

Λειτουργική γονιδιωματική

6ο εργαστήριο

Λειτουργική γονιδιωματική

- Προσπαθεί να κατανοήσει τις λειτουργίες των βιολογικών μορίων, σε επίπεδο ολόκληρου του γονιδιώματος.
- Γίνονται μετρήσεις για το σύνολο των γονιδίων, σε μια συγκεκριμένη στιγμή ή κατάσταση.
- Αρχικά, οι μετρήσεις γίνονταν για ένα βιομόριο. Σήμερα μελετάμε την συμπεριφορά ολόκληρου του συστήματος.
- Η μελέτη της μεταγραφής του συνόλου των γονιδίων ονομάζεται μεταγραφωματική ή transcriptomics.

Transcriptomics

- Expressed sequence tags (ESTs)
- Serial analysis of gene expression (SAGE)
- Μικροσυστοιχίες (microarrays)
- RNA-seq (whole transcriptome shotgun sequencing)

Expressed sequence tags (ESTs)

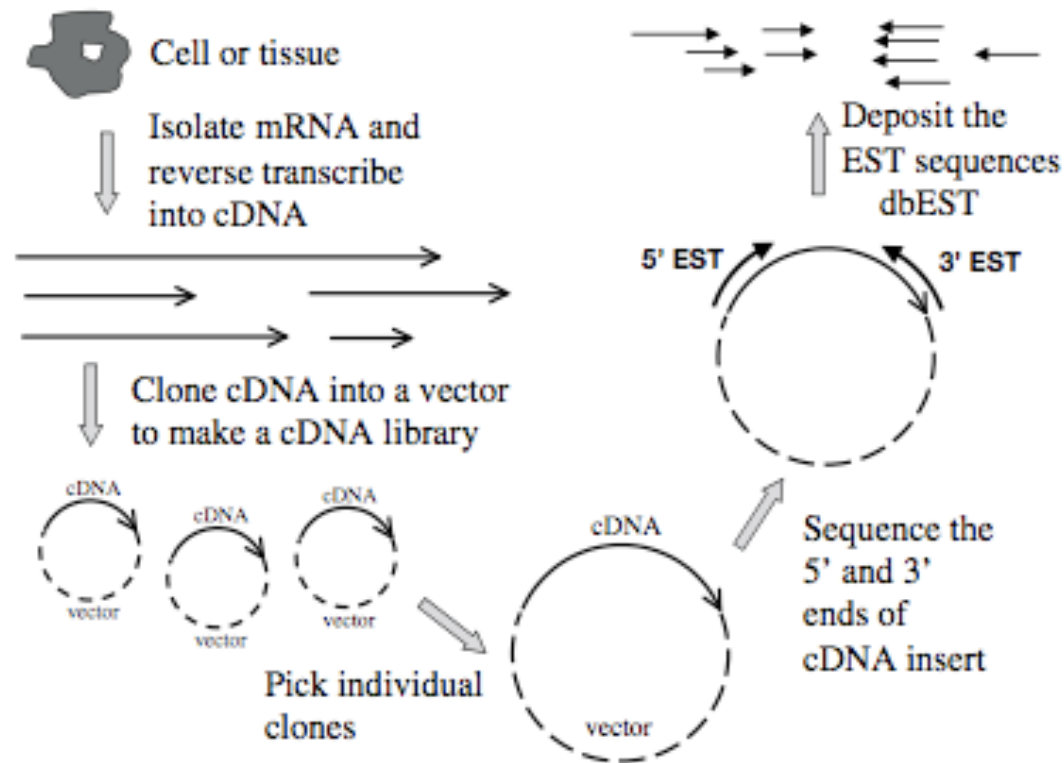


Figure 12.1. Overview of how ESTs are constructed.

Expressed sequence tags (ESTs)

- Το cDNA κλωνοποιείται σε φορείς (vectors) και δημιουργείται μια βιβλιοθήκη.
- Από την βιβλιοθήκη, επιλέγονται τυχαία κλώνοι για αλληλούχιση.
- Η αλληλούχιση ξεκινάει είτε από το 5' ή το 3' άκρο του cDNA.
- Τα παραγώμενα ESTs έχουν μήκος 400-600 νουκλεοτίδια.

Expressed sequence tags (ESTs)

- Η συχνότητα των διαφόρων ESTs που αντιστοιχούν σε ένα γονίδιο αντικατοπτρίζει σε κάποιο βαθμό το επίπεδο της μεταγραφής του γονιδίου.
- Χαμηλά εκφρασμένα γονίδια συνήθως δεν ανιχνεύονται με την μέθοδο των ESTs.
- Κυρίως ανιχνεύονται υψηλά εκφρασμένα γονίδια.
- Οι ακολουθίες των ESTs είναι χαμηλής ποιότητας και μπορεί να εμπεριέχουν λάθη (λάθος νουκλεοτίδια, μετατόπιση αναγνωστικού πλαισίου, λάθος κωδικόνια τερματισμού, χιμαιρικοί κλώνοι).
- Αν και τα επιμέρους ESTs μπορεί να εμπεριέχουν λάθη, μια συλλογή από αλληλο-επικαλυπτόμενα ESTs επιτρέπει την διόρθωση λαθών και ίσως και την ανακατασκευή της ακολουθίας ολόκληρου του cDNA.
- Παρ' όλα τα προβλήματα, τα ESTs χρησιμοποιούνται συχνά, λόγω της ευκολίας που δημιουργούνται οι βιβλιοθήκες σε διάφορες συνθήκες.
- Επιτρέπουν την ανίχνευση νέων γονιδίων.

Expressed sequence tags (ESTs)

- Κατάθεση των δεδομένων στην ΒΔ:
 - NCBI: dbEST (1992)
 - www.ncbi.nlm.nih.gov/dbEST/

Expressed sequence tags (ESTs)

- Οι ακολουθίες:
 - φιλτράρονται.
 - Ομαδοποιούνται και τυχόν λάθη διορθώνονται -> EST contigs.
 - Εντοπίζονται οι περιοχές που κωδικοποιούν την πρωτεΐνη με αλγόριθμους πρόβλεψης γονιδίων.
 - Η προβλεπόμενη πρωτεΐνη επιτρέπει την αναζήτηση ομολόγων σε ΒΔ και την πρόβλεψη της λειτουργίας.
 - Αν είναι διαθέσιμο το γονιδίωμα, τότε τα ESTs βοηθούν στην ανίχνευση γονιδίων και στον καθορισμό των ορίων μεταξύ ιντρονίων-εξονίων.

Expressed sequence tags (ESTs)

- Unigene
- www.ncbi.nlm.nih.gov/UniGene/
- Κάθε ομάδα ESTs είναι αλληλοεπικαλυπτόμενα ESTs που αντιπροσωπεύουν ένα συγκεκριμένο γονίδιο

Unigene

The screenshot shows a web browser window with the address bar containing the URL: <http://www.ncbi.nlm.nih.gov/sites/entrez?db=unigene&cmd=search&term=ESR1+human>. The page title is "uid=163517 UniGene Result".

The NCBI logo is visible on the left, and the UniGene logo is in the center with the tagline "ORGANIZED VIEW OF THE TRANSCRIPTOME". A "My NCBI" button with "Sign In" and "Register" links is on the right.

The navigation bar includes: All Databases, PubMed, Nucleotide, Protein, Genome, Structure, OMIM, PMC, Journals, Books.

The search bar shows "UniGene" selected and "ESR1 human" entered. Buttons for "Go", "Clear", and "Save Search" are present.

Below the search bar are buttons for "Limits", "Preview/Index", "History", "Clipboard", and "Details".

The display options show "Summary" selected, "Show 20", and "Sort By" set to a default. A "Send to" button is also visible.

Taxonomy filters are shown: All: 1, Fungi: 0, Insects: 0, Mammals: 1, Plants: 0.

The search results list one entry:

- 1:** [Estrogen receptor 1](#)
ESR1, *Homo sapiens*
Hs.208124: 217 sequences.
[Order cDNA clone](#)

A "Recent activity" sidebar on the right shows a list of previous searches:

- ESR1 human (1)
- 9606[taxid] AND adult[res...] (2280)
- Estrogen receptor 1
- ESR1 (31)
- ESR1_HUMAN (0)

At the bottom of the sidebar, there is a link: » See more...

Unigene

Estrogen receptor 1 (ESR1)

SELECTED PROTEIN SIMILARITIES

Comparison of sequences in UniGene with selected protein reference sequences. The alignments can suggest function of a gene.

	Reference Protein	Species	Id(%)	Len(aa)
NP_001116212.1	estrogen receptor isoform 2	<i>H. sapiens</i>	100.0	594
XP_001097228.1	PREDICTED: estrogen receptor isoform 1	<i>M. mulatta</i>	99.7	594
NP_001158059.1	estrogen receptor	<i>P. anubis</i>	99.7	594
NP_031982.1	estrogen receptor	<i>M. musculus</i>	94.8	598
NP_001083084.2	estrogen receptor 1	<i>X. laevis</i>	82.4	584
NP_694491.1	estrogen receptor	<i>D. rerio</i>	69.9	531

GENE EXPRESSION

Tissues and development stages from this gene's sequences survey gene expression. Links to other NCBI expression resources.

Restricted Expression: adult [[show more like this](#)]

[EST Profile:](#) Approximate expression patterns inferred from EST sources.
[[Show more entries with profiles like this](#)]

[GEO profiles:](#) Experimental gene expression data (Gene Expression Omnibus).

[cDNA Sources:](#) uterus; mammary gland; ovary; uncharacterized tissue; prostate; pancreas; muscle; kidney; testis; lung; brain; spleen; trachea; thymus; heart; pituitary gland; eye; connective tissue; adrenal gland; embryonic tissue; lymph node

MAPPING POSITION

Genomic location specified by transcript mapping, radiation hybrid mapping, genetic mapping or cytogenetic mapping.

Chromosome: 6
Map position: 6q25.1
UniSTS entry: Chr 6 [PMC108984P2](#)
UniSTS entry: Chr 6 [RH103609](#)
UniSTS entry: [BARC0078](#)
UniSTS entry: [Esr1](#)
UniSTS entry: [Esr1](#)

Unigene

SEQUENCES

Sequences representing this gene; mRNAs, ESTs, and gene predictions supported by transcribed sequences.

mRNA sequences (78)

BX640939.1	Homo sapiens mRNA; cDNA DKFZp686N23123 (from clone DKFZp686N23123)	PA
NM_000125.3	Homo sapiens estrogen receptor 1 (ESR1), transcript variant 1, mRNA	PA
M12674.1	Human estrogen receptor mRNA, complete cds	P
AF120105.1	Homo sapiens alternatively-spliced estrogen receptor alpha mRNA, partial cds	P
AY750962.1	Homo sapiens estrogen receptor alpha mamillary body 1 isoform mRNA, partial cds, alternatively spliced	P
AF258449.1	Homo sapiens estrogen receptor alpha mRNA, complete cds, alternatively spliced	P
AF258450.1	Homo sapiens estrogen receptor alpha mRNA, complete cds, alternatively spliced	P
AF258451.1	Homo sapiens estrogen receptor alpha mRNA, complete cds, alternatively spliced	P
DQ163909.1	Homo sapiens hippocampal estrogen receptor alpha isoform TADDI (ESR1) mRNA, partial cds, alternatively spliced	
BC128574.1	Homo sapiens estrogen receptor 1, mRNA (cDNA clone MGC:157709 IMAGE:40128595), complete cds	P

EST sequences (10 of 140) [[Show all sequences](#)]

AI025006.1	Clone IMAGE:1631712	mammary gland	3' read	P
AI073549.1	Clone IMAGE:1640294	testis	3' read	PA
AI127412.1	Clone IMAGE:1705898	heart	3' read	P
BX108369.1	Clone IMAGp998C181779_._; IMAGE:725321	ovary		A
AI202659.1	Clone IMAGE:1943657	prostate	3' read	A
AI274727.1	Clone IMAGE:1986505	uterus	3' read	A
AI273871.1	Clone IMAGE:1964251	ovary	3' read	A
AI370308.1	Clone IMAGE:1987501	uterus	3' read	A
CB215772.1	Clone IMAGE:5937582	uterus	5' read	
AI524356.1	Clone IMAGE:2118591	prostate	3' read	A

[Download Sequences](#)

Key to Symbols

- P** Has similarity to known **P**roteins (after translation)
- A** Contains a poly-**A**denylation signal
- S** Sequence is a **S**uboptimal member of this cluster
- M** Clone is putatively CDS-complete by **M**GC criteria

Unigene

EST Profile

Hs.208124 - ESR1: Estrogen receptor 1

Breakdown by Body Sites

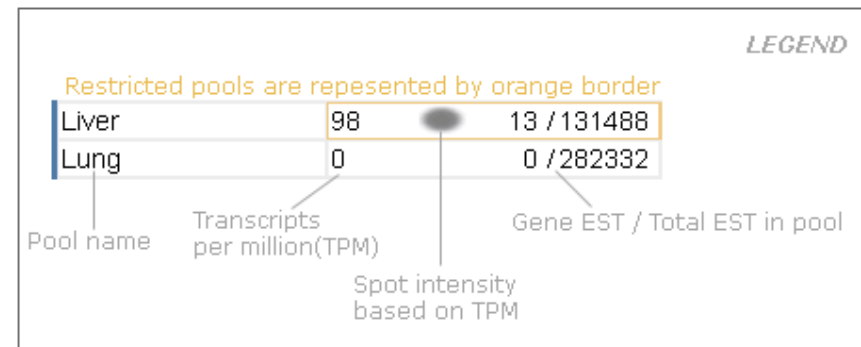
Hs.208124		
adipose tissue	0	0 / 13106
adrenal gland	30	1 / 33197
ascites	0	0 / 40015
bladder	0	0 / 29757
blood	0	0 / 123478
bone	0	0 / 71655
bone marrow	0	0 / 48801
brain	0	1 / 1100989
cervix	0	0 / 48171
connective tissue	6	1 / 149255
ear	0	0 / 16212

Breakdown by Developmental Stage

Hs.208124		
embryoid body	0	0 / 70761
blastocyst	0	0 / 62319
fetus	5	3 / 564012
neonate	0	0 / 31097
infant	0	0 / 23620
juvenile	0	0 / 55556
adult	21	42 / 1939121

Breakdown by Health State

Hs.208124		
adrenal tumor	78	1 / 12794
bladder carcinoma	0	0 / 17475
breast (mammary gland) tumor	63	6 / 94178
cervical tumor	0	0 / 34366
chondrosarcoma	12	1 / 82823
colorectal tumor	0	0 / 114246
esophageal tumor	0	0 / 17290
gastrointestinal tumor	0	0 / 119369
germ cell tumor	0	0 / 263845
glioma	0	0 / 106883



Unigene

- Οι βιβλιοθήκες είναι οργανωμένες στο library browser

[http://www.ncbi.nlm.nih.gov/UniGene/lbrowse2.cgi?TAXID=9606&log\\$=BlueSideBar](http://www.ncbi.nlm.nih.gov/UniGene/lbrowse2.cgi?TAXID=9606&log$=BlueSideBar)

Library Browser: Homo sapiens

Search UniGene

Libraries for with minimum sequences

[Collapse All](#) | [Expand All](#)

Body Sites

▼ **adipose tissue** 18 libraries

Lib. ID	Library Name	Sequences
Lib.10983	Human Fat Cell 5'-Stretch Plus cDNA Library	9638
Lib.886	NCI_CGAP_Lip2	1740
Lib.16445	Sugano cDNA library, adipose tissue	1665
Lib.816	Adipose tissue, white II	1195

Not Shown: 14 libraries having fewer than 1000 sequences

▼ **adrenal gland** 30 libraries

Lib. ID	Library Name	Sequences
Lib.18302	ADRGL2	10385
Lib.7317	NIH_MGC_84	7572
Lib.6791	ADB	6475
Lib.927	NCI_CGAP_AA1	3363
Lib.16377	Sugano cDNA library, adrenal gland	2772
Lib.6815	cdA	2460
Lib.6792	ADC	1995
Lib.6793	Cu	1649
Lib.993	NCI_CGAP_Phe1	1356
Lib.766	Adrenal gland tumor	1183

Not Shown: 20 libraries having fewer than 1000 sequences

▼ **amniotic fluid** 63 libraries

Lib. ID	Library Name	Sequences
Lib.7332	AN0080	1112

Not Shown: 62 libraries having fewer than 1000 sequences

Digital differential display

- <http://www.ncbi.nlm.nih.gov/UniGene/help.cgi?item=DDD>
- Συγκρίνει ομάδες βιβλιοθηκών ESTs μεταξύ τους και βρίσκει ποιά γονίδια είναι περισσότερο εκφρασμένα στην κάθε μια από τις βιβλιοθήκες.
- Η σύγκριση γίνεται με Fisher's exact test.
- Για να γίνουν οι συγκρίσεις, θα πρέπει ο αριθμός των ESTs που αντιστοιχούν σε ένα γονίδιο, για την κάθε βιβλιοθήκη, να ξεπερνάει ένα κατώφλι, αλλιώς δεν μπορούν να φανούν οι διαφορές.
- Οι συγκρίσεις μπορούν να αποκαλύψουν γονίδια που παίζουν σημαντικό ρόλο σε κάποιο ιστό ή περιβαλλοντική κατάσταση ή ασθένεια.

Digital differential display

Digital Differential Display (DDD)

DDD is a tool for comparing EST profiles in order to identify genes with significantly different expression levels ([More about DDD](#)).

Species: *Homo sapiens* (human)

[Start Over](#)

Pool A: Muscle

2 libraries, 5084 ESTs

[Edit Pool](#)

Pool B: Skin

4 libraries, 35274 ESTs

[Edit Pool](#)

[New Pool](#)

Differential Display Results

The following genes (UniGene entries) display statistically significant differences in EST counts by the Fisher Exact Test.

A Muscle	B Skin	UniGene Entry
0.0519 	0.0000	Hs.728212 Transcribed locus, strongly similar to NP_001091.1 actin, alpha skeletal muscle [Homo sapiens]
0.0228 	0.0001	Hs.134602 Titin (TTN)
0.0142 	0.0000	Hs.726317 Transcribed locus, strongly similar to NP_036662.1 creatine kinase M-type [Rattus norvegicus]
0.0134 	0.0000	Hs.631558 Troponin T type 1 (skeletal, slow) (TNNT1)
0.0124 	0.0000	Hs.320890 Troponin I type 1 (skeletal, slow) (TNNI1)
0.0104 	0.0000	Hs.719946 Myosin, heavy chain 7, cardiac muscle, beta (MYH7)
0.0094 	0.0000	Hs.517586 Myoglobin (MB)
0.0092 	0.0001	Hs.598320 Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)

Άσκηση

- Πηγαίνετε στον ιστότοπο του DDD και διαβάστε πρώτα την αρχική σελίδα, που εξηγεί πώς δουλεύει το DDD.

<http://www.ncbi.nlm.nih.gov/UniGene/help.cgi?item=DDD>

- Χρησιμοποιείστε το DDD για να βρείτε πιθανούς μοριακούς δείκτες για τον ανθρώπινο καρκίνο του ήπατος.
- Μπορείτε να συγκρίνετε τις ακόλουθες βιβλιοθήκες με dbEST ids 18416 & 18525.
- Πριν κάνετε την σύγκριση, δείτε τι πληροφορίες υπάρχουν για την κάθε βιβλιοθήκη.
- Αφού κάνετε την σύγκριση, ποιοί είναι πιθανοί δείκτες;
- Ελέγξτε την α -fetoprotein
 - Ποιό είναι το EST profile στο Unigene
 - Δείτε τις πληροφορίες στο Uniprot
 - Δείτε τις πληροφορίες στο MIM (Mendelian inheritance in Man)
 - Δείτε τις σχετικές πληροφορίες στο KEGG και KEGG diseases.
 - Θεωρείται μοριακός δείκτης για τον καρκίνο του ήπατος;
 - Ελέγξτε τι υπάρχει για την πρωτεΐνη αυτή στη ΒΔ Human Protein Atlas
<http://www.proteinatlas.org/>