.384	0.081
Akt	0.999	1.000	1.001	1.000	1.000	1.000	1.000	1.435	0.999	1.109	0.149	0.357
Myc	3.592	2.679	2.629	2.204	2.776	2.716	4.639	7.004	5.217	4.894	0.818	0.063
Hif1a	0.796	0.642	1.862	1.691	1.248	1.283	1.344	1.000	1.090	1.179	-0.081	0.837

Ο υπολογισμός του βαθμού της μεταβολής της έκφρασης του ίδιου γονιδίου μεταξύ δύο διαφορετικών συνθηκών γίνεται με τη χρήση του λογαρίθμου του λόγου των τιμών έκφρασης στη συνθήκη μελέτης (test) προς τη συνθήκη ελέγχου (control).

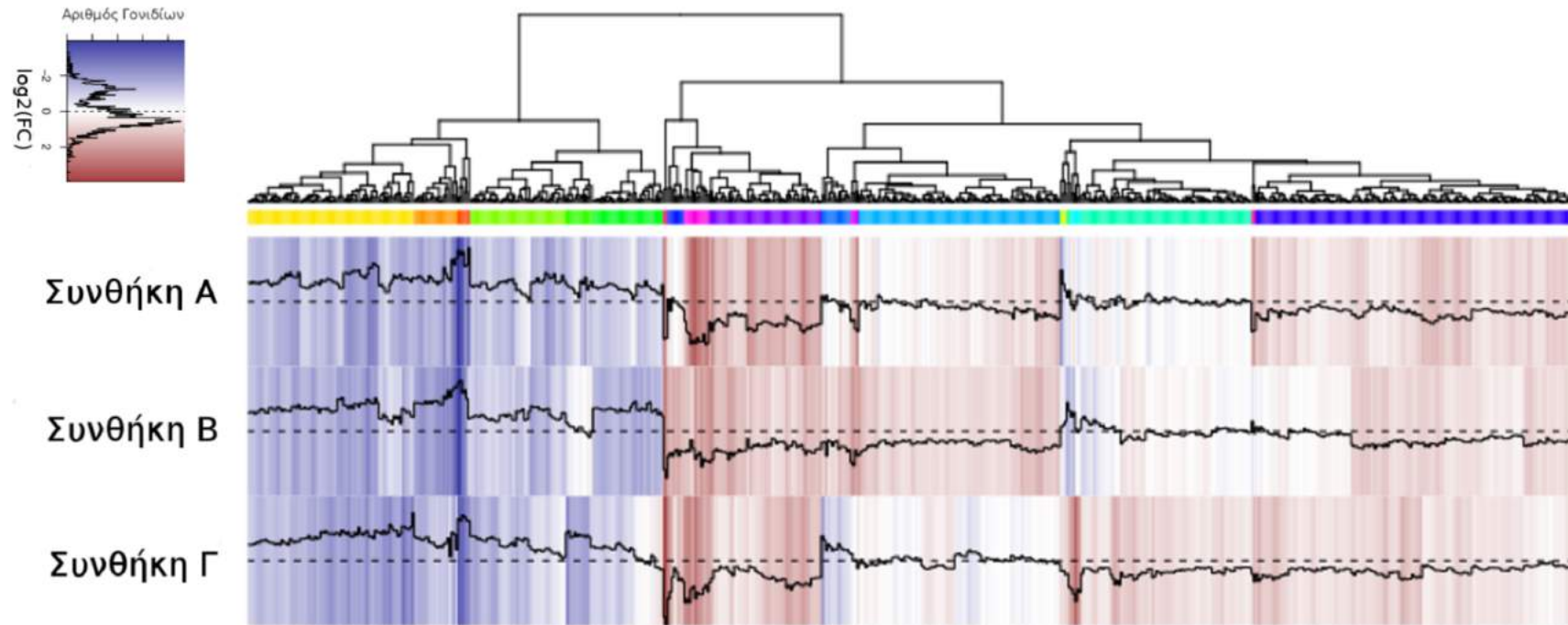
$$\log_2 FC(g) = \log_2 \frac{E(g)_{test}}{E(g)_{control}}$$

Τι σημαίνει $\log_2 FC = 1$ και $\log_2 FC = -1$

RNA-seq / Transcriptomics

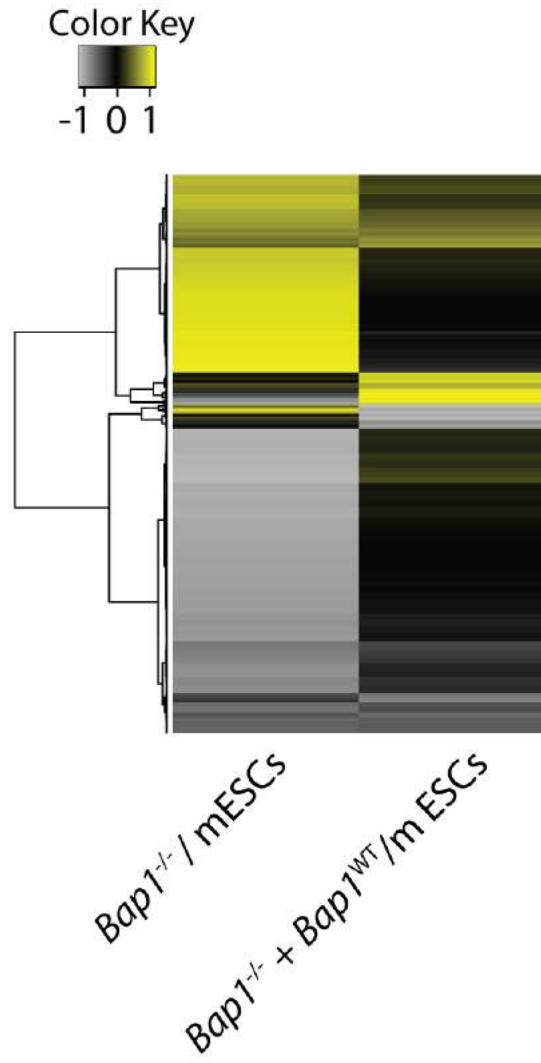


Ανάλυση γονιδίων με κοινά χαρακτηριστικά πρότυπα έκφρασης



Εικόνα 7.6: Θερμικός χάρτης που αναπαριστά τις σχετικές τιμές έκφρασης 650 γονιδίων όπως αυτές μετρήθηκαν σε τρεις διαφορετικές συνθήκες (Α, Β και Γ). Το γαλάζιο αντιστοιχεί σε χαμηλότερη και το κόκκινο σε υψηλότερη έκφραση σε σχέση με την κατάσταση ελέγχου, καθώς στο θερμικό χάρτη εμφανίζονται μόνο σχετικές τιμές έκφρασης. Ο χάρτης συνοδεύεται από ιεραρχική ομαδοποίηση (βλ. Παρακάτω) των γονιδίων με βάση τα πρότυπα έκφρασής τους στις τρεις συνθήκες. Γονίδια που βρίσκονται στον ίδιο κλάδο του δέντρου εμφανίζουν μεγαλύτερη ομοιότητα σε ό,τι αφορά την αυξομείωση των επιπέδων έκφρασης μεταξύ των συνθηκών.

Heatmap



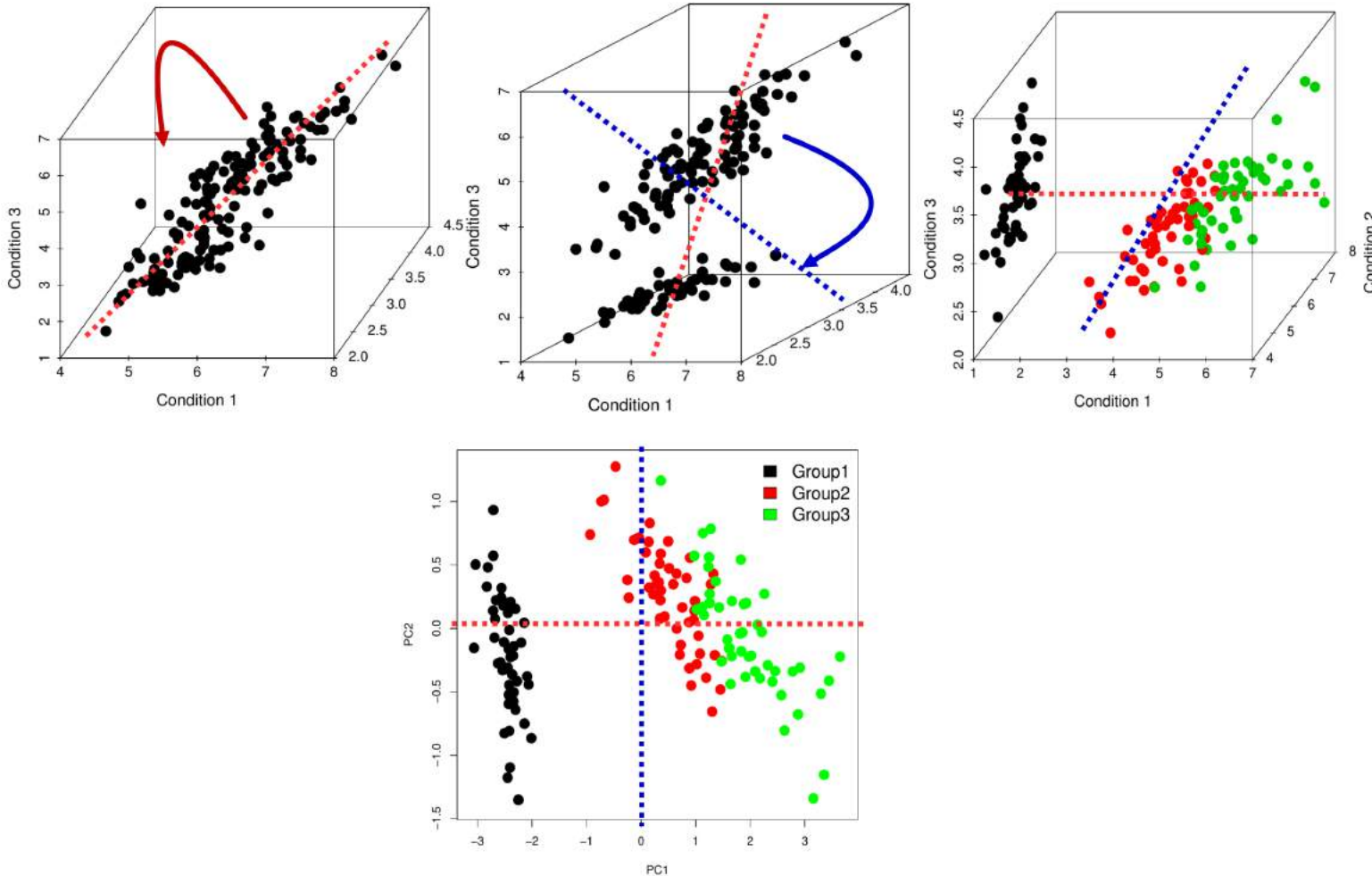
Ανάλυση Κύριων Συνιστωσών (Principal Component Analysis, PCA)

- ✓ Το βασικότερο ποιοτικό χαρακτηριστικό των δεδομένων γονιδιακής έκφρασης είναι η **πολυπλοκότητά τους σε ότι αφορά τις διαστάσεις τους**.
- ✓ Με το όρο “**διαστάσεις**” ή καλύτερα “**διαστασιμότητα**” (dimensionality) αναφερόμαστε στη δομή των δεδομένων που περιλαμβάνουν πολύ συχνά τιμές έκφρασης για χιλιάδες ή δεκάδες χιλιάδες γονιδίων και για μεγάλο αριθμό διαφορετικών συνθηκών.
- ✓ Στην προσπάθειά μας να διακρίνουμε ομάδες, υποσύνολα γονιδίων που συμπεριφέρονται με παρόμοιο τρόπο, μια πρώτη προσέγγιση που θα πρέπει να αναλογιστούμε είναι να προσπαθήσουμε να **μειώσουμε την πολυπλοκότητα στο χώρο των διαστάσεων**.
- ✓ Τεχνικές που αποσκοπούν σ' αυτό ονομάζονται τεχνικές “**μείωσης διαστασιμότητας**” (dimensionality reduction techniques).
- ✓ Η πιο χαρακτηριστική από αυτές είναι η **Ανάλυση Κύριων Συνιστωσών** (Principal Component Analysis, στο εξής **PCA**)



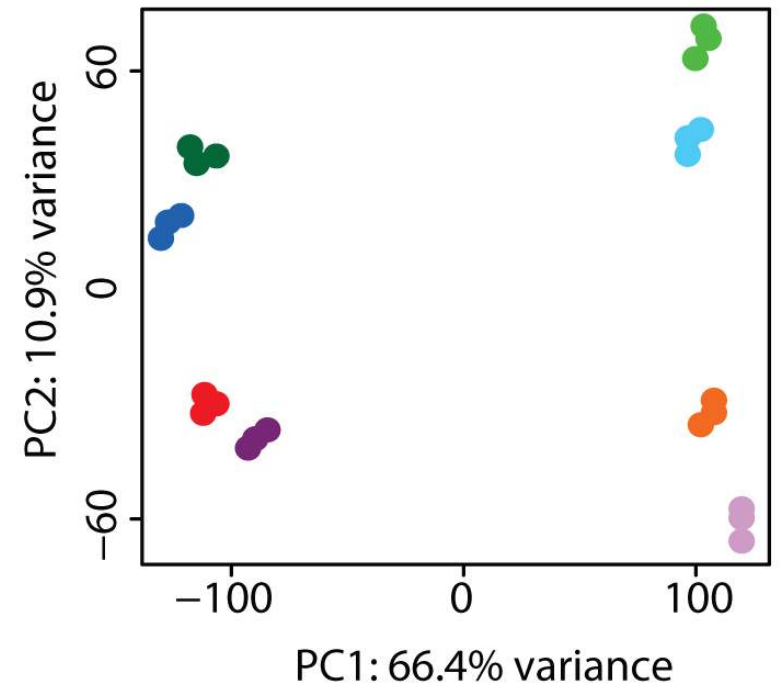
Αναδεικνύει τις **σημαντικότερες διαφορές** μεταξύ των στοιχείων μιας δομής δεδομένων (στην περίπτωσή μας, τα πρότυπα έκφρασης χιλιάδων γονιδίων) κρατώντας και **παρουσιάζοντας εκείνα τα χαρακτηριστικά τους που ευθύνονται για τη μεγαλύτερη ποικιλομορφία στο δείγμα μας**

Ανάλυση Κύριων Συνιστωσών (Principal Component Analysis, PCA)



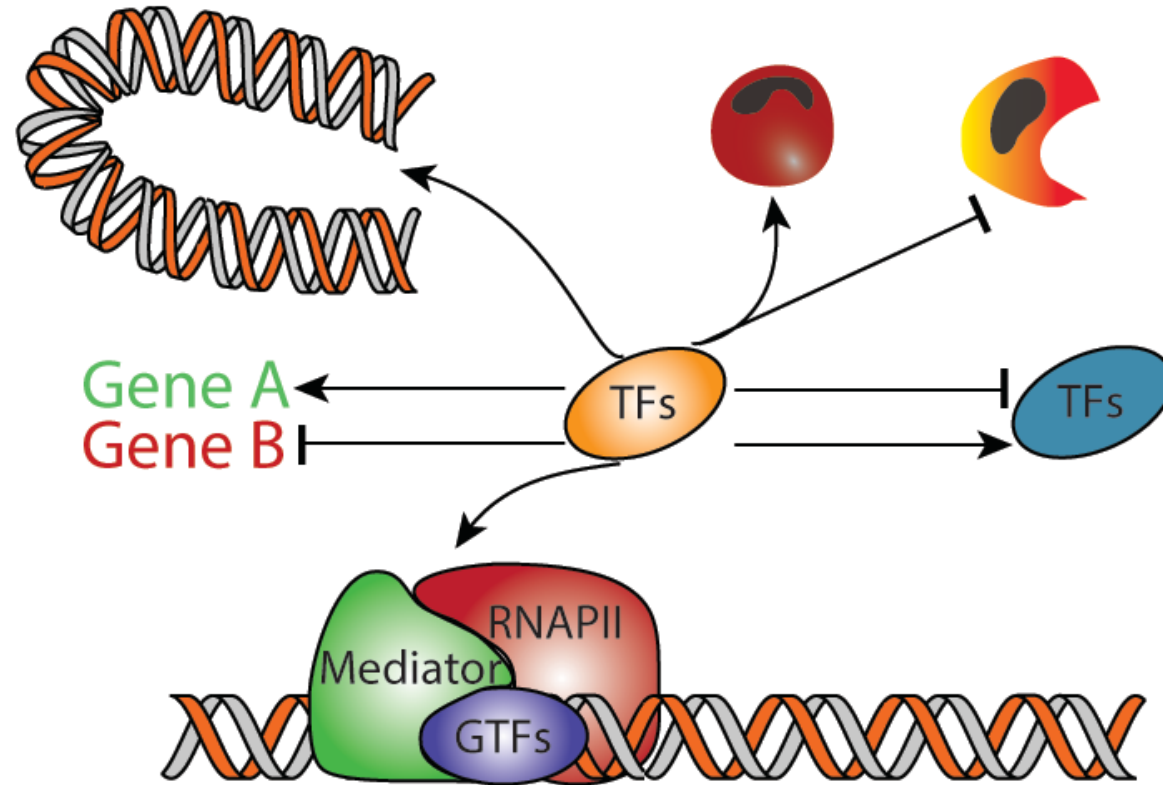
Εικόνα 7.9: Σχηματική αναπαράσταση της διαδικασίας της PCA σε ένα σύνολο τιμών έκφρασης 150 γονιδίων σε τρεις διαφορετικές συνθήκες. Ενώ αρχικά δεν φαίνεται να υπάρχουν διακριτά υποσύνολα, περιστροφή του “νέφους” γύρω από τους άξονες των κύριων συνιστωσών (κόκκινη και γαλάζια διακεκομμένη γραμμή) αναδεικνύει τρία διακριτά υποσύνολα.

Ανάλυση Κύριων Συνιστωσών (Principal Component Analysis, PCA)

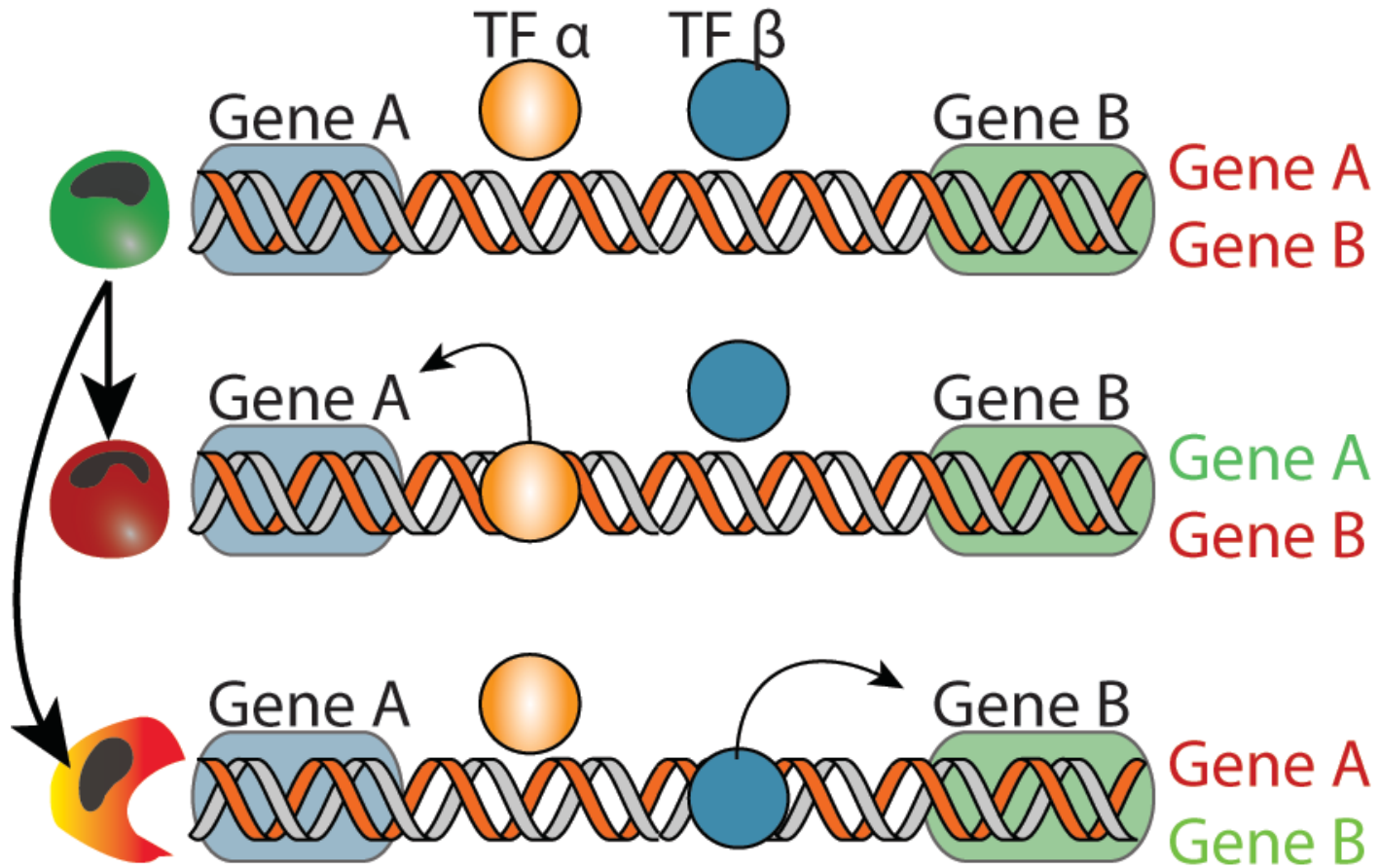


- mESCs
- mESCs^(+ATRA)
- *Bap1*^{-/-}
- *Bap1*^{-/-}(+ATRA)
- *Asx1/2/3*^{-/-}
- *Asx1/2/3*^{-/-}(+ATRA)
- *Foxk1/2*^{-/-} 0h
- *Foxk1/2*^{-/-}(+ATRA)

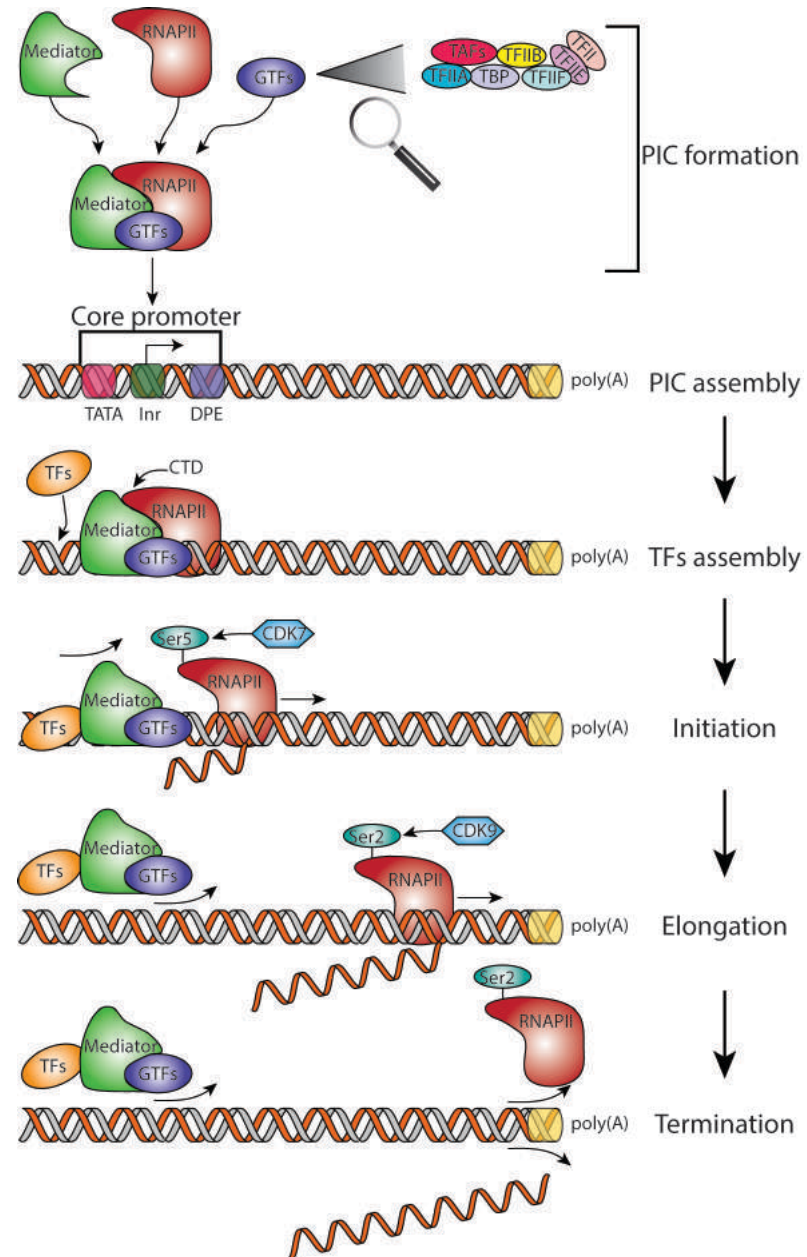
Μεταγραφικοί παράγοντες



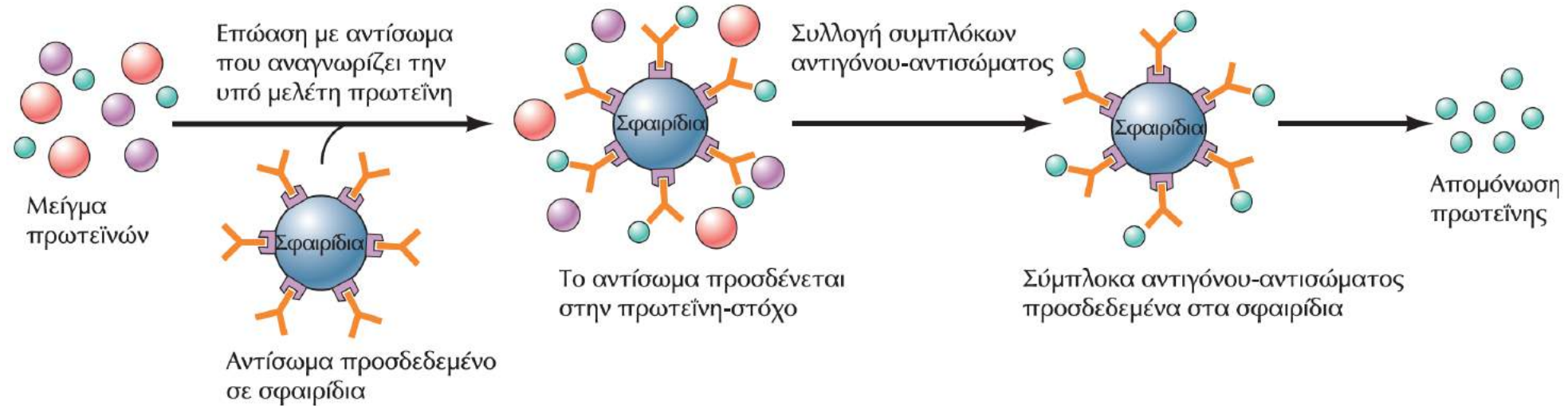
Μεταγραφικοί παράγοντες



Οι ΜΠς συμμετέχουν στην μεταγραφή



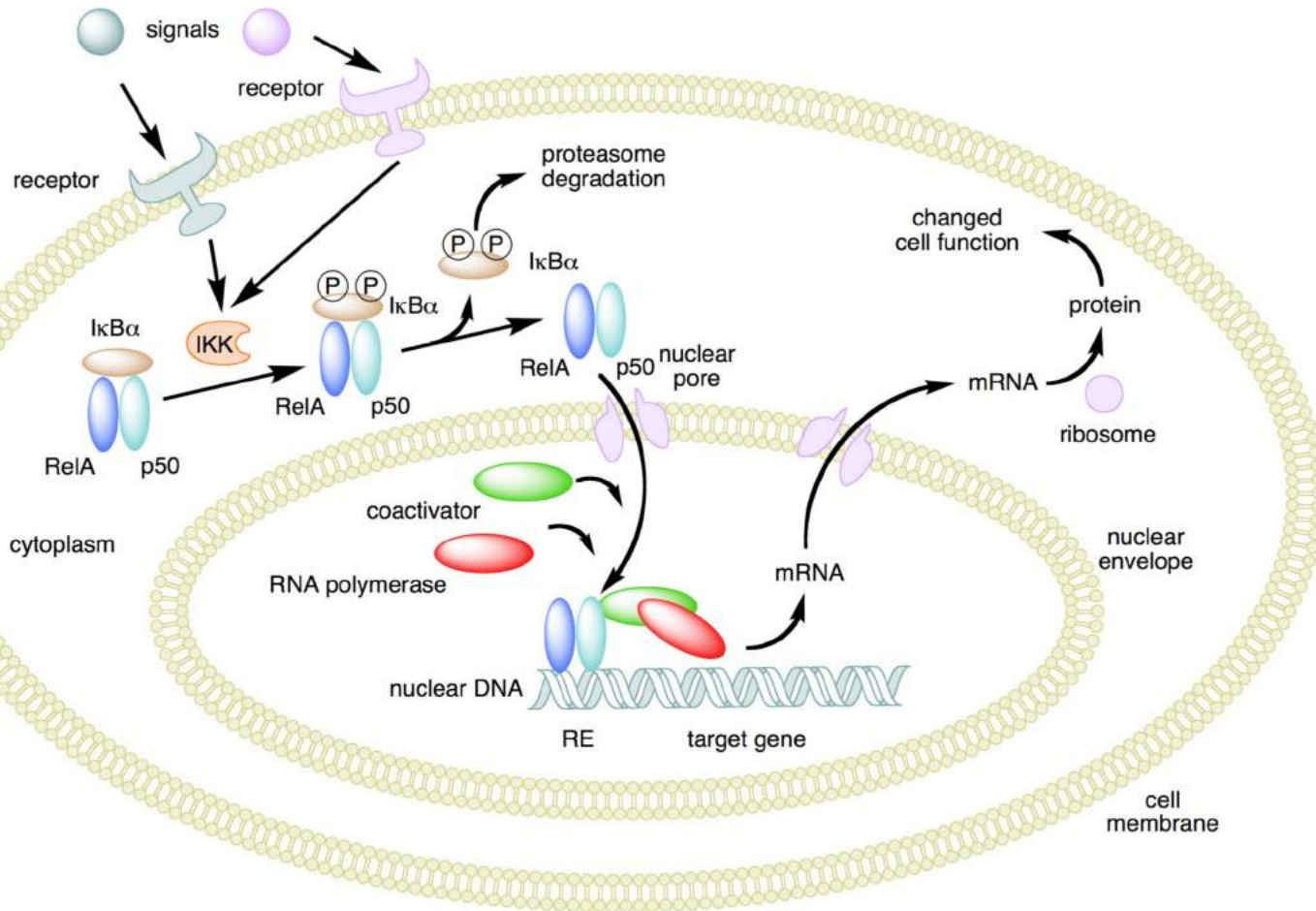
ChIP-seq



ΕΙΚΟΝΑ 5.11 Ανοσοκατακρήμνιση. Ένα μείγμα κυτταρικών πρωτεϊνών επωάζεται με ένα αντίσωμα που είναι προσδεδεμένο σε σφαιρίδια. Το αντίσωμα σχηματίζει σύμπλοκο με την πρωτεΐνη-αντιγόνο που αναγνωρίζει (υπο-δεικνύεται με πράσινο χρώμα). Τα σφαιρίδια συλλέγονται και έτσι απομονώνονται τα σύμπλοκα αντιγόνου-αντισώματος, από τα οποία κατόπιν απομονώνεται η πρωτεΐνη-στόχος του αντισώματος.



Μονοπάτι ενεργοποίησης του NF-κB



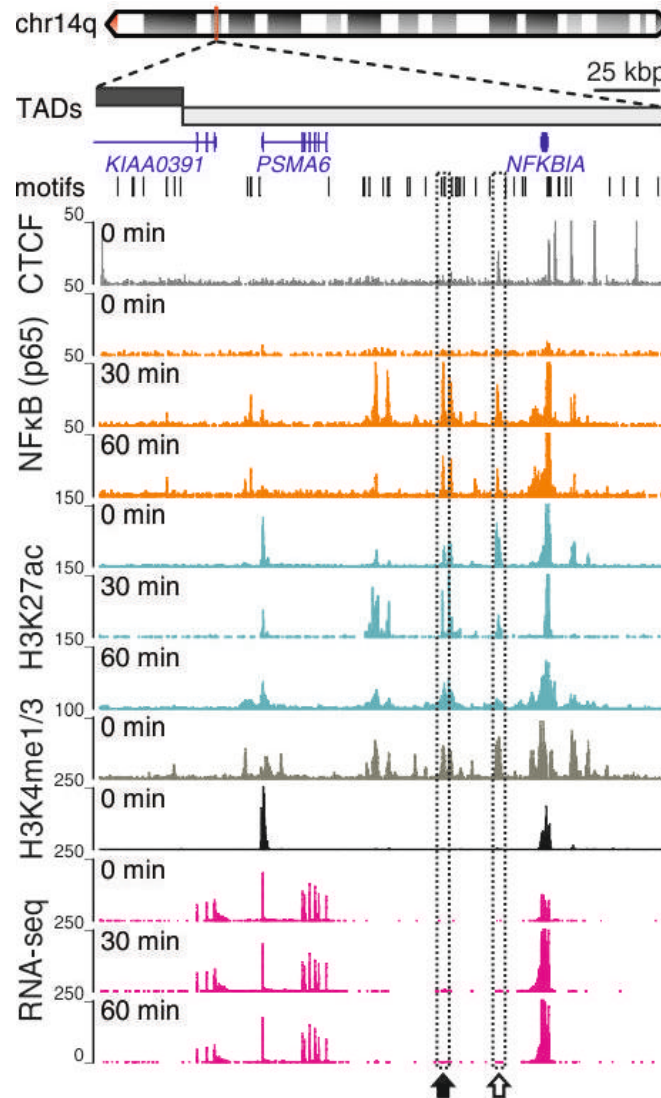
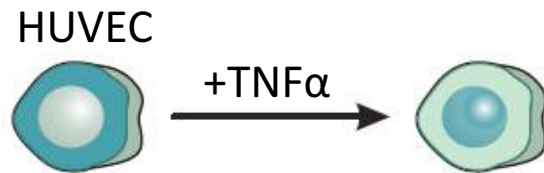
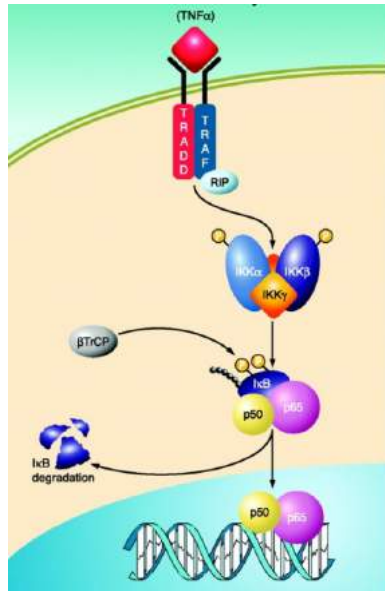
Ο μεταγραφικός παράγοντας NF-κB ρυθμίζει τη μεταγραφή γονιδίων στην περίπτωση ανοσοαπόκρισης σε διάφορα ευκαρυωτικά συστήματα.

Για να προσδεθεί στο DNA, αναγνωρίζει μια αλληλουχία 10 νουκλεοτιδίων: **GGGAATTCCC**

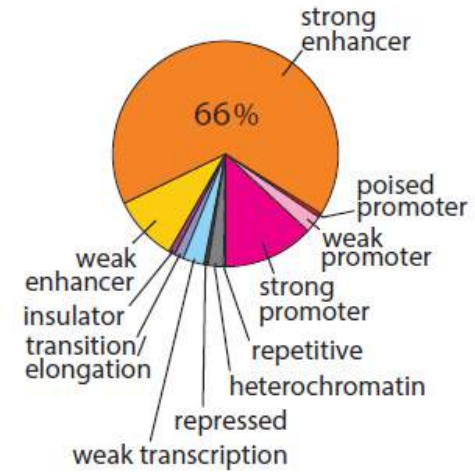
Πως προέκυψε αυτή η πληροφορία;

ChIP-seq

Ο NFκB προσδένεται σε ήδη ενεργοποιημένους ενισχυτές



HMM motifs of 30-min p65 peaks



DNA (binding) motif



Στην Εικόνα 3.2 φαίνεται σχηματικά η δομή του διμερούς NF-κΒ (p65, με τα στοιχεία της δομής του χρωματισμένα ως γαλάζιες β-πτυχωτές επιφάνειες και πορτοκαλί α-έλικες) καθώς βρίσκεται προσδεμένος σε μια περιοχή δίκλωνου DNA (με βιολετί)

(Εικόνα από DonabelSDSU.bot: CC0 2.5, από Wikimedia Commons).

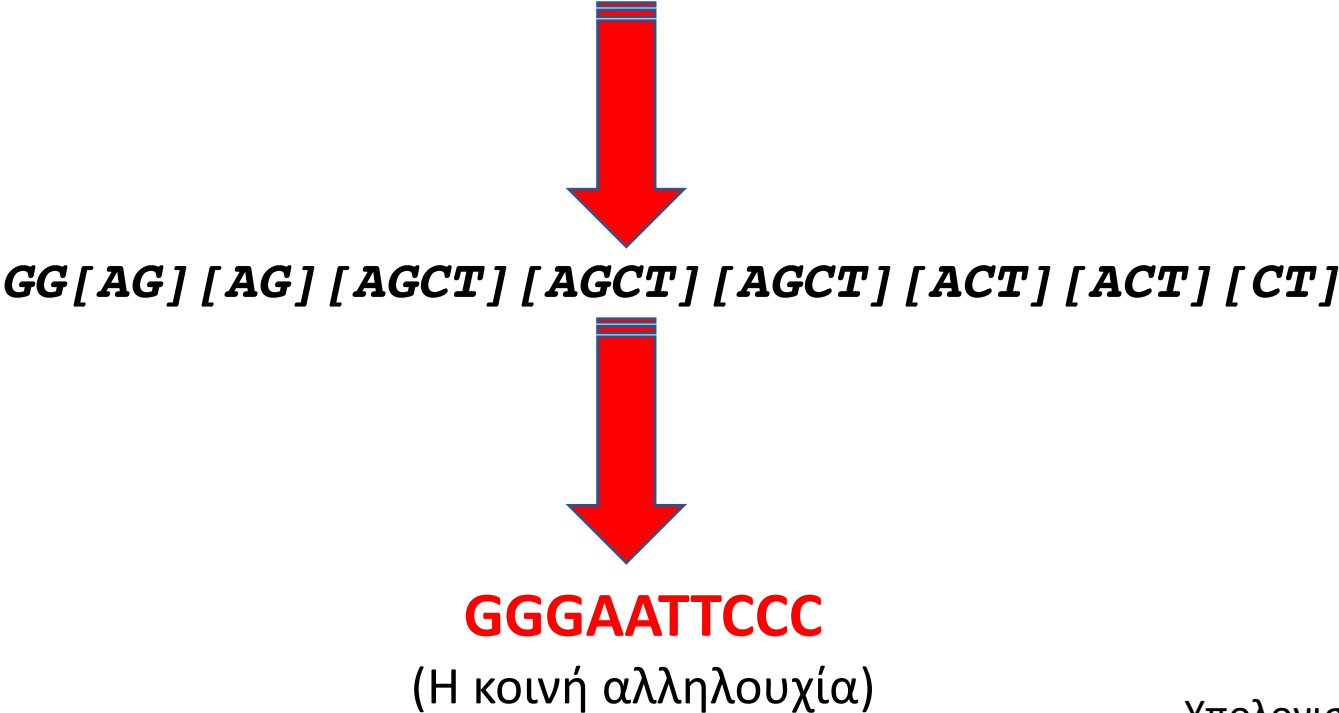
Οι ΜΠς προσδέονται στο DNA σε συγκεκριμένες θέσεις που ονομάζονται “σημεία πρόσδεσης μεταγραφικών παραγόντων” (transcription factor binding sites ή για συντομία TFBS).

Οι θέσεις αυτές έχουν μήκος 4-12 bp και είναι χαρακτηριστικές για κάθε ΜΠ. Έχουν δηλαδή χαρακτηριστικά μοτίβου. **GGGAATTCCC** -> **DNA (binding) motif**

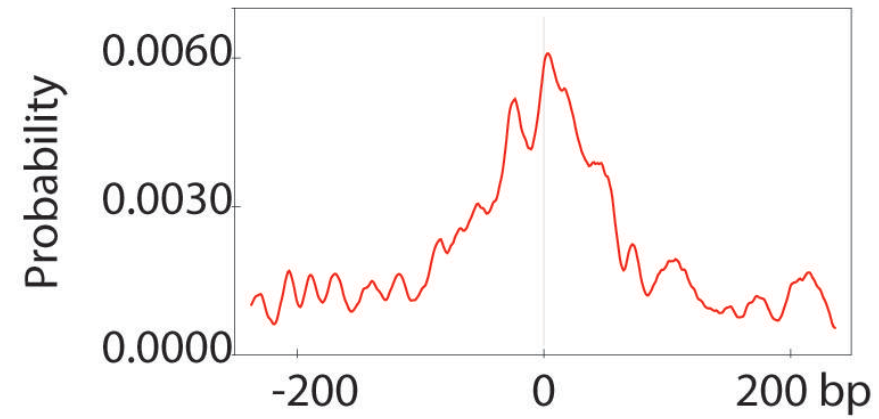
Ορισμός ενός βιολογικού μοτίβου συναινετικών αλληλουχιών (consensus)

GGGGCATTCC	GGGATATCCC	GGGAATTCCC	GGGAATGTCC	GGGATATTTT	GGGGCCTCCC	GGGAATTTCC	GGGACTGCCC
GGGAAATTCC	GGGAAATCCC	GGGAATTCCC	GGGACTTACC	GGGGATTTCC	GGGAATTTCC	GGGACATTTCC	GGGAAATTTCC
GGAAATTTCC	GGGAATTTCC	GGGGATTTCC	GGGGTTTAC	GGGAAGGTCC	GGGGCTTTCC	GGGGCTTTCC	GGGAAATTTCC
GGGGCTTTCC	GGGACTTTCC	GGGACATTTCC	GGGAATTTCC	GGGACATTTCT	GGGACAGCCC	GGGGCTTTAC	GGGACTTTCC
GGGAATTCAC	GGGAAATCCC	GGAGCTTTCC	GGGACTTTCC	GGGAAACCCC	GGGGCTTTCC	GGGAATTTCC	GGGAAATTTCC
GGGACTTTCC	GGGAAATTTCT	GGGAATTTCC	GGGACTTTCC	GGGACTTTCC	GGGGATTTCC	GGGACATTTCC	GGGAAATTTCC
GGGATGTTCC	GGGGTCTCCC	GGGACTGTCC	GGGAATTTCC	GGGACTTTTAC	GGGAATTTCC	GGGACTTTCC	GGGGCGTTCC
GGGGTTTTCC	GGGAAATTTCC	GGGAATTTCC	GGGGATTTCC	GGGAATGCCC	GGGGATTTCC	GGGAATTTCC	GGGATTTTTCC
GGGGAAATTTCC	GGGACTTTCC	GGGATTTTCC	GGGAAGTTCC	GGGAAATTTCC	GGGAATTTCC	GGGAATTTTAC	GGGAAATTTCC
GGGGGTTTAC	GGGACTTTCC	GGGAATTTCC	GGGAATTTCC	GGGACATTTCC	GGGAATTTTAC	GGGACTTTCC	GGGACTTTCC
GGGAATTTCC	GGGACTTTCC	GGGGACTTTCC	GGGACTTTTAC	GGGACTTTTCC	GGGATACTTTCC	GGGGATGTTAC	GGGATATTTCC
GGGAATTTCCC	GGGACTTTCCC	GGGACTTTTAC	GGGGTTATTTCC	GGGAATCTTTCC	GGGAATTTTCC	GGGACATTTCT	GGAAATTTCCC
GGGAAACTCT	GGGGTTTTCC	GGGATTTTCC	GGGGCGTTCC	GGGAAACTCT	GGGGTTTTCC	GGGATTTTCC	GGGGCGTTCC

Πίνακας 3.1: 104 σημεία πρόσδεσης του NF-κB από το γονιδίωμα του ποντικίου (*Mus musculus*)



Παράδειγμα motif analysis



GTAAACA $p=3.9e-55$

Πίνακες Βαθμονόμησης ανά θέση (Position-specific Scoring Matrices, PSSM)

Μοτίβο NF-κΒ (<i>P</i>)											Πίνακας Υποβάθρου (<i>Q</i>)										
Νουκλεοτίδιο	1	2	3	4	5	6	7	8	9	10	Νουκλεοτίδιο	1	2	3	4	5	6	7	8	9	10
A	0.00	0.00	0.03	0.76	0.49	0.23	0.01	0.01	0.10	0.00	A	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
C	0.00	0.00	0.00	0.00	0.37	0.03	0.04	0.38	0.88	0.97	C	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27
G	1.00	1.00	0.97	0.24	0.01	0.05	0.07	0.00	0.00	0.00	G	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33
T	0.00	0.00	0.00	0.00	0.13	0.69	0.88	0.61	0.02	0.03	T	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23

$$R = \log_2(P_{i,j}/Q_{i,j})$$

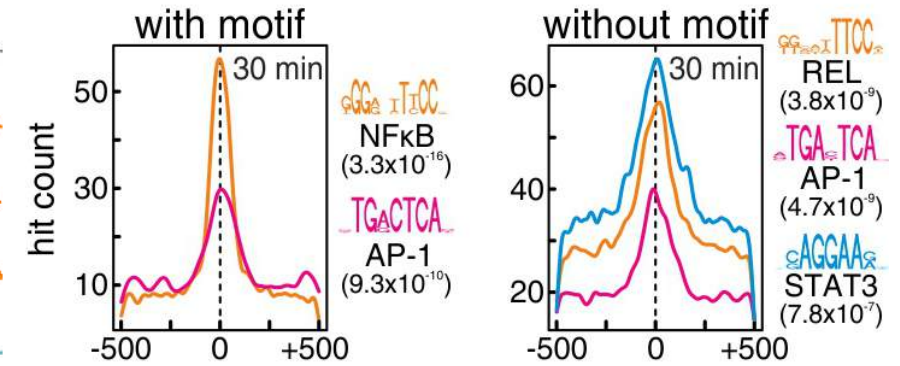
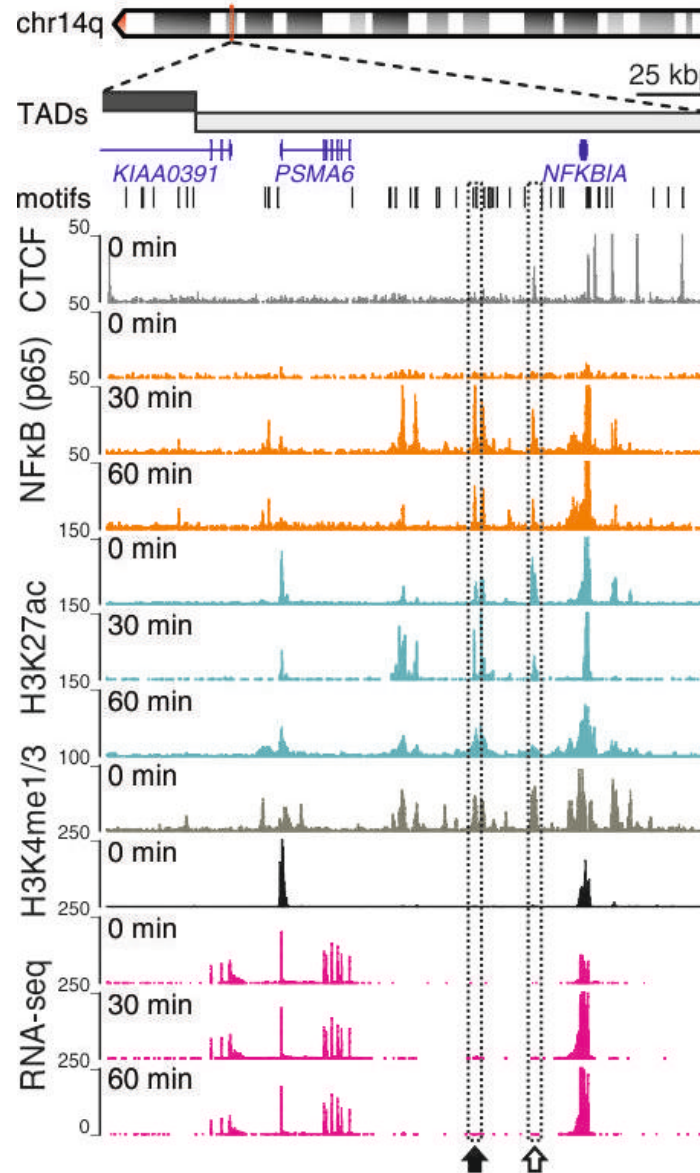
Νουκλεοτίδιο	1	2	3	4	5	6	7	8	9	10
A	-7.3	-7.3	-2.4	2.2	1.6	0.5	-3.9	-3.9	-0.7	-7.3
C	-8.1	-8.1	-8.1	-8.1	0.5	-3.1	-2.7	0.5	1.7	1.8
G	1.6	1.6	1.6	-0.5	-4.9	-2.7	-2.2	-8.4	-8.4	-8.4
T	-7.8	-7.8	-7.8	-7.8	-0.8	1.6	1.9	1.4	-3.5	-2.9

Position-Specific Scoring Matrix, PSSM

Εικόνα 3.4: Δημιουργία Πίνακα Βαθμονόμησης ανα Θέση (PSSM).

Κάθε μοτίβο είναι σημείο πρόσδεσης; Συγκριτική γονιδιωματική

Μια αναζήτηση μοτίβων με της χρήση της δειγματοληψίας Gibbs σε ένα σύνολο T αλληλουχιών μπορεί συχνά να αποφέρει ένα μεγάλο αριθμό πιθανών σημείων πρόσδεσης, αυτό δε σημαίνει όμως σε καμία περίπτωση ότι αυτά είναι και πραγματικά, φυσικά σημεία αλληλεπίδρασης/πρόσδεσης της πρωτεΐνης με το DNA.

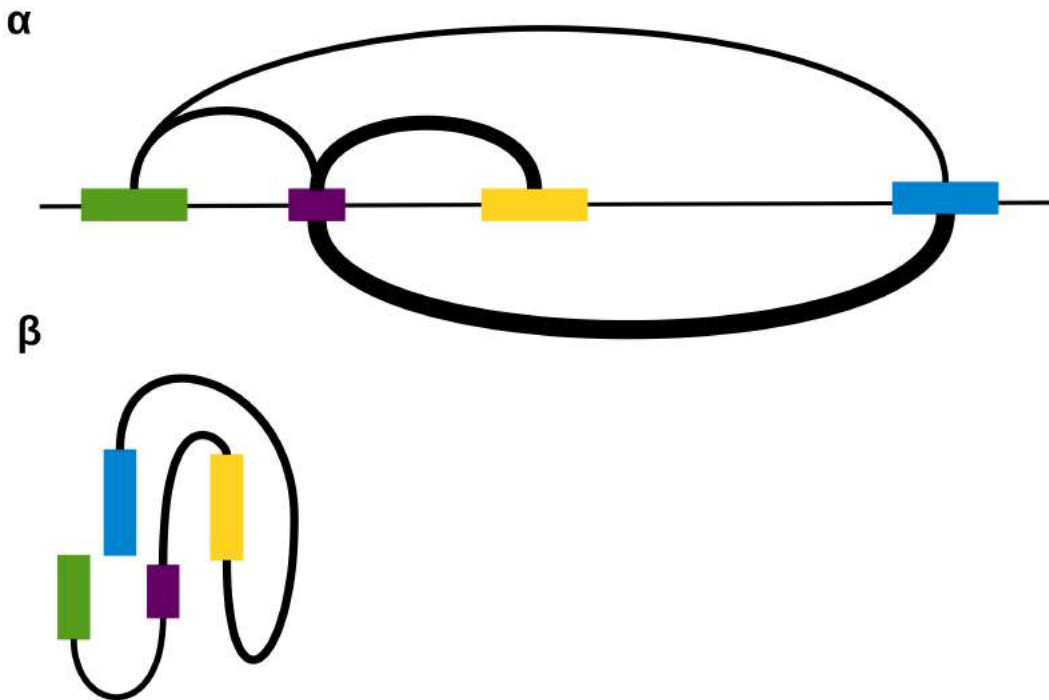


Kolovos et al., 2016

Η μεγάλη κλίμακα. Δομή της χρωματίνης σε τρεις διαστάσεις



Η μεγάλη κλίμακα. Δομή της χρωματίνης σε τρεις διαστάσεις

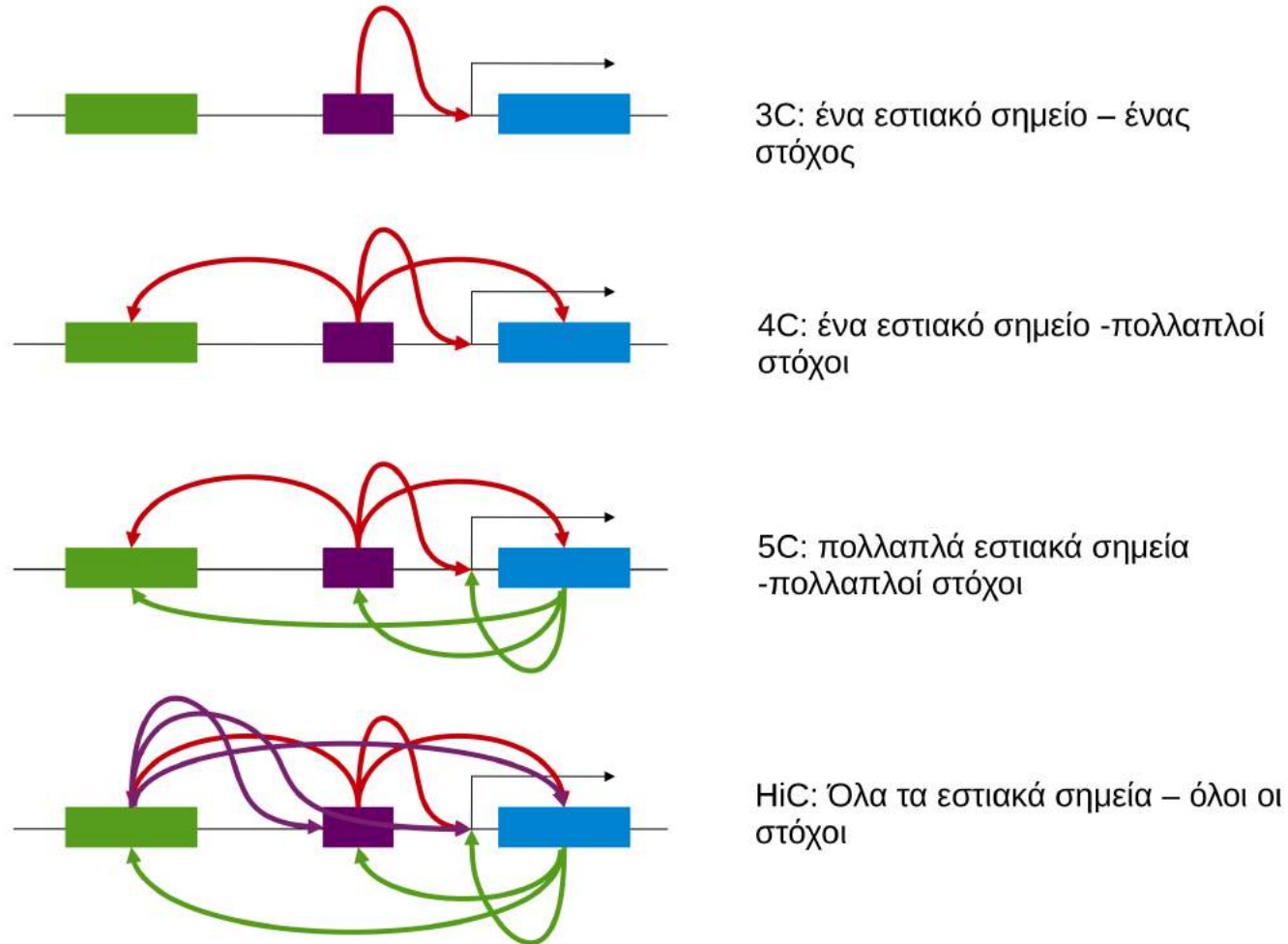


Εικόνα 5.9: Σχηματική αναπαράσταση της διερεύνησης της δομής της χρωματίνης μέσω χρωμοσωμικών αλληλεπιδράσεων. α) Οι αλληλεπιδράσεις μεταξύ τεσσάρων στοιχείων του γονιδιώματος ποσοτικοποιούνται με βάση τη συχνότητα των επαφών τους σε πειράματα σύλληψης χρωμοσωμικής διαμόρφωσης. Εδώ το πάχος κάθε καμπύλης αναλογεί στη συχνότητα της αλληλεπίδρασης. β) Ανασύσταση μιας πιθανής ανώτερης διαμόρφωσης των τεσσάρων στοιχείων με βάση τη συχνότητα των αλληλεπιδράσεων. Η εγγύτητα των στοιχείων αντιστοιχεί στο πάχος των καμπύλων γραμμών του α).

Κάτω από αυτό το αδιανόητο επίπεδο συμπύκνωσης, το γονιδίωμα επιτρέπει τη συντονισμένη διενέργεια μιας σειράς πολύπλοκων μοριακών λειτουργιών.

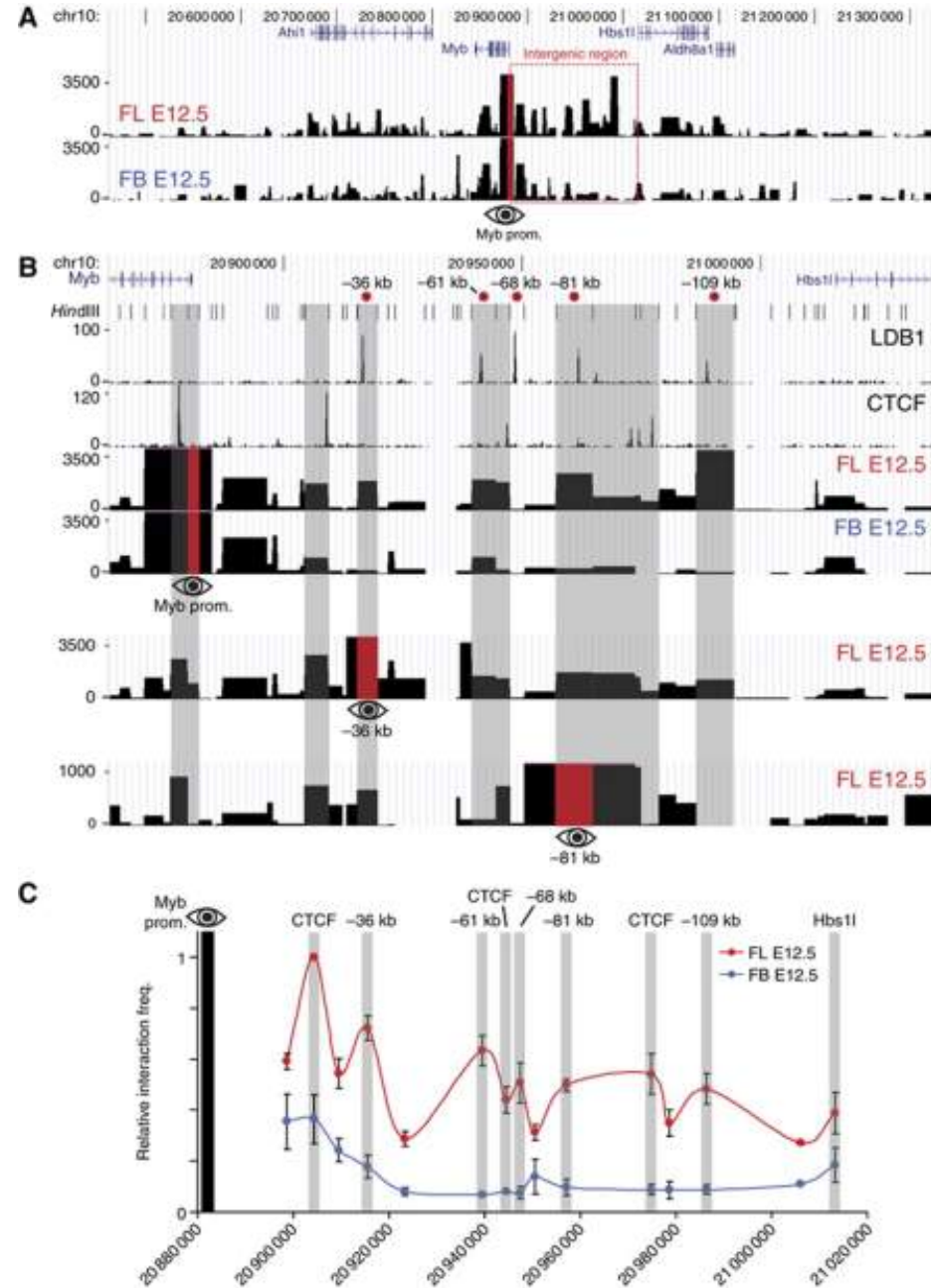
Μελέτη της **τριδιάστατης οργάνωσης του γονιδιώματος** -> **3D Chromatin Organization**

Η μεγάλη κλίμακα. Δομή της χρωματίνης σε τρεις διαστάσεις

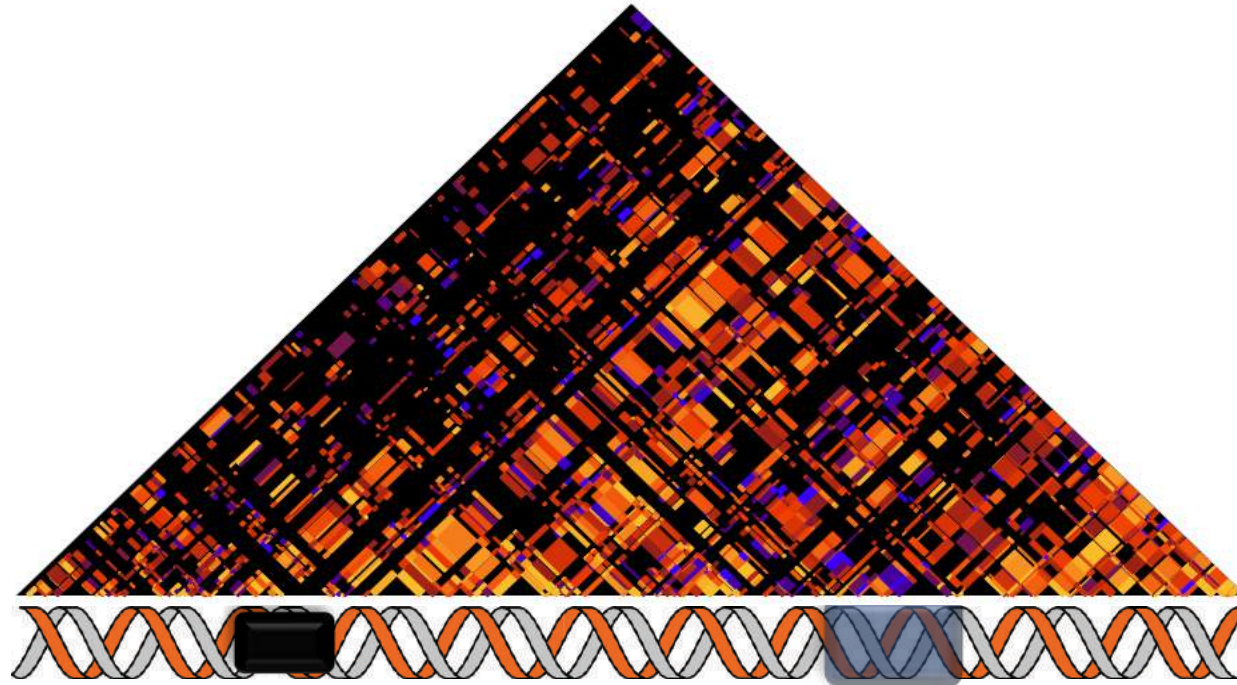


Εικόνα 5.10: Σχηματική αναπαράσταση των ειδών αλληλεπιδράσεων που εντοπίζονται με τις διαφορετικές παραλλαγές της 3C μεθοδολογίας.

Πως διαβάζουμε 3C-seq/4C-seq



Πως διαβάζουμε τους T2C/Hi-C χάρτες;

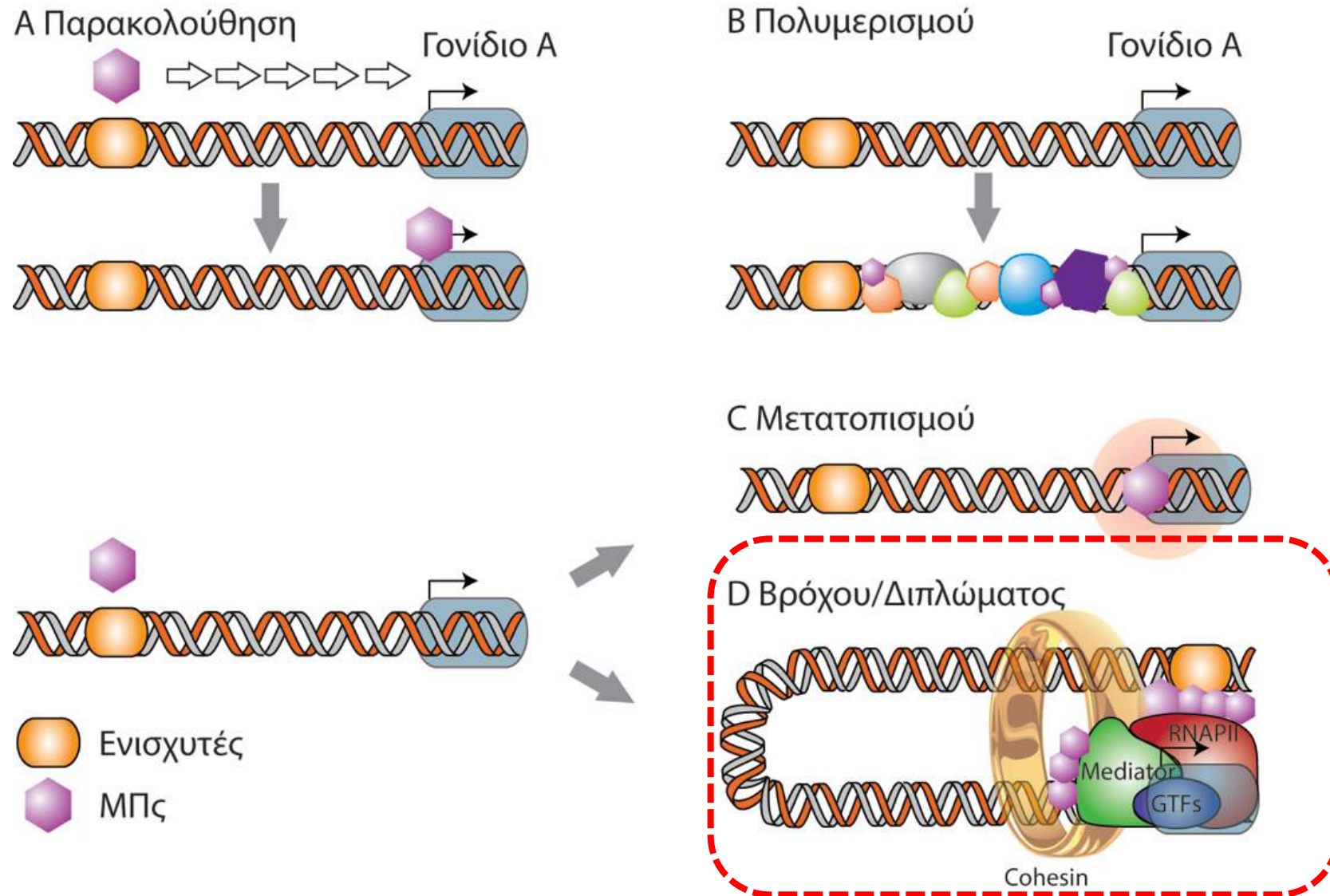


Γονίδιο

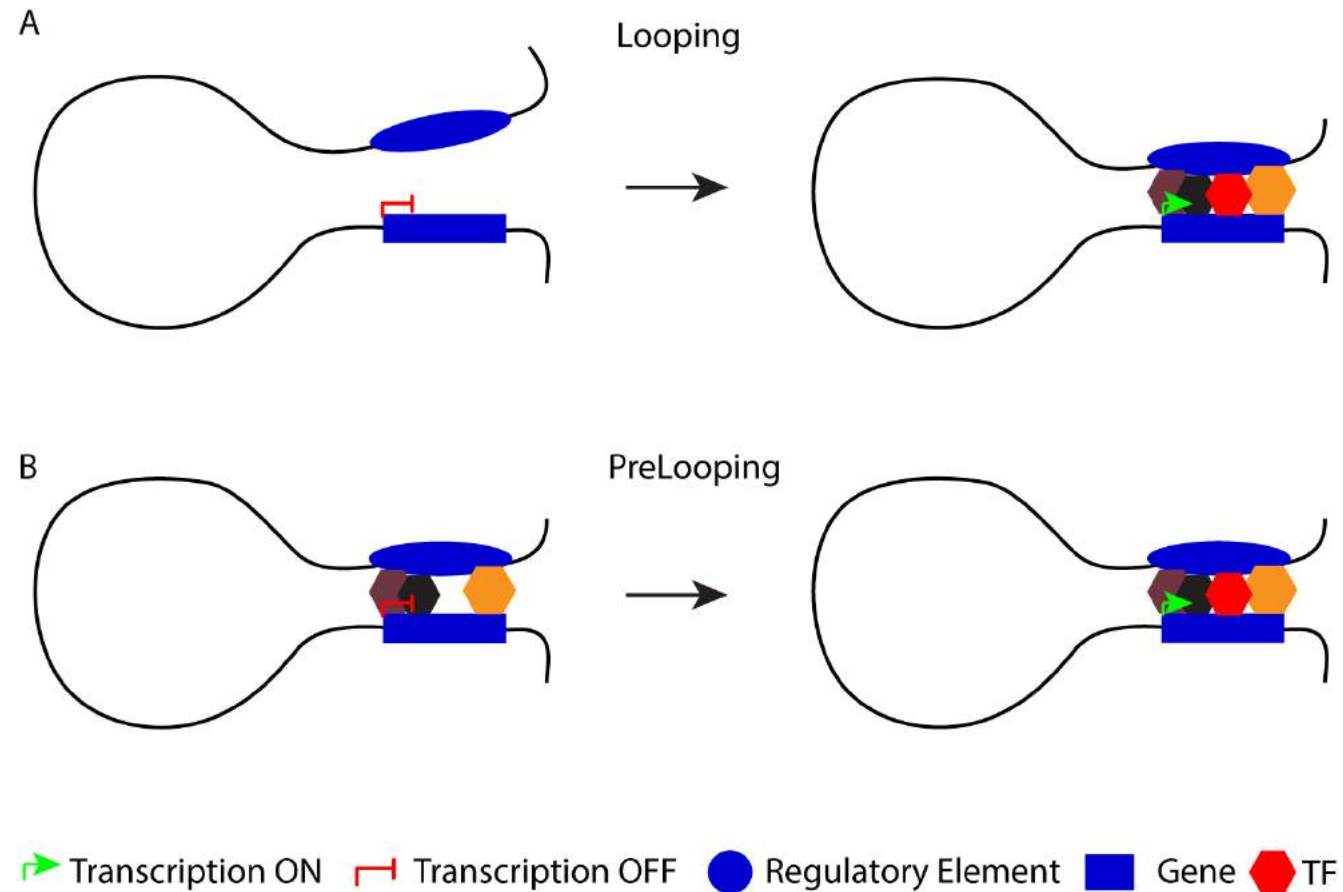


Ρυθμιστικό στοιχείο

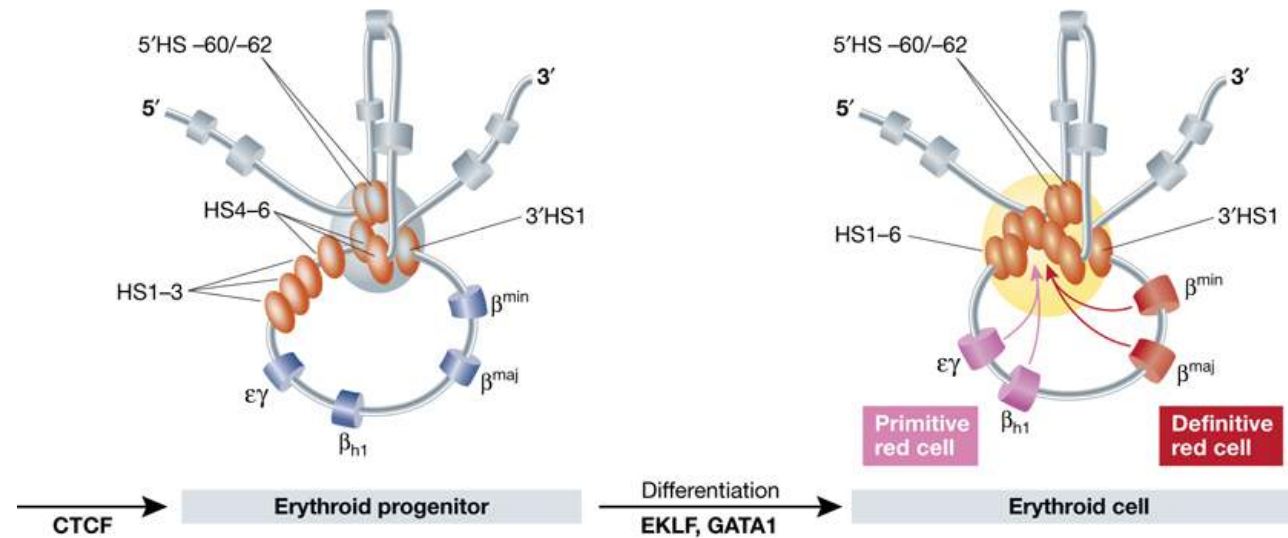
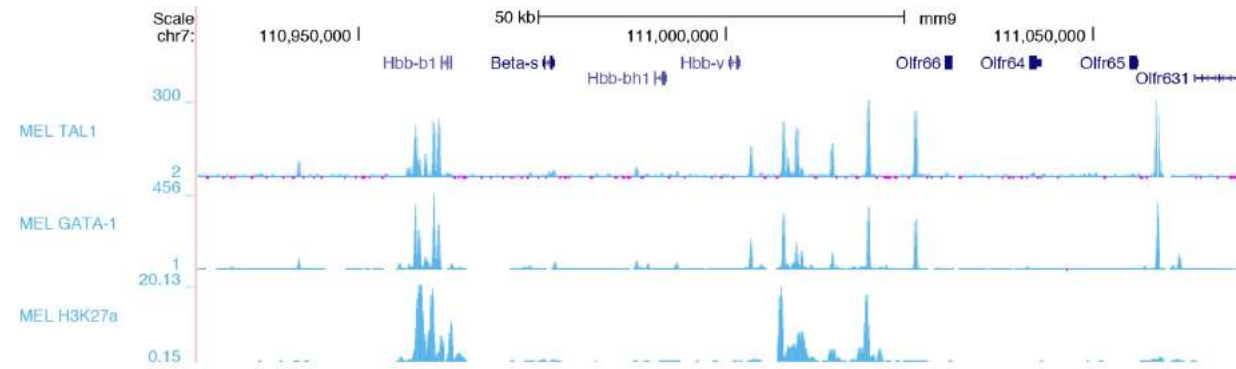
Υπάρχοντα μοντέλα για την λειτουργία των ρυθμιστικών στοιχείων



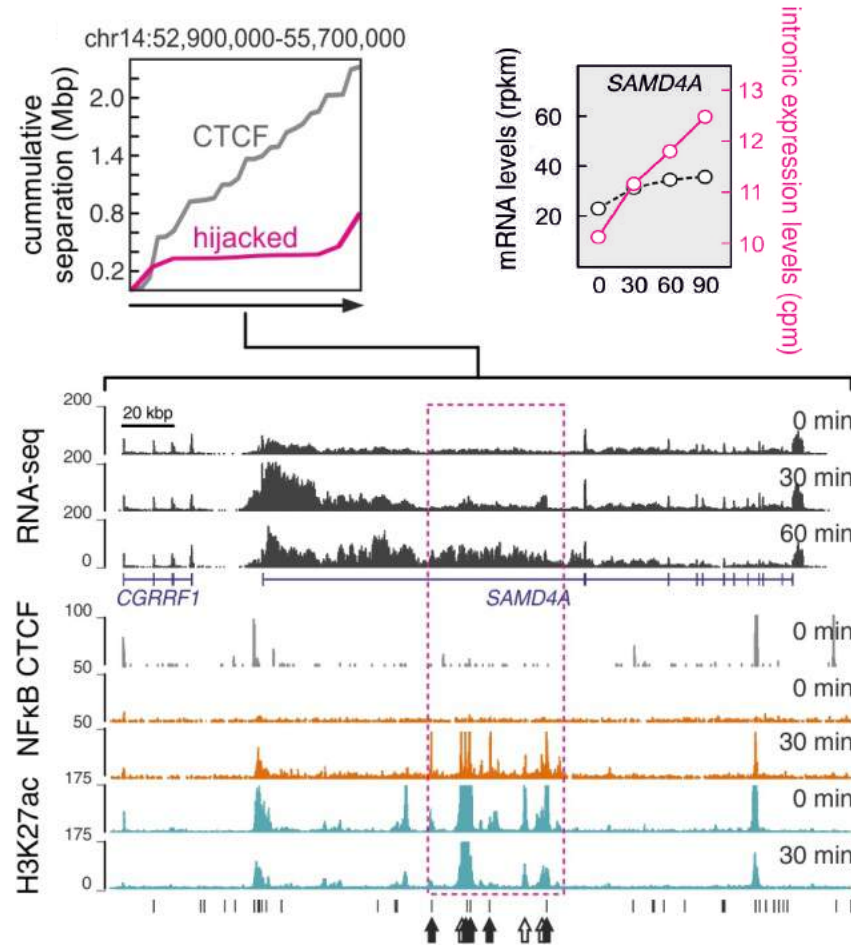
Καινούριοι τρόποι για την ρύθμιση της έκφρασης των γονιδίων



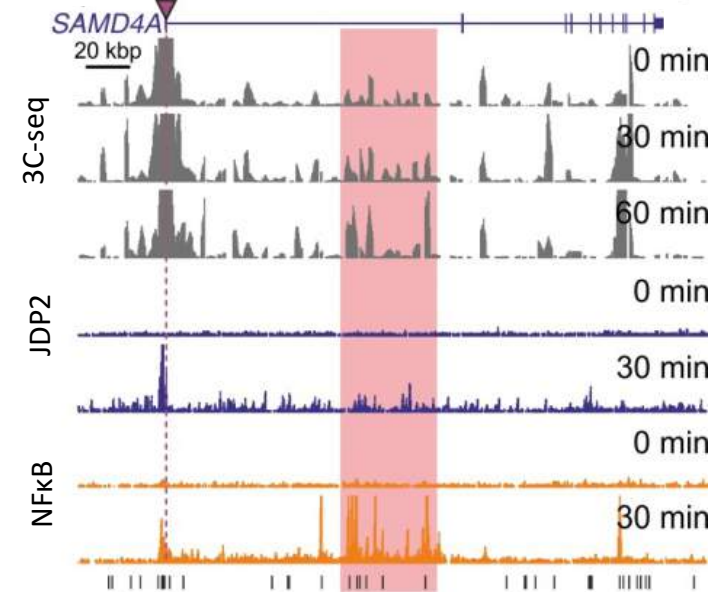
Το Looping μοντέλο ρυθμίζει την έκφραση των globin (Hbb)



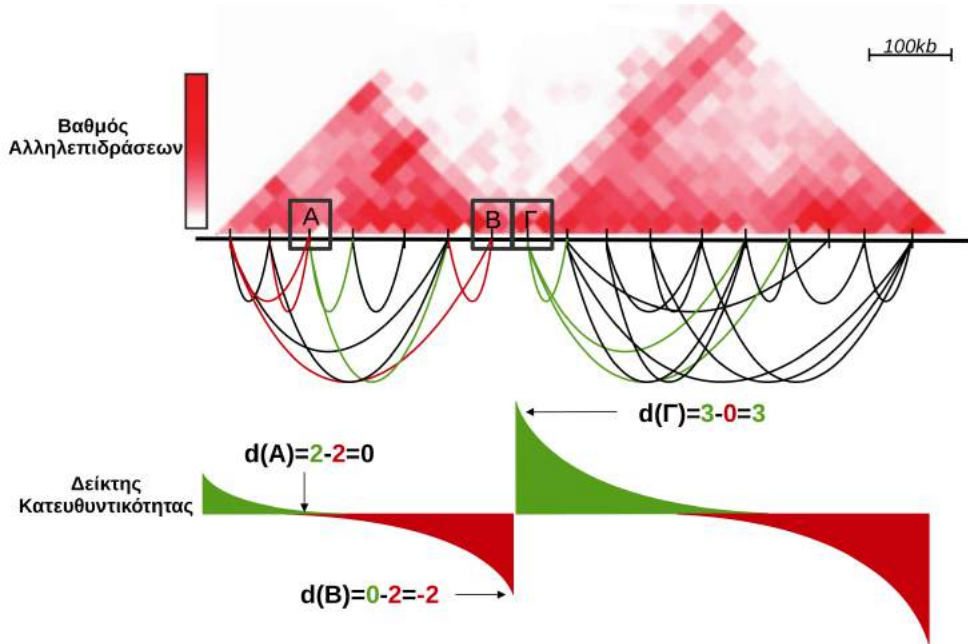
Pre-looped μοντέλο



Pre-looped ενισχυτές



Τοπολογικά διασυνδεδεμένες χρωσωμικές επικράτειες



Εικόνα 5.11: Οι Τοπολογικά Διασυνδεδεμένες Επικράτειες (TAD) ορίζονται ως περιοχές του γονιδιώματος μεταξύ των οποίων ο βαθμός αλληλεπιδράσεων είναι σημαντικά υψηλός. Ο δείκτης κατευθυντικότητας για μια περιοχή X ισούται με τον αριθμό των καθοδικών (downstream) αλληλεπιδράσεων μείον αυτόν των ανοδικών (upstream). Σημεία απότομης αλλαγής του δείκτη αυτό (όπως το σημείο Β-Γ στο σχήμα) αποτελούν όρια μεταξύ συνεχόμενων TAD.

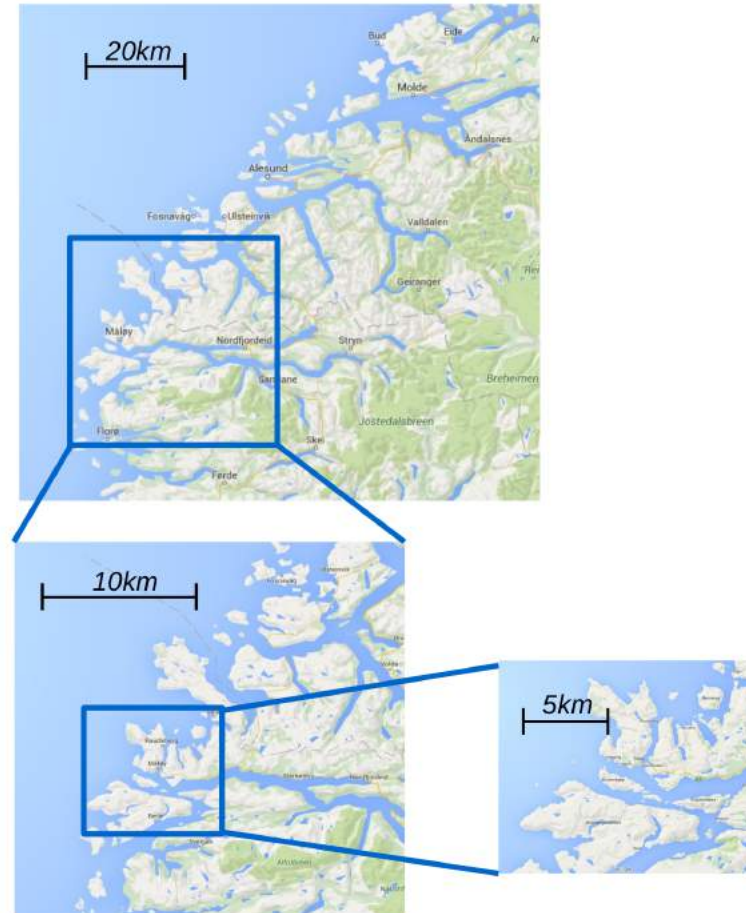
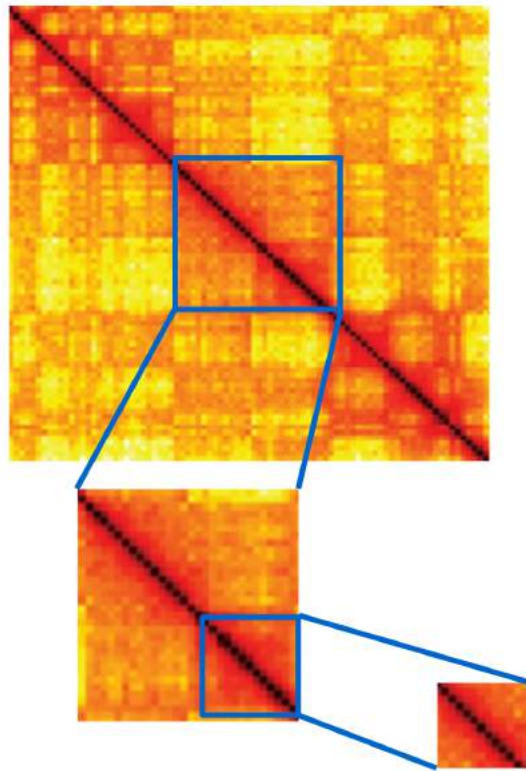
Με ποιον τρόπο καθορίζονται τα όρια των περιοχών, μέσα στις οποίες η χρωματίνη θα έχει συγκεκριμένη διαμόρφωση;

Υπαρξη γονιδιωματικών επικρατειών με μεγάλη συχνότητα εσωτερικών αλληλεπιδράσεων, μεταξύ των οποίων όμως οι αλληλεπιδράσεις είναι σπάνιες

➤ **Τοπολογικά Διασυνδεδεμένες Επικράτειες** (Topologically associated Domains, ή για συντομία **TAD**)

- Έχουν μέσο μήκος μερικές εκατοντάδες χιλιάδες βάσεις
- Ορίζουν διακριτά τμήματα του γονιδιώματος μεταξύ των οποίων οι αλληλεπιδράσεις είναι ελάχιστες
- Τα γονίδια μέσα στα TADs, εκφράζονται συνήθως με τον ίδιο τρόπο
- “Δείκτης κατευθυντικότητας” (directionality index) -> ορισμός ορίων των TADs

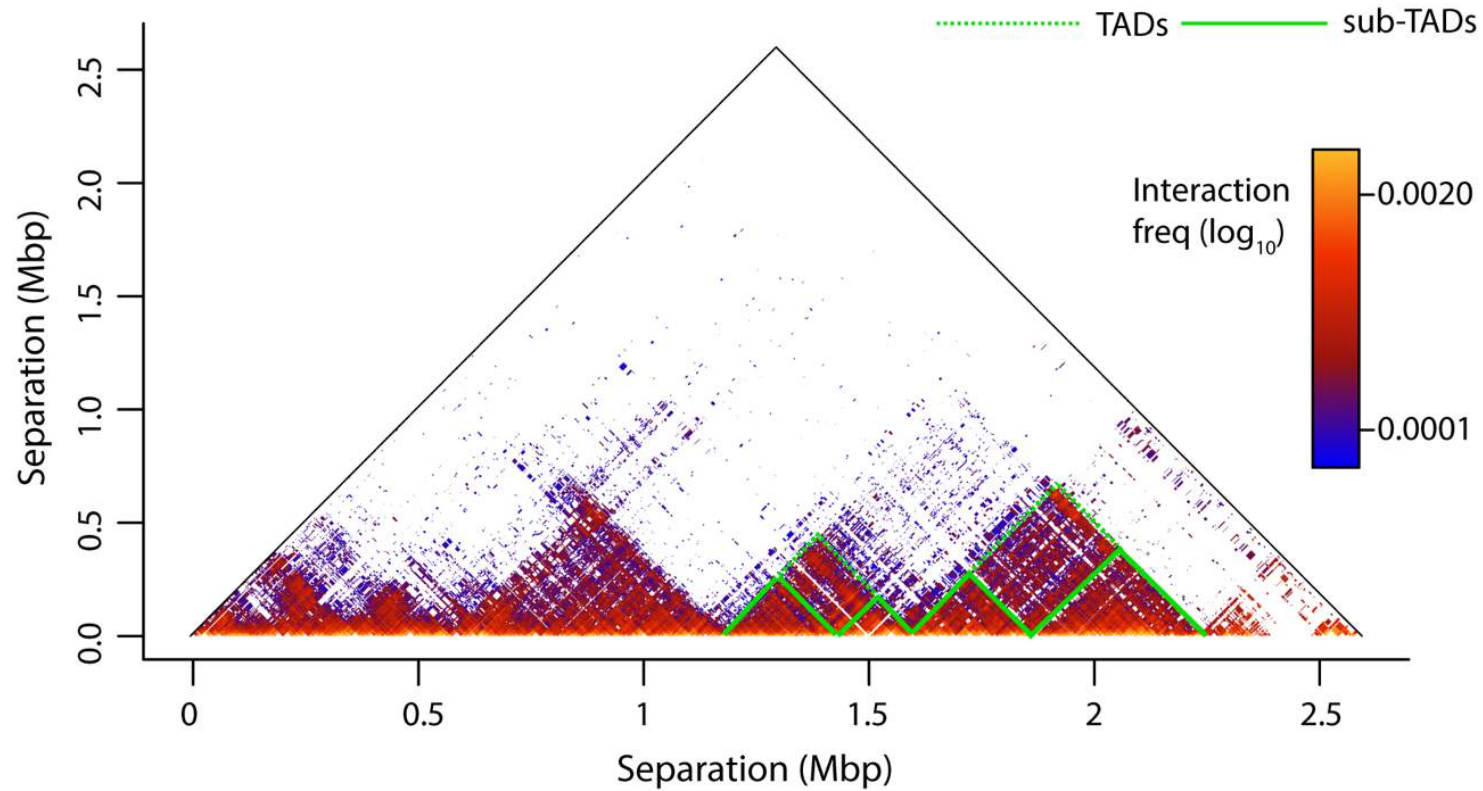
Χρωσωμικοί χάρτες σε τρεις διαστάσεις. Το μορφοκλασματικό γονιδίωμα



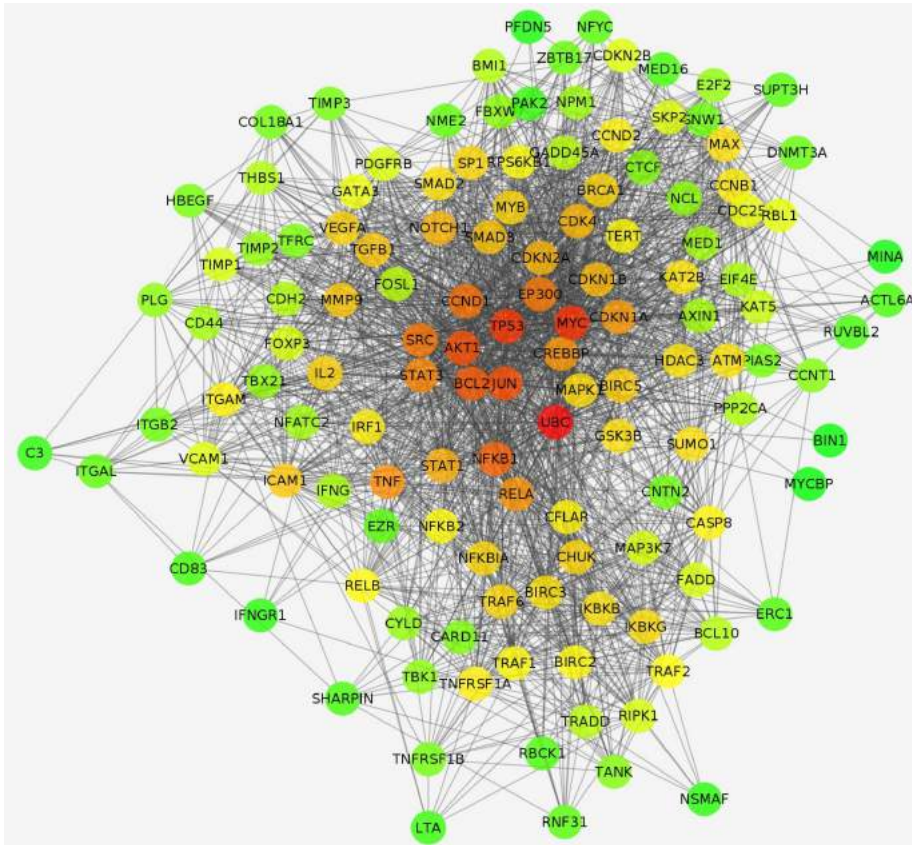
Εικόνα 5.12: Αυτο-ομοιότητα α) σε χάρτες χρωσωμικών αλληλεπιδράσεων στο ανθρώπινο γονιδίωμα (Εικόνα από ένα HiC πείραμα) β) στο περίγραμμα της ακτογραμμής των Νορβηγικών φιορδ. Ανάλογες εικόνες αναπαράγονται σε διαφορετικές κλίμακες.

Μορφοκλασματικά (**fractal**) αντικείμενα και αυτο-ομοιότητα

Ποια είναι η διαμερισματοποίηση του γονιδιώματος?



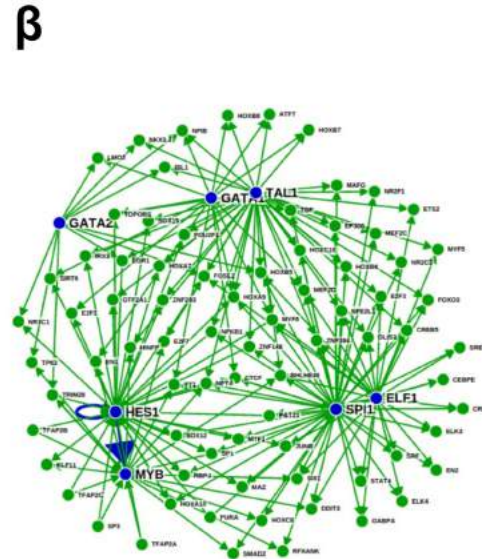
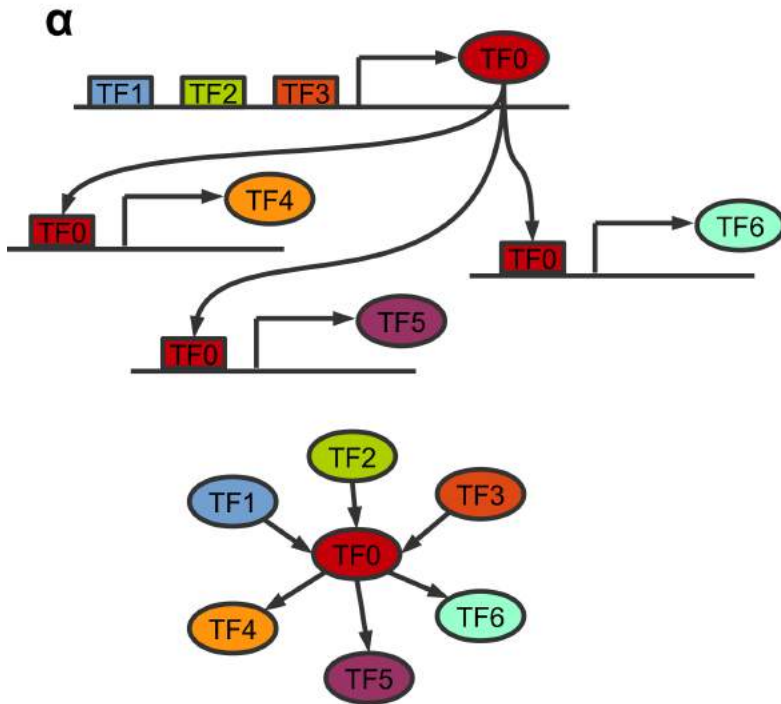
Δίκτυα



Εικόνα 9.1: Ένα παράδειγμα δικτύου πρωτεϊνικών αλληλεπιδράσεων. Οι πρωτεΐνες αναπαρίστανται ως κόμβοι και οι μεταξύ τους σχέσεις με ακμές, που αντιστοιχούν σε σχέσεις ενεργοποίησης. Το χρώμα των κόμβων είναι ανάλογο των αριθμητικών τιμών χαρακτηριστικών ιδιοτήτων του δικτύου που θα συζητηθούν στη συνέχεια του κεφαλαίου. Η Εικόνα δημιουργήθηκε με τη χρήση του Cytoscape (Shannon et al., 2003).

- Απλός ορισμός δικτύου: Μια θεωρητική κατασκευή που περιγράφει τις **σχέσεις** μεταξύ ομοειδών ή ομάδων ομοειδών στοιχείων αναπαριστώντας τις με σχηματικό τρόπο.
- Απομονωμένα στοιχεία οποιουδήποτε είδους δε συνιστούν δίκτυο αλλά ένα απλό σύνολο.
- Αυτό που ορίζει ένα δίκτυο είναι το πλέγμα των σχέσεων ή των “συνδέσεων” μεταξύ τους. Τα δίκτυα είναι στην ουσία γραφικές αναπαραστάσεις πολύπλοκων συστημάτων που μας βοηθούν να τα κατανοήσουμε καλύτερα και πληρέστερα

Δίκτυα μεταγραφικής ρύθμισης



Εικόνα 11.10: α) Δημιουργία δικτύων μεταγραφικής ρύθμισης μέσω της ανάλυσης ενός πειράματος DNaseI DGF. Στον υποκινητή του γονιδίου του μεταγραφικού παράγοντα TF0 εντοπίζονται σημεία πρόσδεσης για τους TF1, TF2 και TF3 ενώ ο ίδιος ο TF0 προσδένεται στους υποκινητές των TF4, TF5 και TF6. Συνδυασμός των παρατηρήσεων οδηγεί στο δίκτυο των επτά κόμβων που φαίνεται στο κάτω αριστερά μέρος της εικόνας. β) Το εξαγόμενο δίκτυο μεταγραφικής ρύθμισης για μυοβλάστες σκελετικών μυών του ανθρώπου, όπως προέκυψε από την ανάλυση του συγκεκριμένου κυτταρικού τύπου στο πλαίσιο του προγράμματος ENCODE. Εικόνα από το regulatorynetworks.org (Neph, Stergachis, et al., 2012).

Εργασία

- Περιγράψτε σε max ½-1 σελίδα ένα ερώτημα, για την επίλυση του οποίου πρέπει να χρησιμοποιήσετε μια από τις τεχνικές NGS, την εφαρμογή της και πως η ανάλυση αυτής της τεχνικής οδηγεί στην επίλυση του ερωτήματός σας.

