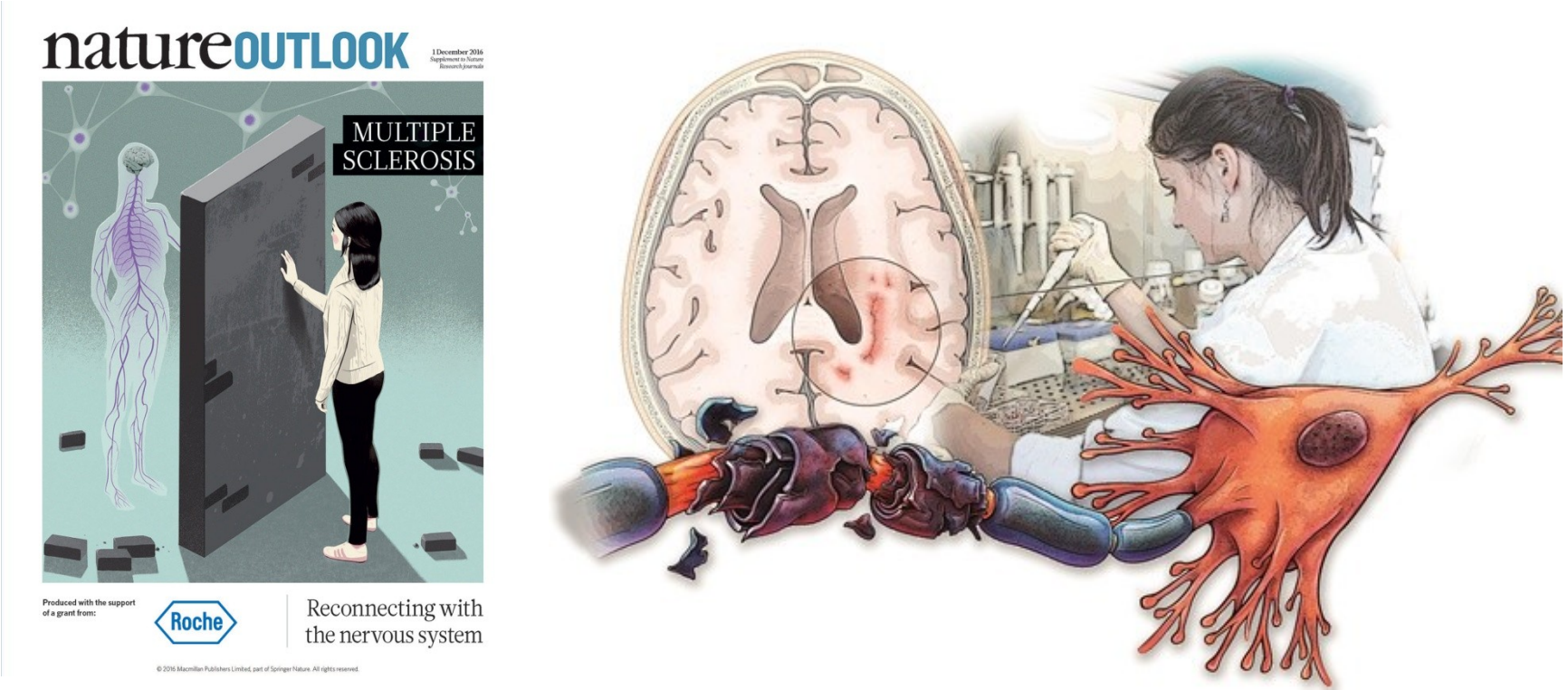
Advanced Molecular Biology

Modulo B - "**APPLICATIONS IN MEDICINE**"

SANTINA CUTRUPI
Professor of Molecular Biology
University of Turin
Dept. Clinical & Biological Sciences
CIR Molecular Systems Biology
santina.cutrupi@unito.it



" APPLICATIONS IN MEDICINE MODULE"

DEVELOPS SKILLS IN:

- **PROBLEM SOLVING in the application of molecular biology IN MEDICINE**

- **EXPERIMENTAL DESIGN to understand molecular mechanisms linked to disease**

" APPLICATIONS IN MEDICINE MODULE"

How can we improve these skills?

- PROBLEM SOLVING IN MOLECULAR BIOLOGY FIELD
- EXPERIMENTAL DESIGN

- PAPERS ANALYSIS
  - METHODS
  - EXPERIMENT STEPS
  - MEANING OF DATA

- Design experiments and molecular biology methods

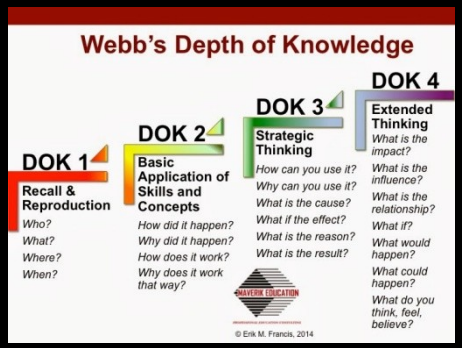
COURSE STRUCTURE:

- **Lesson: presentation/ discussion of main concepts**
- **Quiz: problem solving, experiment design**

We 'll use a "TRAINING TASK"

What is the meaning of this approach?

- Help you to understand the main concept in the deep way
- Help you to remember the main concept
- Help you to apply the main concept to solve problem



In this lesson

- What is the main focus of the course
- Definition of Functional Genomics
- Refresh molecular biology pre-requisite: DNA genomic elements as cell-type specific regulatory regions
- SNPs meaning in the disease
- How Functional Genomics is the basis for understanding diseases

The main focus of this course is
functional genomics

APPLICATIONS IN MEDICINE

Task 1: Search the definition of

FUNCTIONAL GENOMICS

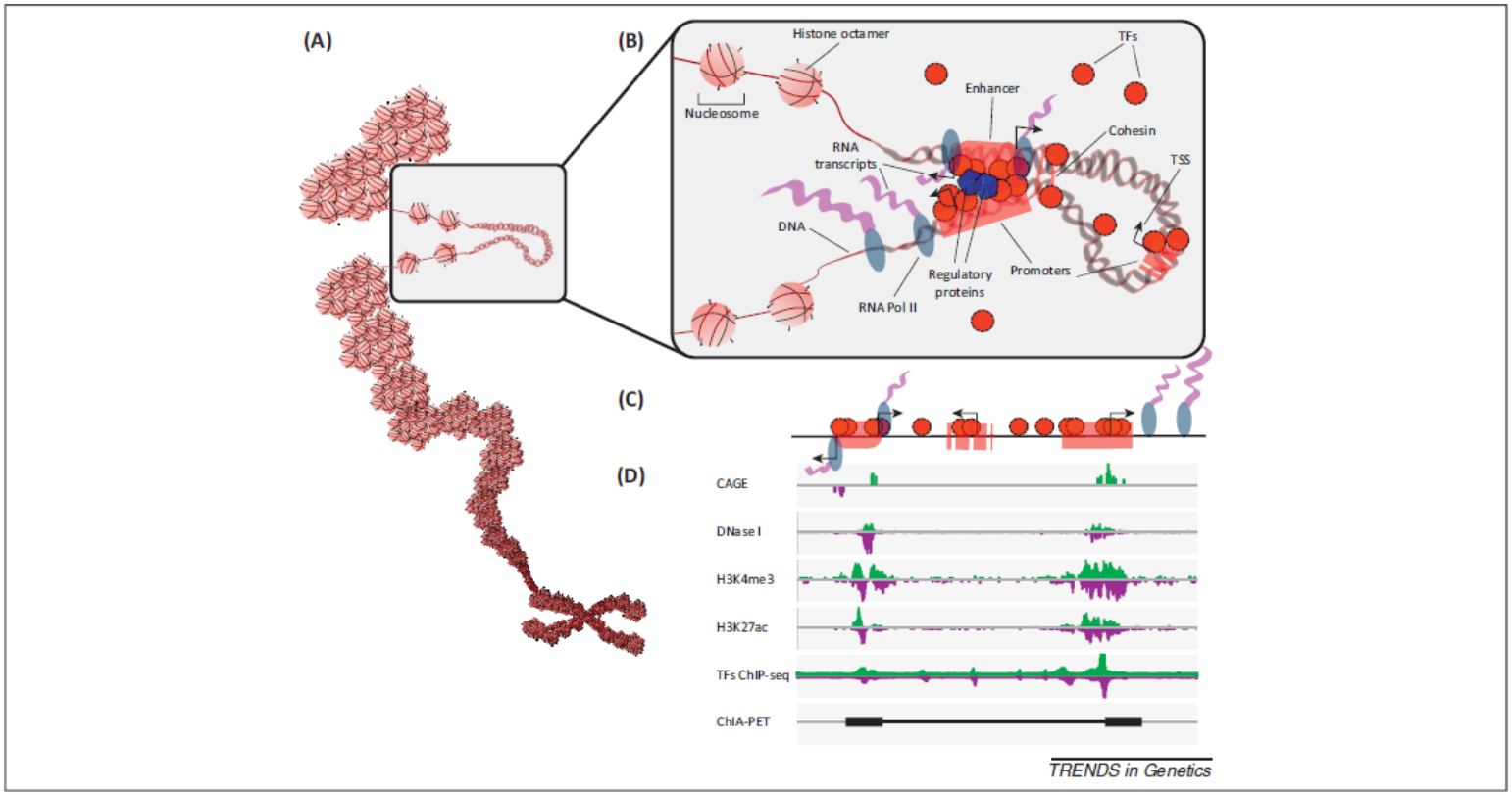
Copy the definition

Answer the questions:

1) **What type of data are used?**
2) **What type of tecniques are used?**
3) **What is the impact?**

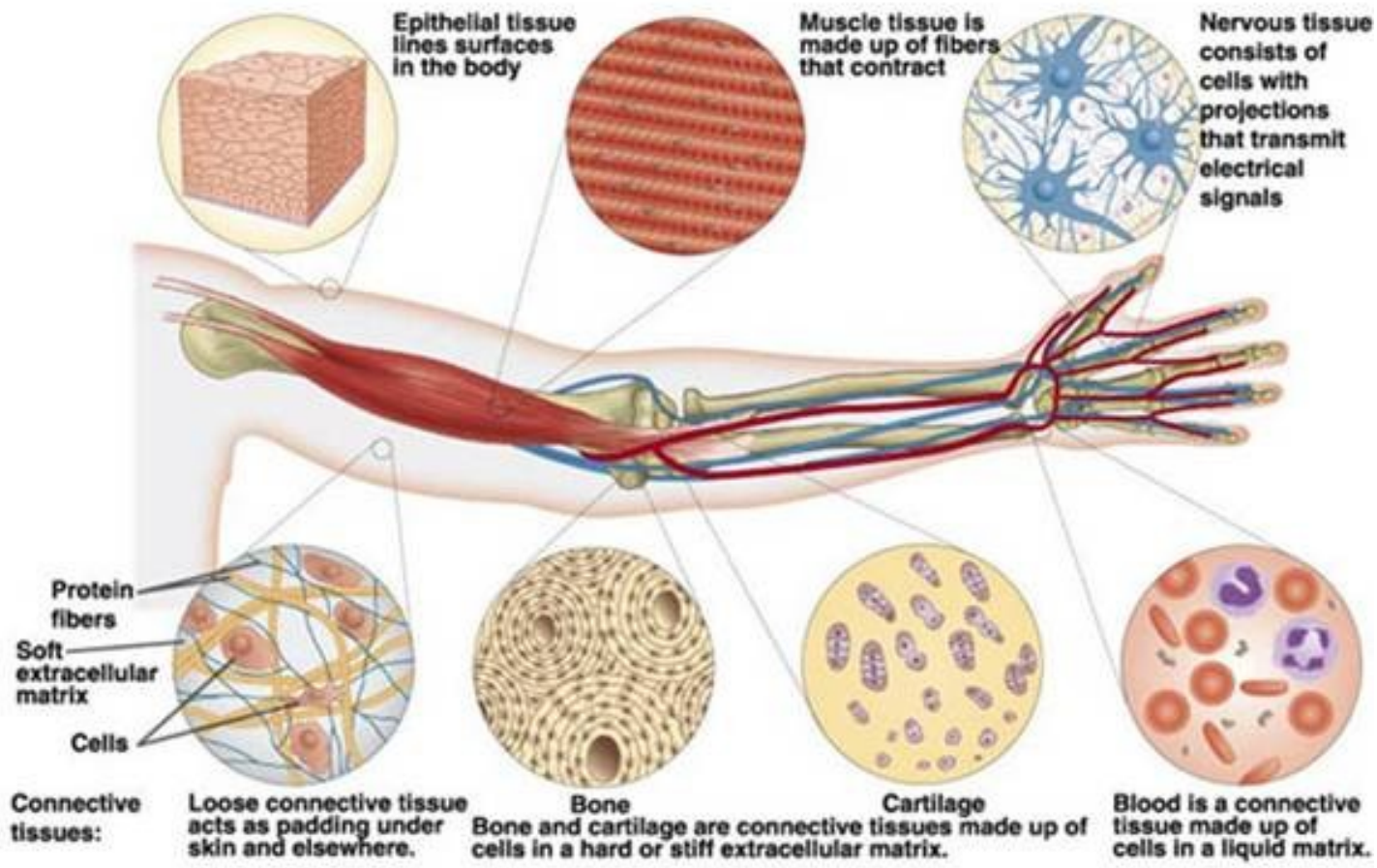
Functional genomics

Functional genomics uses genomic data to study gene expression, regulation and biological functions on a global scale (genome-wide or system-wide), focusing on gene transcription, epigenetic modifications, chromatin remodelling enzymes, transcription factors association involving high-throughput methods.



Task 2- Cell-type specific gene expression

Each cell has a specific pattern of genes. Why?
Did you find in the previous course or lessons the elements that are involved in cell-type specific gene expression?
Write your answer on Moodle "Training Task".



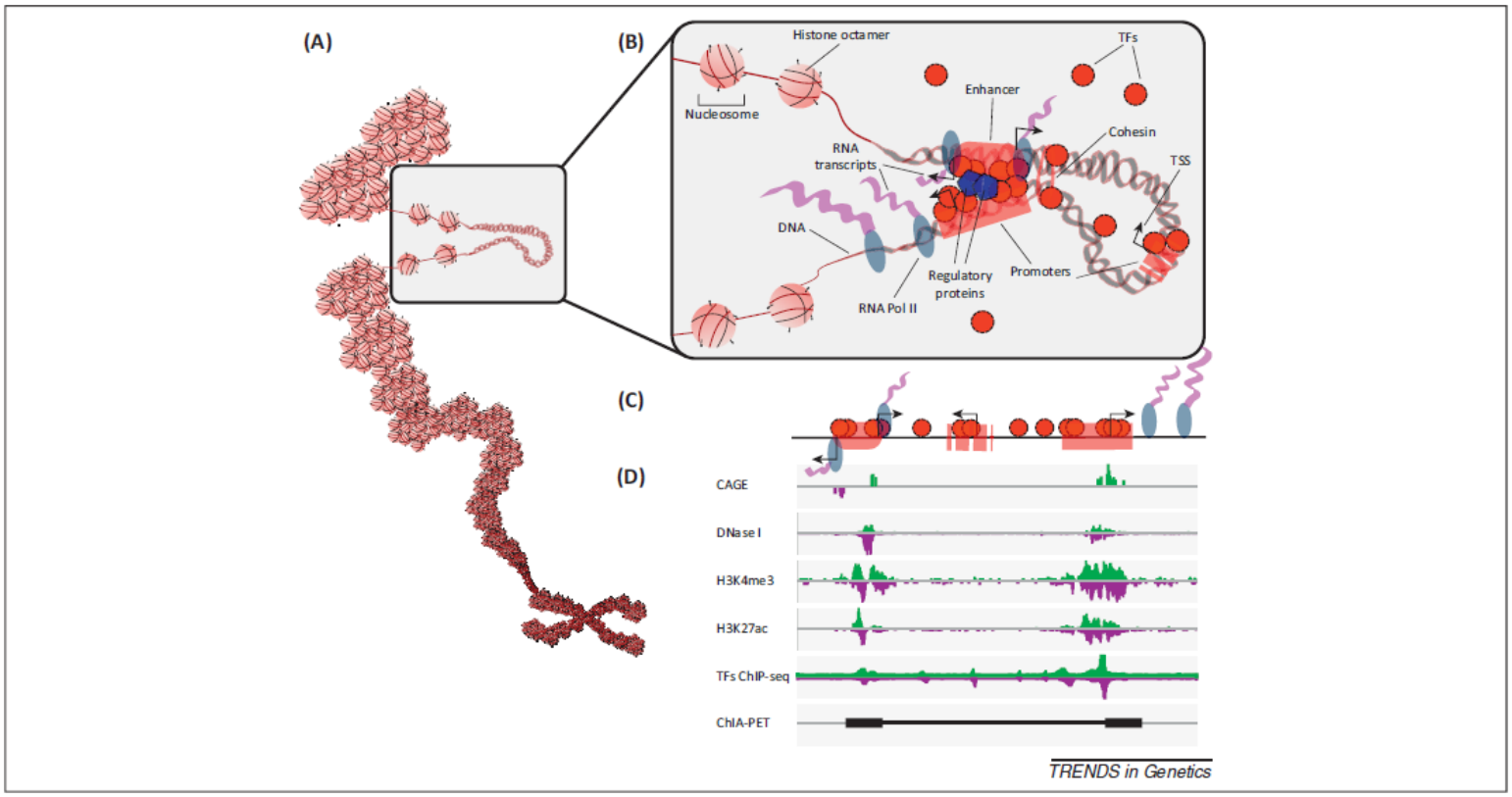
Task 3- What is a genomic regulatory region

Describe the picture:

- What is the meaning of figure B

- Describe the techniques that is used in fig D. Indicate if you don't know some of them.

- Does the figure B link with the data of figure D and C?



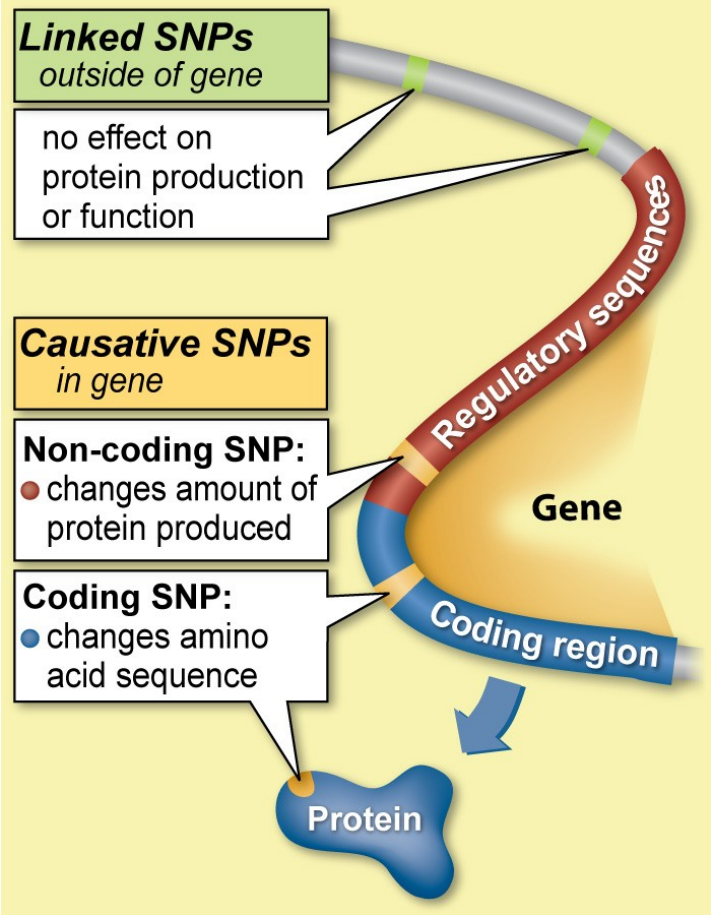
GENOMIC REGULATORY REGIONS

are defined by:

- → **EPIGENETIC MARKS**
- → **TRANSCRIPTION FACTORS BINDING**
- → **CHROMATIN REMODELLING ENZYMES**
- → **NUCLEOSOME POSITIONING**

Task 4- What is the impact of single nucleotide variants

Describe how the change in the single nucleotide in the DNA sequence has an impact in the biological functions.



You have **30 min** to complete this quiz on the Moodle Platform.

Search in the web or in the previous material that you have use during the course

You can write on your book and send me a photo-pdf by Moodle.

- What is the main focus of the course
- Definition of Functional Genomics
- **Focus: DNA genomic elements as cell-type specific regulatory regions**
- How Functional Genomics is the basis for undestanding diseases
- Genome-wide sequencing methods to annotate DNA genomic elements. Storing in Databases.

GENE EXPRESSION REGULATION

BY

ENHACERS

PROMOTER

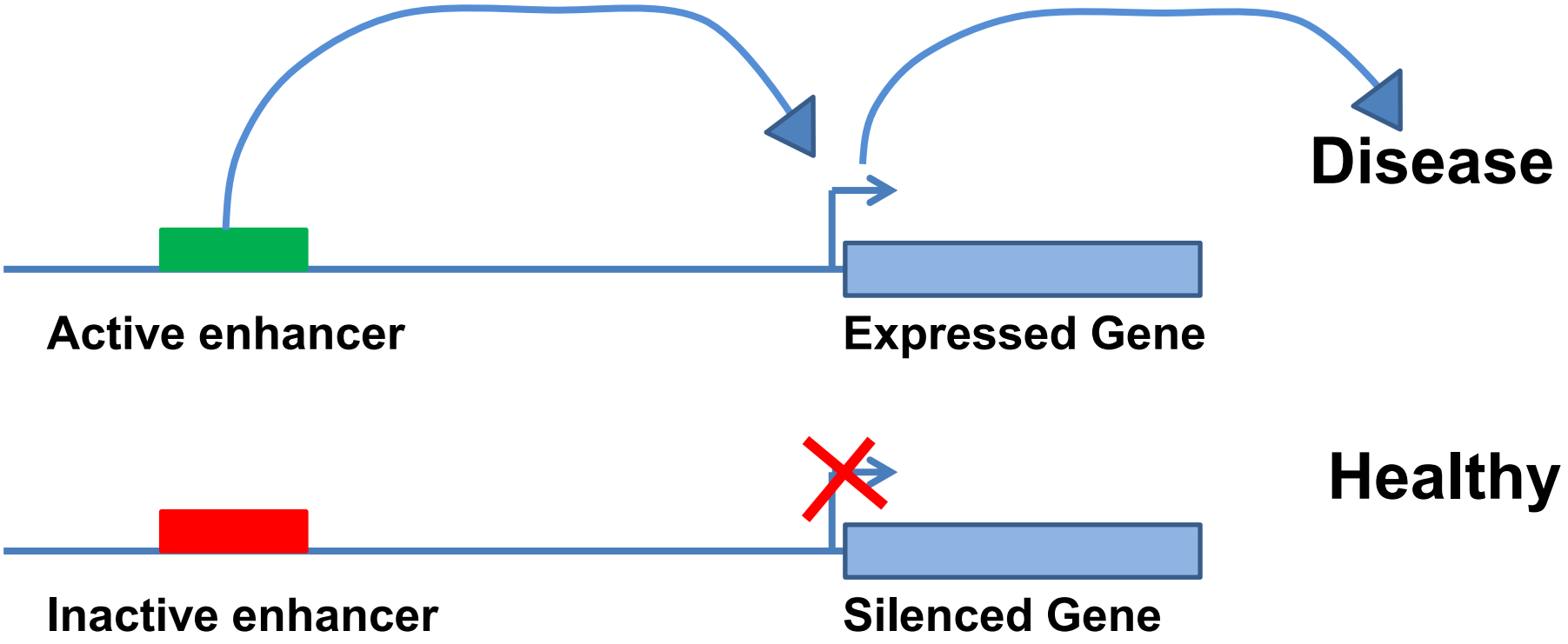
GENOMIC REGULATORY REGIONS

are defined by:

- → EPIGENETIC MARKS
- → TRANSCRIPTION FACTORS BINDING
- → CHROMATIN REMODELLING ENZYMES
- → NUCLEOSOME POSITIONING

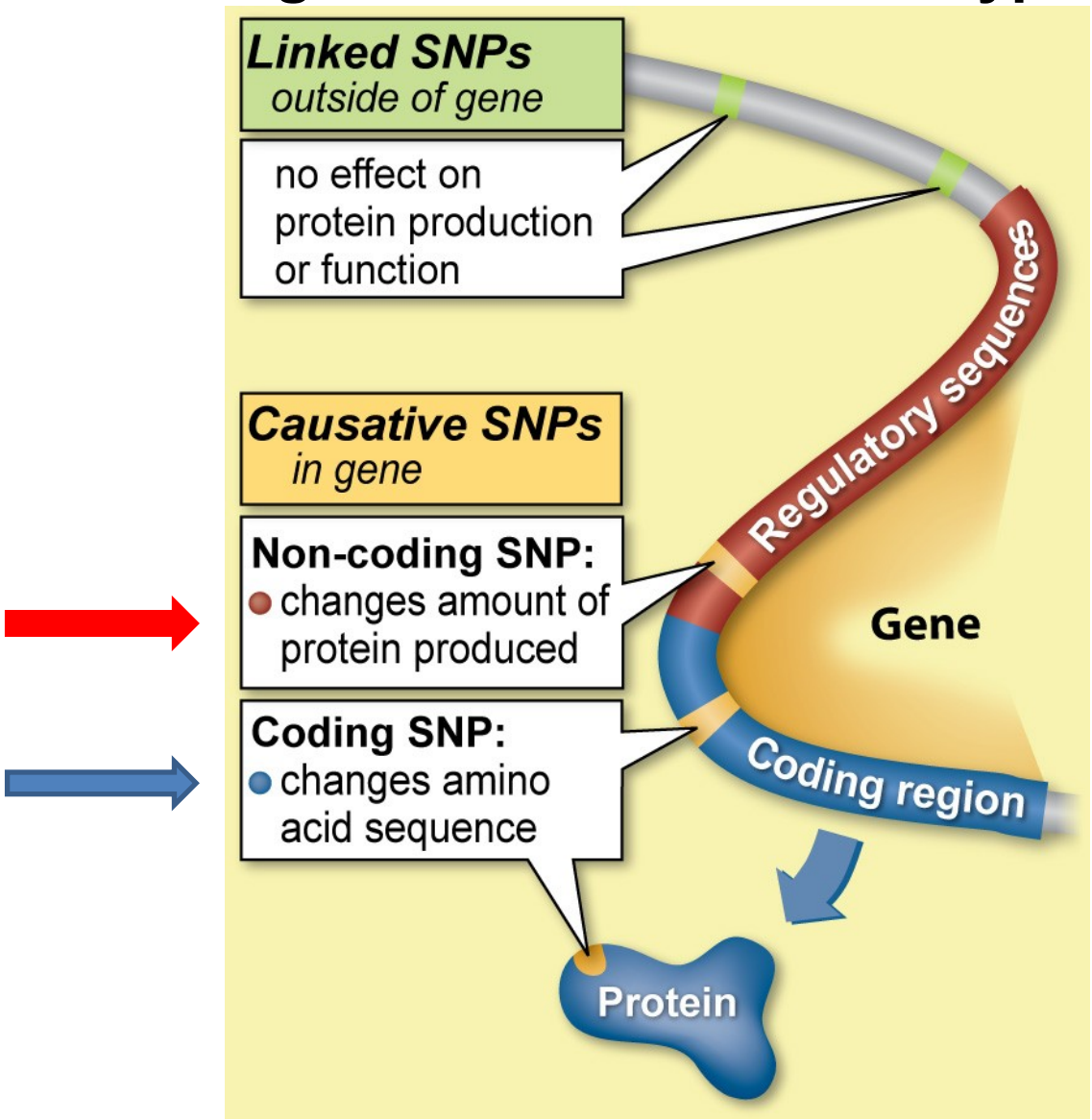
FROM GENOMIC REGULATORY REGIONS TO MOLECULAR MECHANISMS

Genomic regulatory regions control gene expression and specific activation may be associated with disease: One possible Scenario



FROM GENOMIC REGULATORY REGIONS TO MOLECULAR MECHANIMS

Single nucleotide variants types



FROM GENOMIC REGULATORY REGIONS TO MOLECULAR MECHANISMS

What types of alterations in the molecular mechanism could induce diseases:

- **Single nucleotide variations into the genomic regulatory regions change the consensus sequences for transcription factors binding**

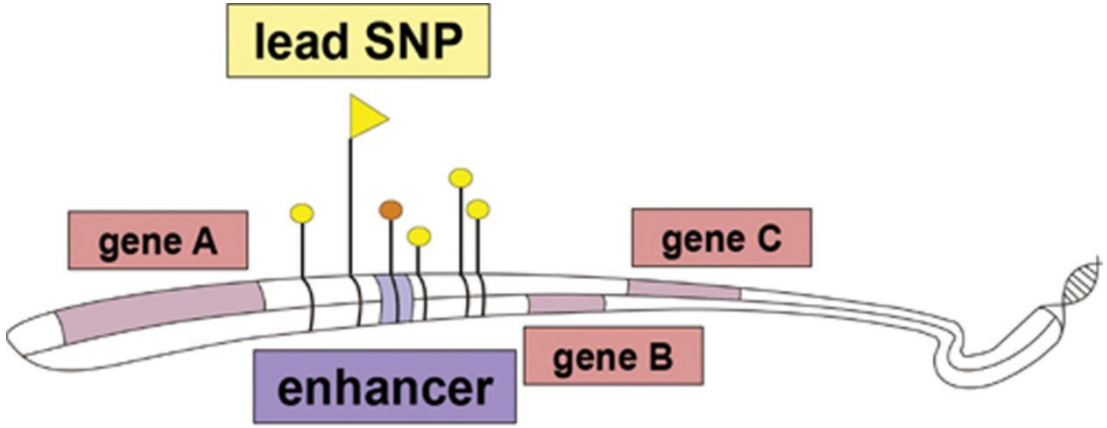
- **Single nucleotide variations into the genomic regulatory regions change long range interactions between two regulatory regions**

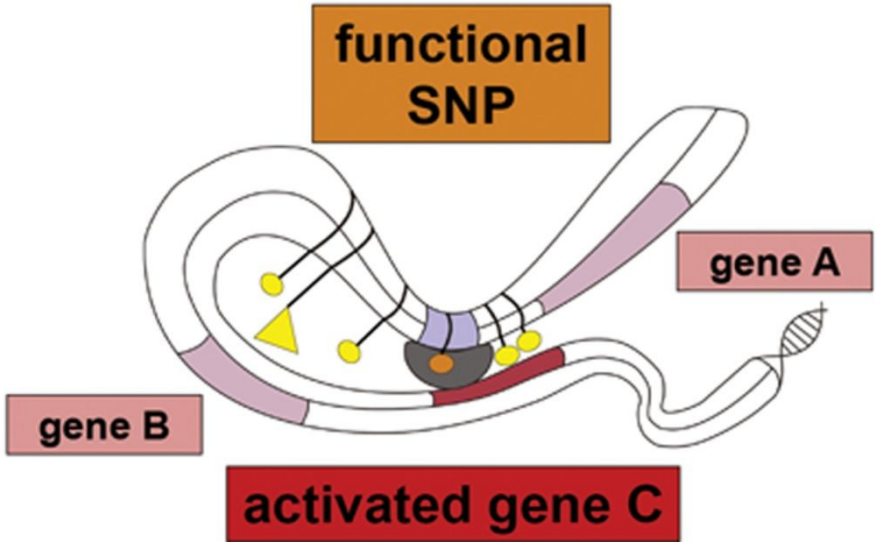
- **Single nucleotide variations in the coding sequence of proteins change:**

a) **Enzimatic activity**
b) **Protein-protein interactions**
c) **Cofactors binding**

FROM GENOMIC REGULATORY REGIONS TO MOLECULAR MECHANIMS

Single nucleotide variants in genomic regulatory regions

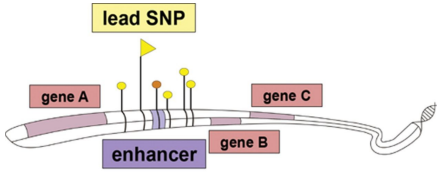




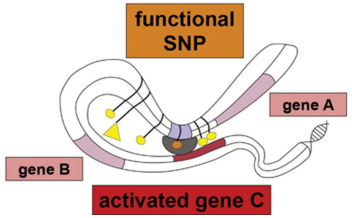
IDENTIFICATION AND CHARCTERIZATION

MOLECULAR MECHANISMS

GENOMIC REGULATORY REGIONS



GENE EXPRESSION REGULATION



CELL IDENTITY

BIOLOGICAL FUNCTIONS

TO UNDERSTAND DISEASES

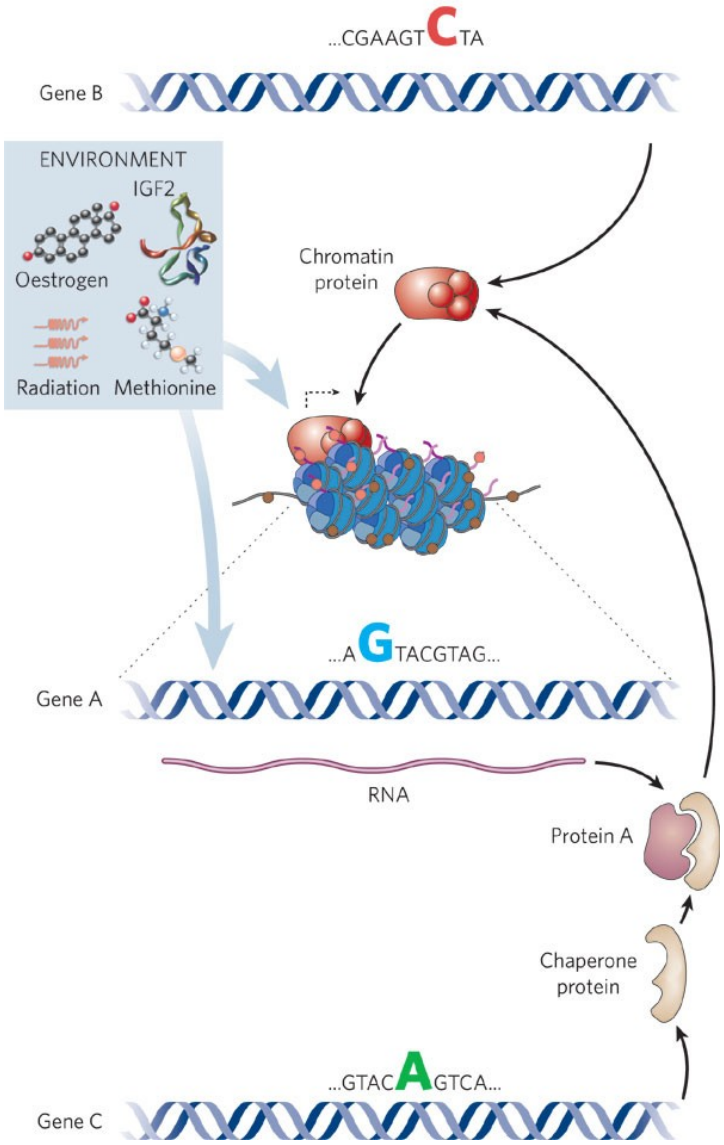
Genome-wide association studies (GWAS) have capitalized on the millions of common single nucleotide polymorphisms (SNPs) to identify those SNPs that are genome-wide significantly associated with a disease or trait.

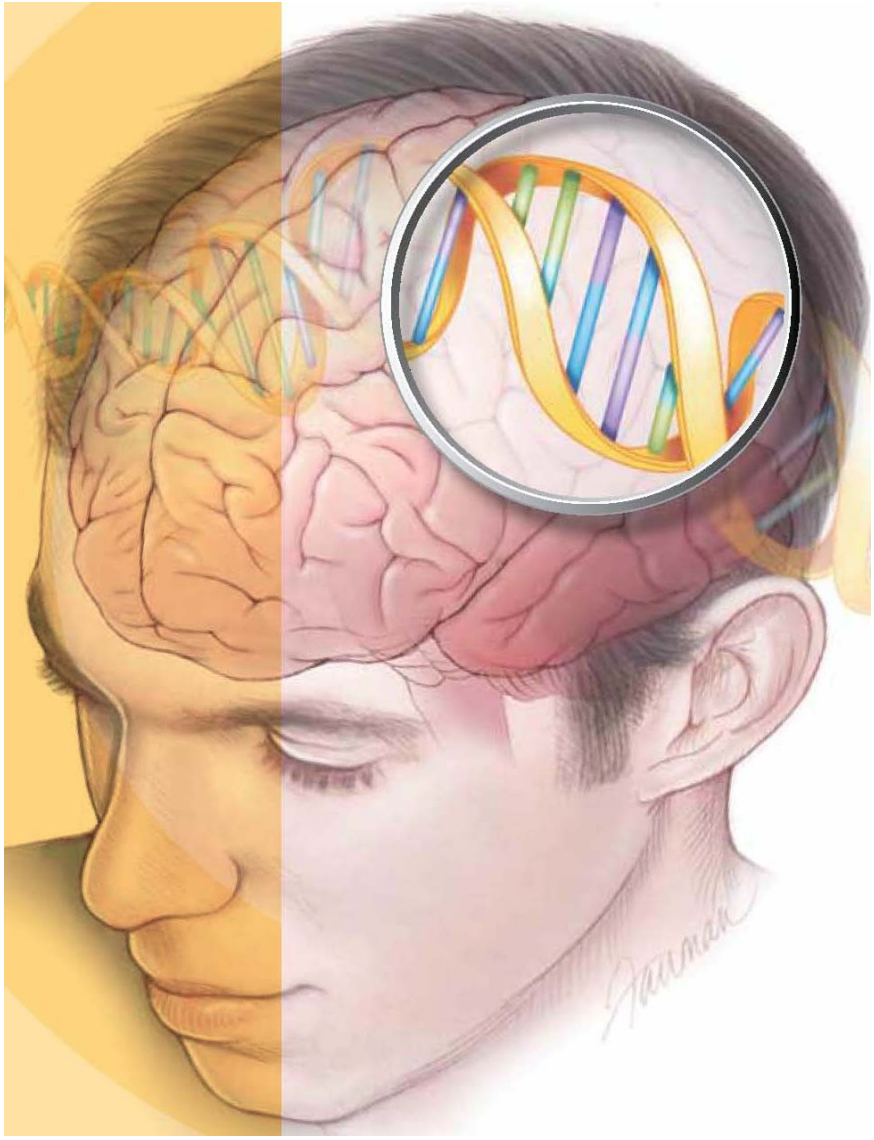
- What is the main focus of the course
- Definition of Functional Genomics
- Focus: DNA genomic elements as cell-type specific regulatory regions
- **How Functional Genomics is the tool to understand diseases**
- Genome-wide sequencing methods to annotate DNA genomic elements. Storing in Databases.

How does functional genomics help us to understand disease?

In this number of Nature Journal there are several articles that show the connection between SNPs and psychiatric disorders

SNPs may have an impact on chromatin remodelling to control neuron activity and network





Which are the steps to understand the SNPs meaning?

Framework for interpretation of individual disease-associated variants

- Single nucleotide polymorphisms (SNPs) is the nucleotide variations associated with disease

- Genome-wide association studies (GWAS) have successfully identified thousands of common genetic variants associated with complex diseases (http://www.ebi.ac.uk/gwas/)

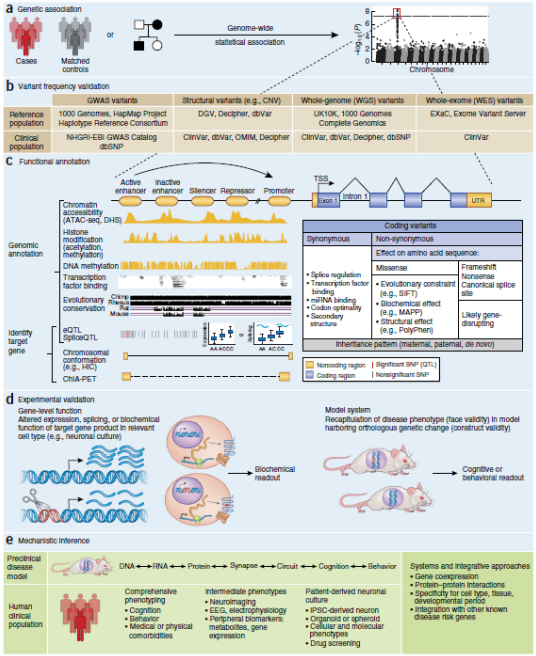
- Functional annotation: to define genomic regulatory regions by genome-wide integration data

- Experimental validation

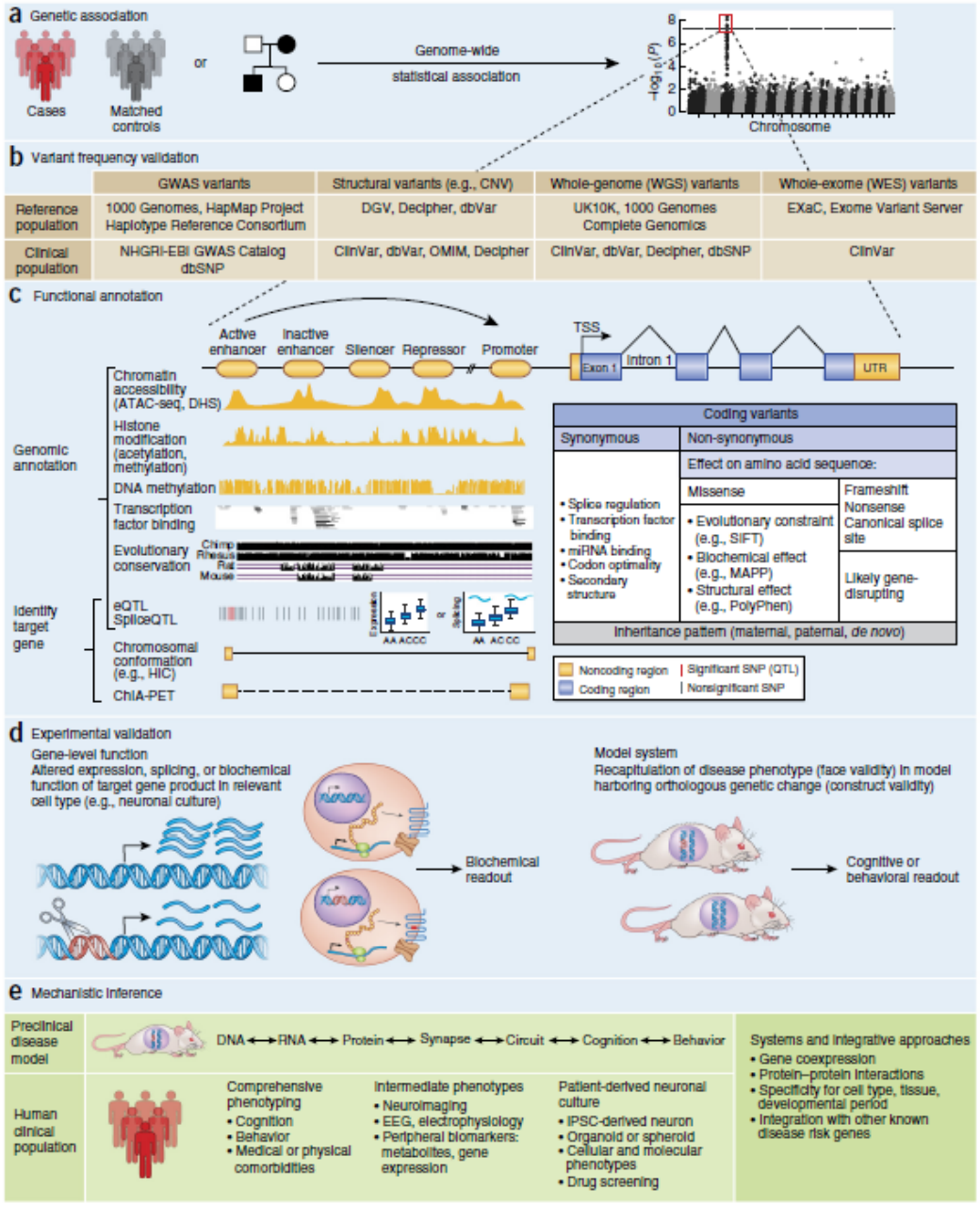
- Disease animal models

- Correlation between molecular mechanisms and disease symptoms

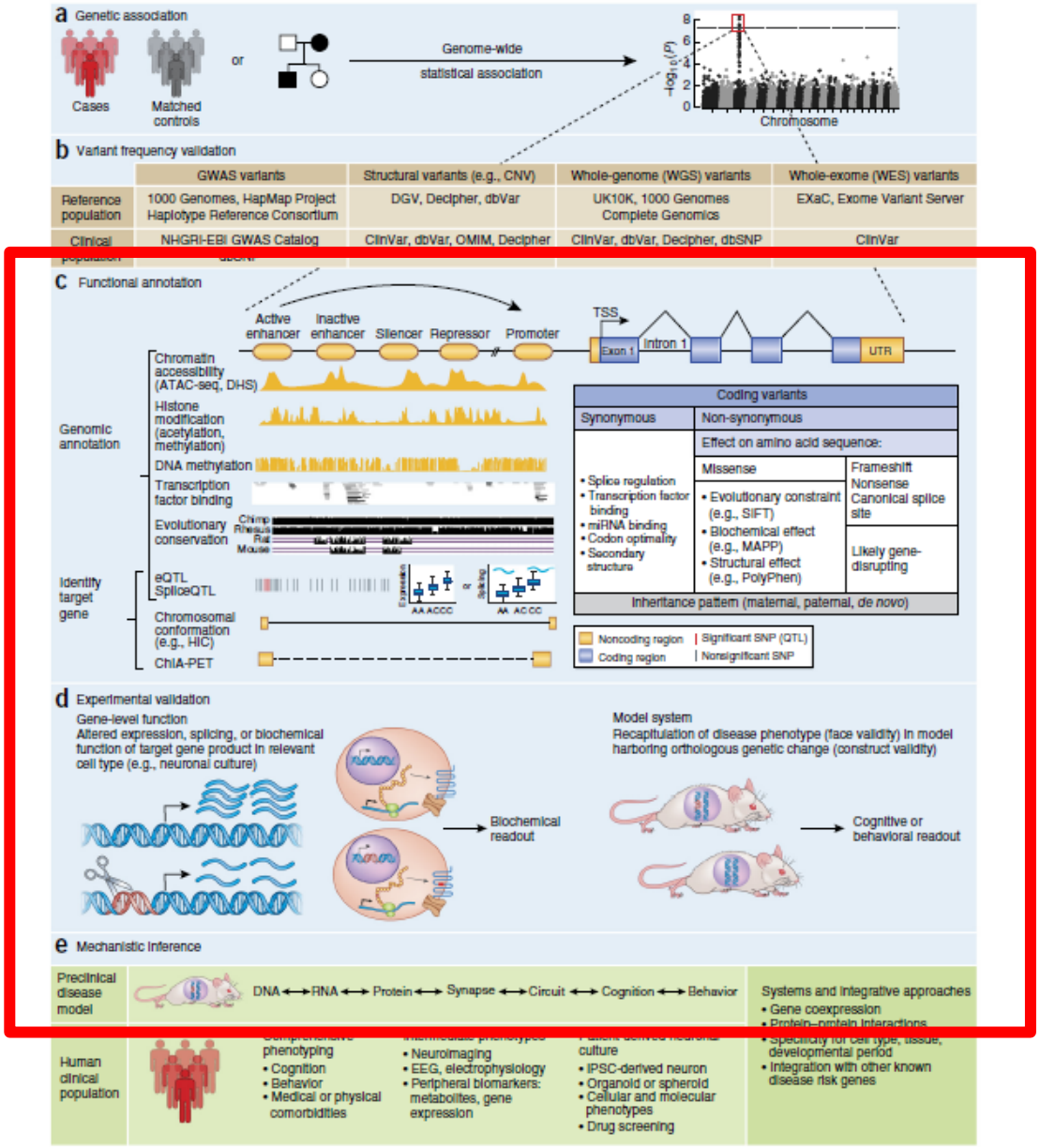
- Drug Discovery



Framework for interpretation of individual disease-associated variants



Point attention: Genomic regulatory regions annotation, Experimental Validation and Animal models



Framework for interpretation of individual disease-associated variants

- Single nucleotide polymorphisms (SNPs) is the nucleotide variations associated with disease

- Genome-wide association studies (GWAS) have successfully identified thousands of common genetic variants associated with complex diseases (http://www.ebi.ac.uk/gwas/)

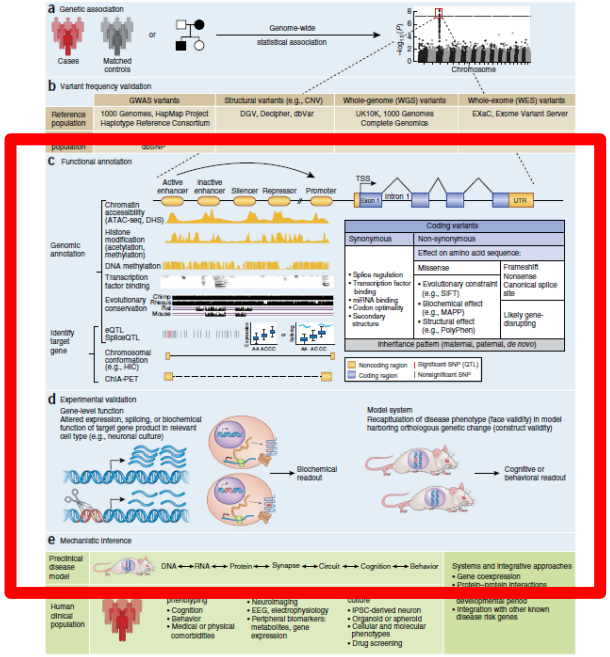
- Functional annotation: to define genomic regulatory regions by genome-wide integration data

- Experimental validation

- Disease Animal models

- Correlation between molecular mechanisms and disease symptoms

- Drug Discovery



Framework for interpretation of individual disease-associated variants

- Single nucleotide polymorphisms (SNPs) is the nucleotide variations associated with disease

- Genome-wide association studies (GWAS) have successfully identified thousands of common genetic variants associated with complex diseases (http://www.ebi.ac.uk/gwas/)

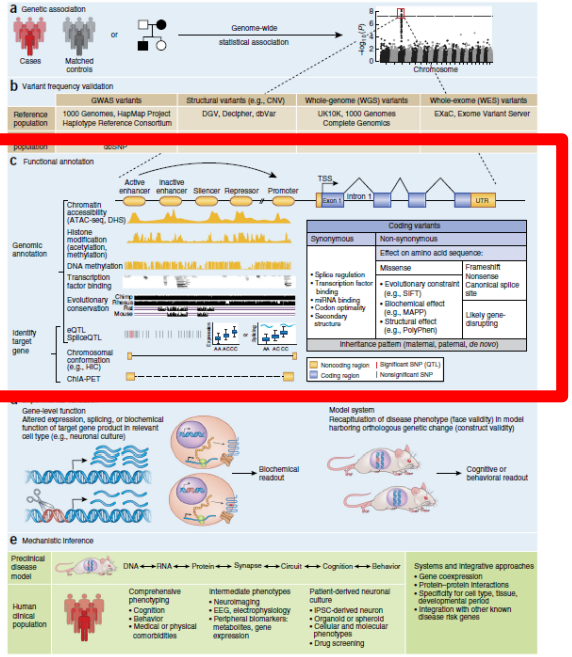
- Functional annotation: to define genomic regulatory regions by genome-wide integration data

- Experimental validation

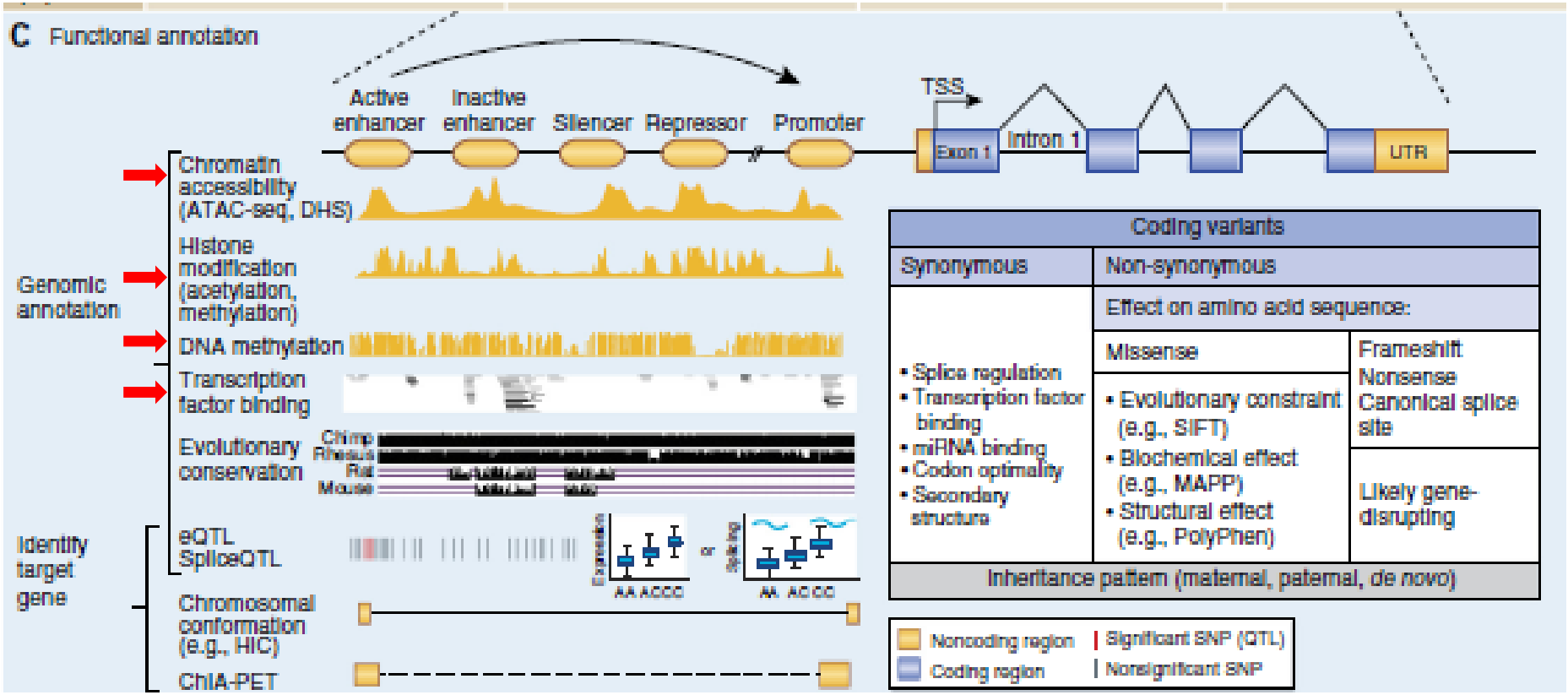
- Disease Animal models

- Correlation between molecular mechanisms and disease symptoms

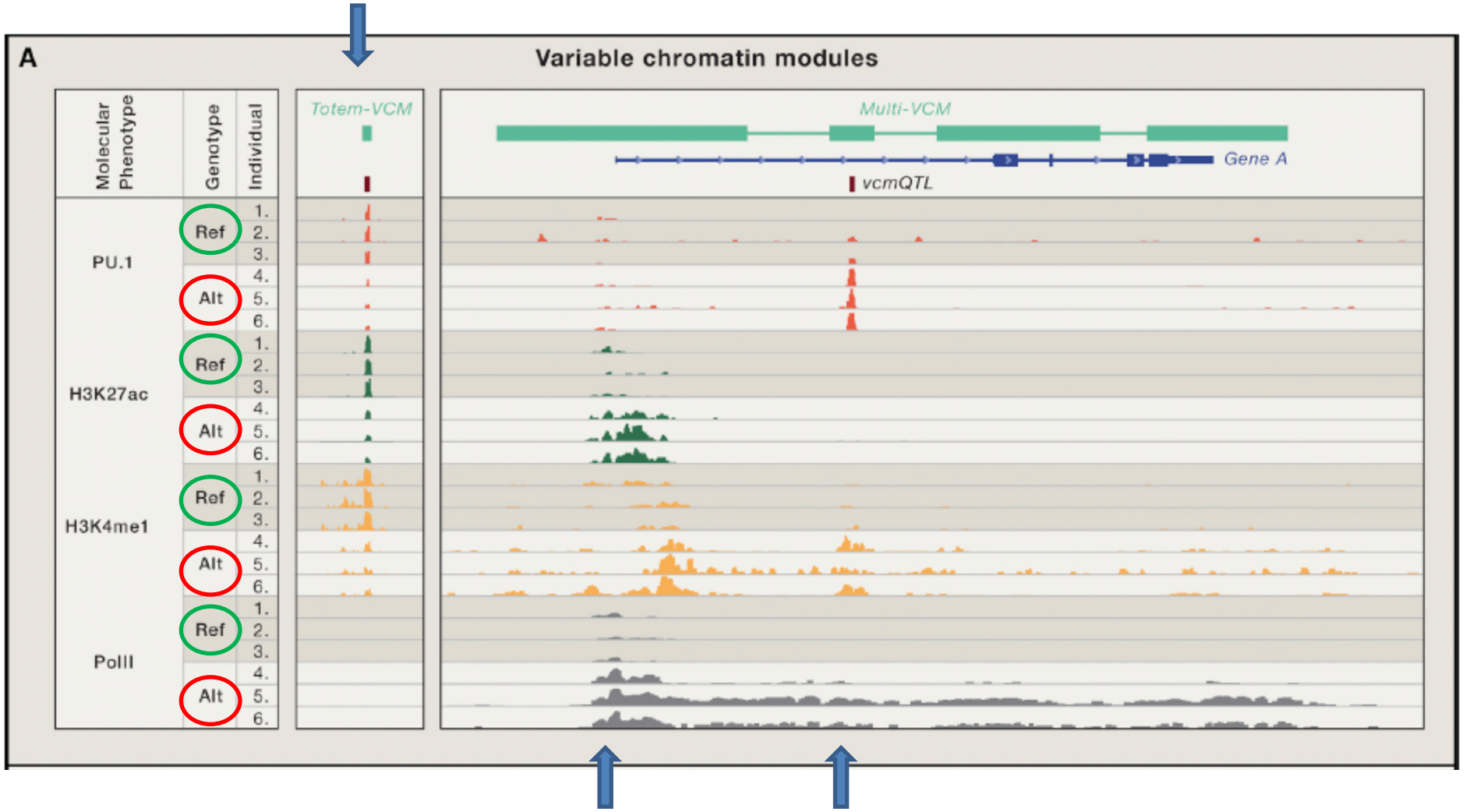
- Drug Discovery



Framework for interpretation of individual disease-associated variants
FUNCTIONAL ANNOTATION

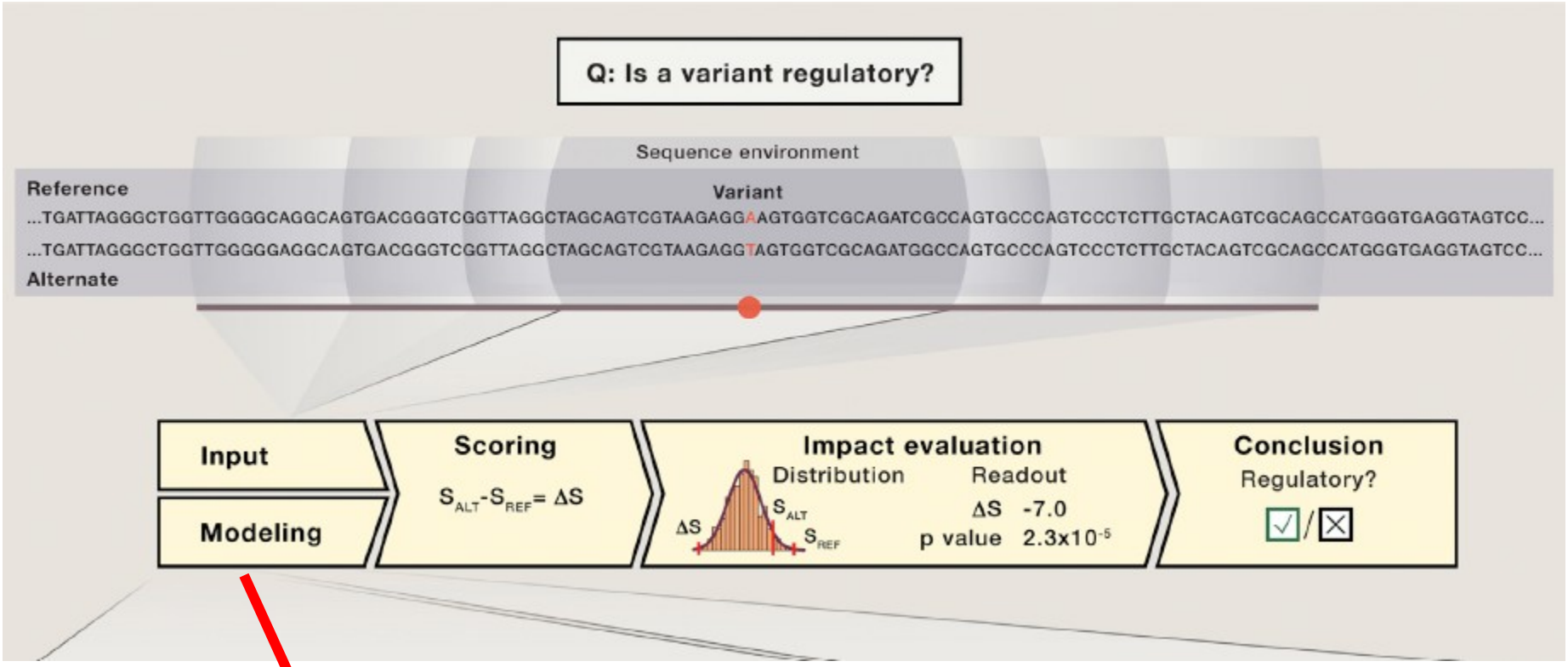


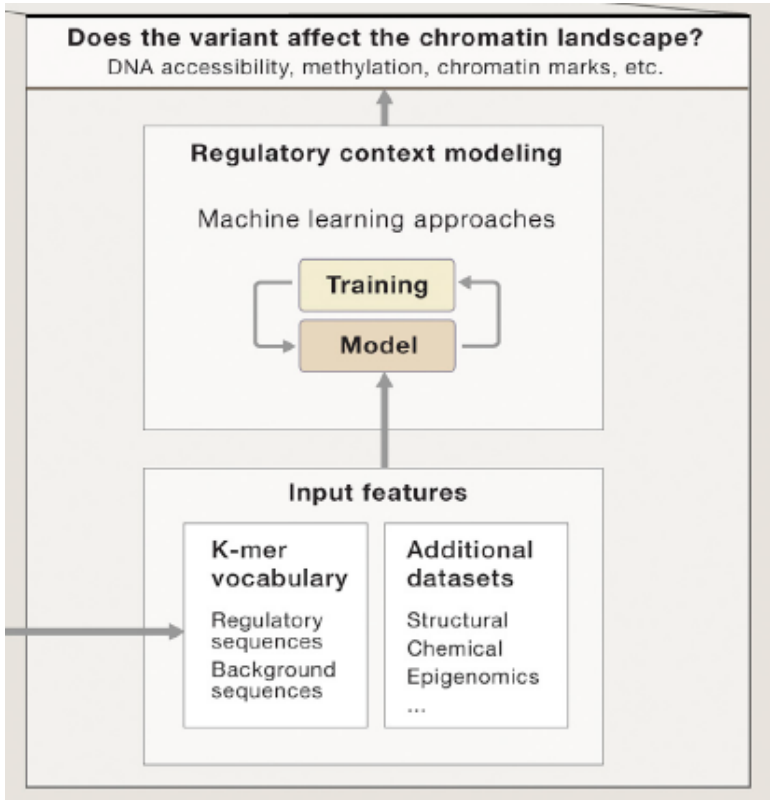
SNPs in the genomic regions may alter a binding site of a specific TFs, such as PU.1 and chromatin states change in the same region

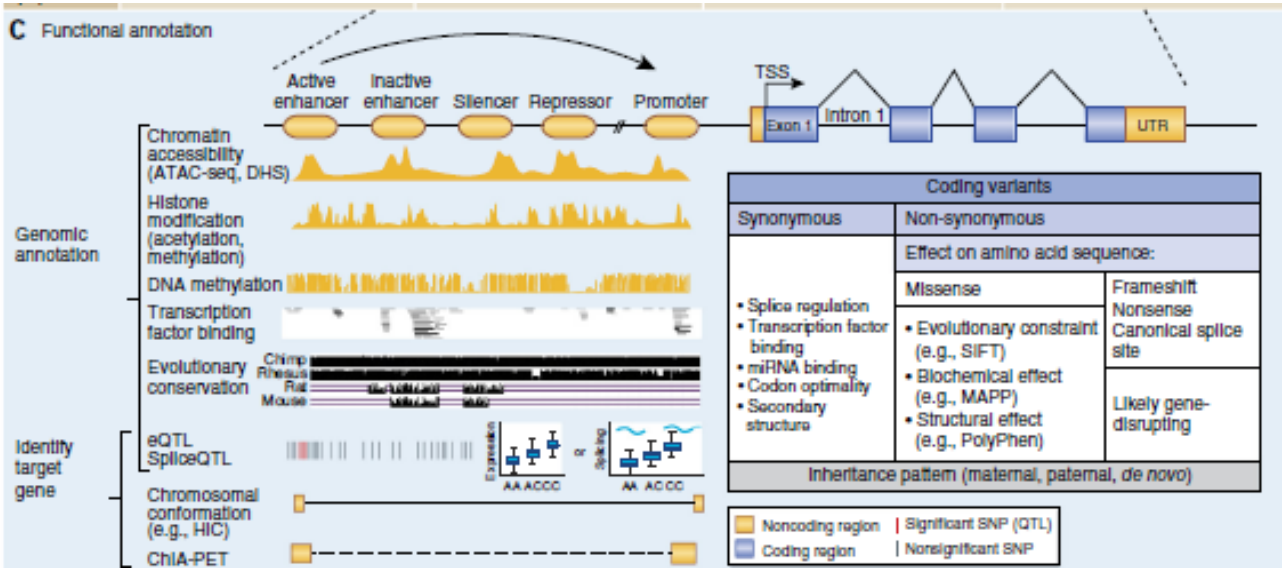


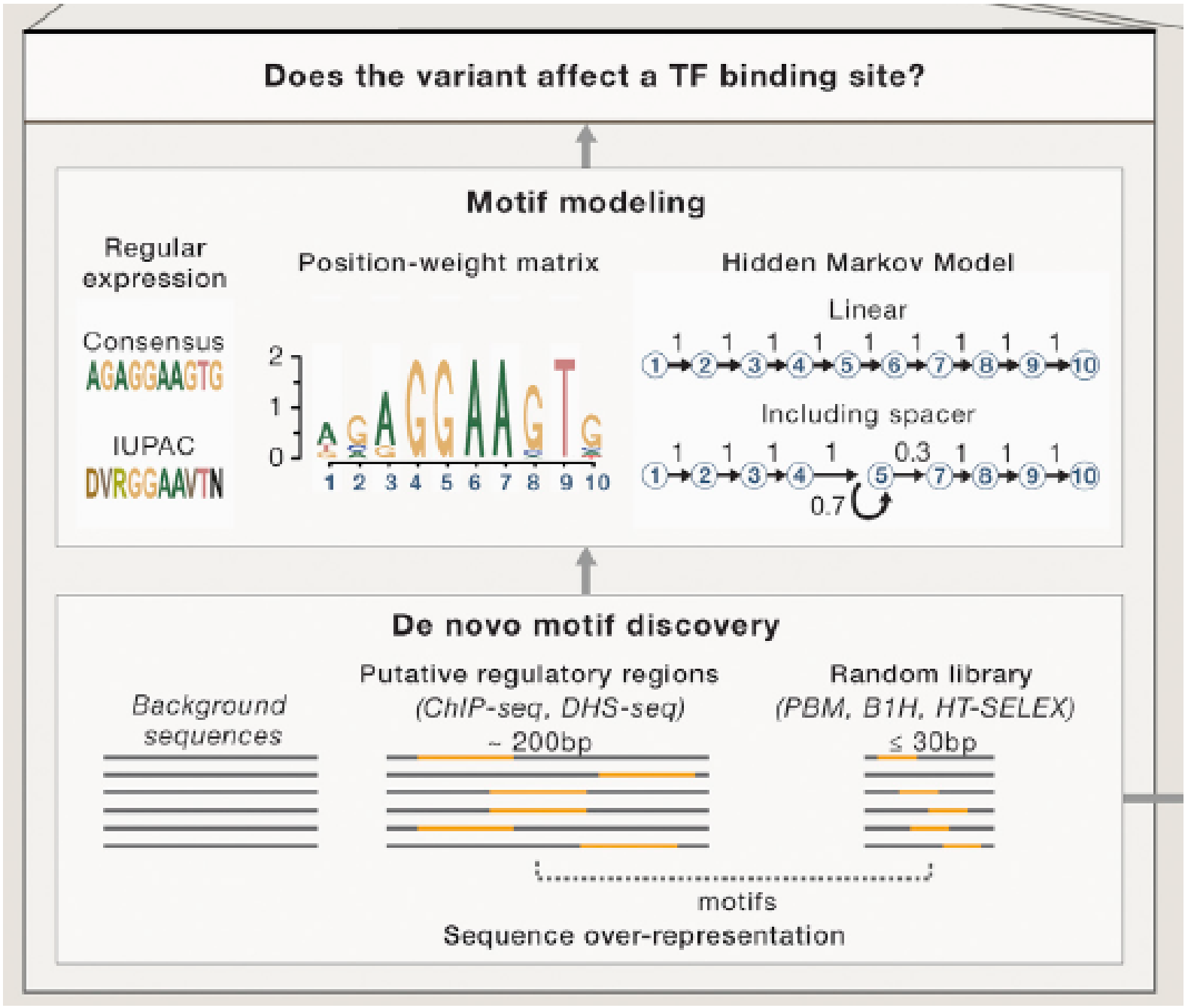
SNPs in the genomic regions may alter a binding site of a specific TFs, such as PU.1 while chromatin states change a whidespread region

SNPs in the genomic regulatory regions DEFINITION

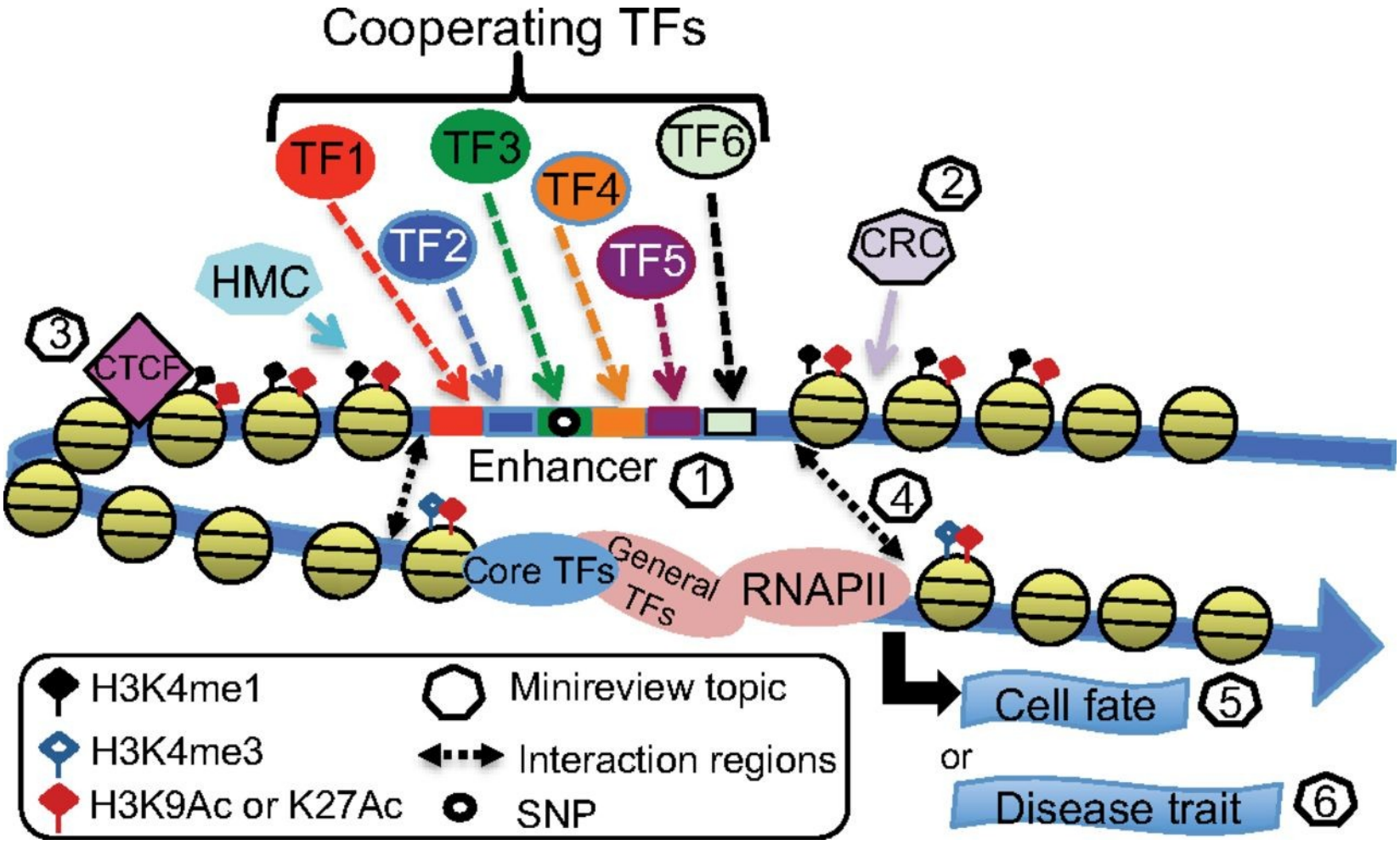



Does the variant affect the chromatin landscape?
DNA accessibility, methylation, chromatin marks, etc.
Regulatory context modeling
Machine learning approaches
Training
Model
Input features
K-mer vocabulary
Regulatory sequences
Background sequences
Additional datasets
Structural
Chemical
Epigenomics
...


C Functional annotation
Active enhancer
Inactive enhancer
Silencer
Repressor
Promoter
TSS
Exon 1
Intron 1
UTR
Genomic annotation
Chromatin accessibility (ATAC-seq, DHS)
Histone modification (acetylation, methylation)
DNA methylation
Transcription factor binding
Evolutionary conservation
Chimp
Rhesus
Rat
Mouse
Identify target gene
eQTL
SpliceQTL
Expression
Splicing
AA ACCC
AA AC CC
Chromosomal conformation (e.g., HIC)
ChIA-PET
Coding variants
Synonymous
Non-synonymous
Effect on amino acid sequence:
Missense
Frameshift
Nonsense
Canonical splice site
• Splice regulation
• Transcription factor binding
• miRNA binding
• Codon optimality
• Secondary structure
• Evolutionary constraint (e.g., SIFT)
• Biochemical effect (e.g., MAPP)
• Structural effect (e.g., PolyPhen)
Likely gene-disrupting
Inheritance pattern (maternal, paternal, de novo)
Noncoding region
Coding region
Significant SNP (QTL)
Nonsignificant SNP

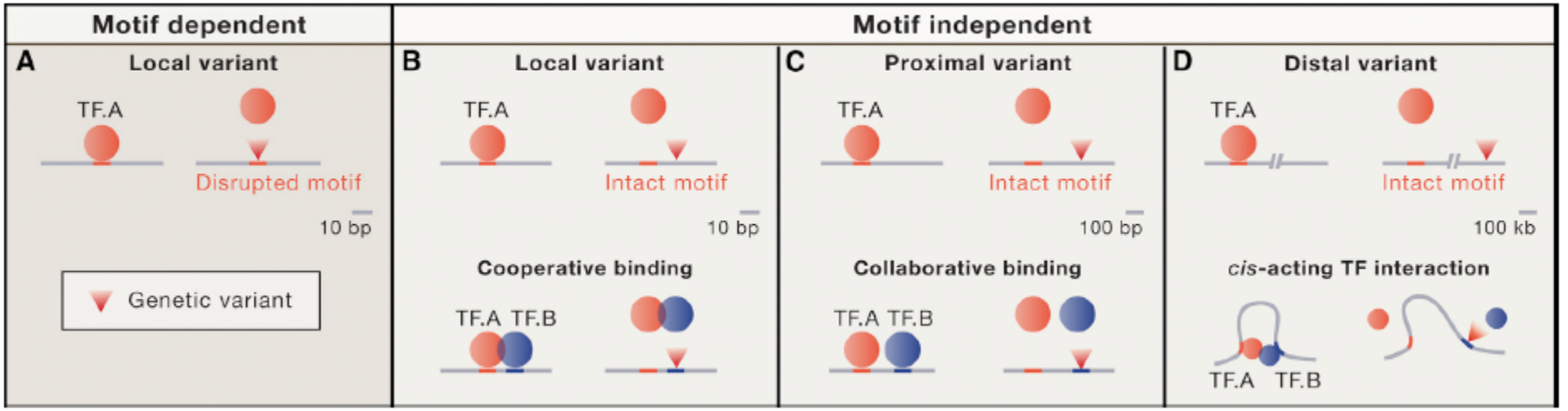

Does the variant affect a TF binding site?
Motif modeling
Regular expression
Consensus
AGAGGAAGTG
IUPAC
DVRGGAAVTN
Position-weight matrix
2
1
0
1 2 3 4 5 6 7 8 9 10
Hidden Markov Model
Linear
1 1 1 1 1 1 1 1 1 1
1 2 3 4 5 6 7 8 9 10
Including spacer
1 1 1 1 0.3 1 1 1
1 2 3 4 5 7 8 9 10
0.7
De novo motif discovery
Background sequences
Putative regulatory regions
(ChIP-seq, DHS-seq)
~ 200bp
Random library
(PBM, B1H, HT-SELEX)
≤ 30bp
motifs
Sequence over-representation

Genome-wide characterizations of regulatory regions.

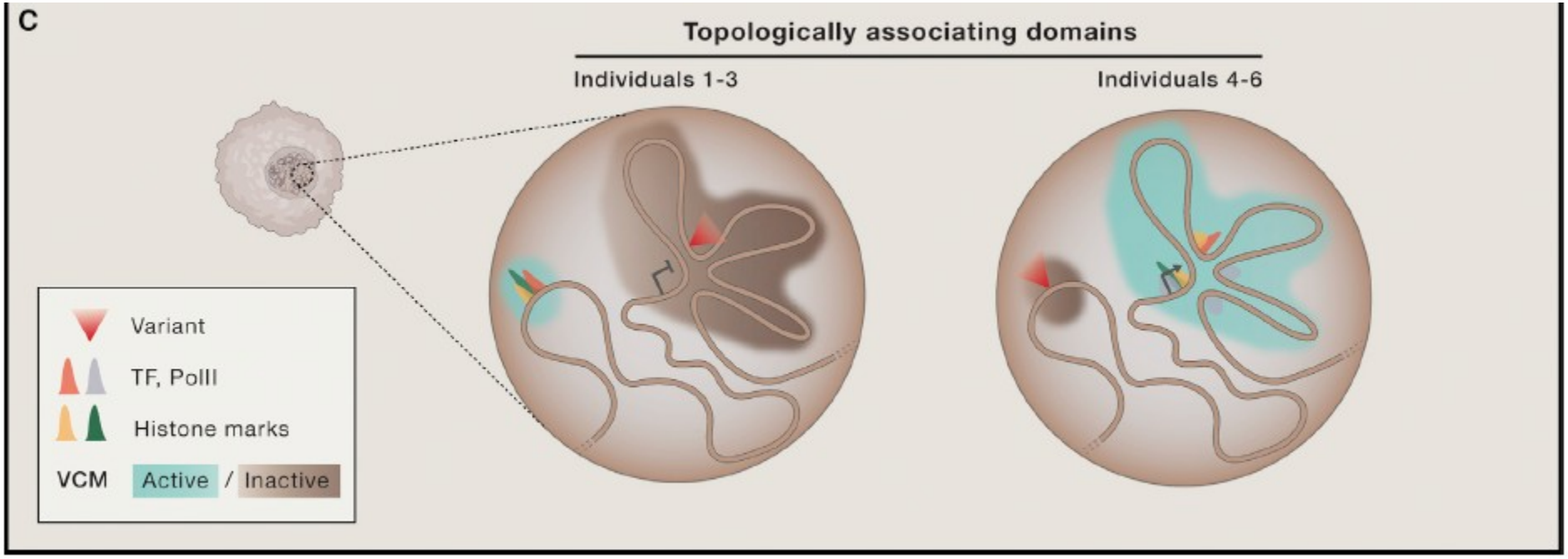


Peggy J. Farnham J. Biol. Chem. 2012;287:30885-30887

SNPs mechanims for alteration of regulatory transcription factors complexes



SNPs may change long range interactions



Framework for interpretation of individual disease-associated variants

- Single nucleotide polymorphisms (SNPs) is the nucleotide variations associated with disease

- Genome-wide association studies (GWAS) have successfully identified thousands of common genetic variants associated with complex diseases (http://www.ebi.ac.uk/gwas/)

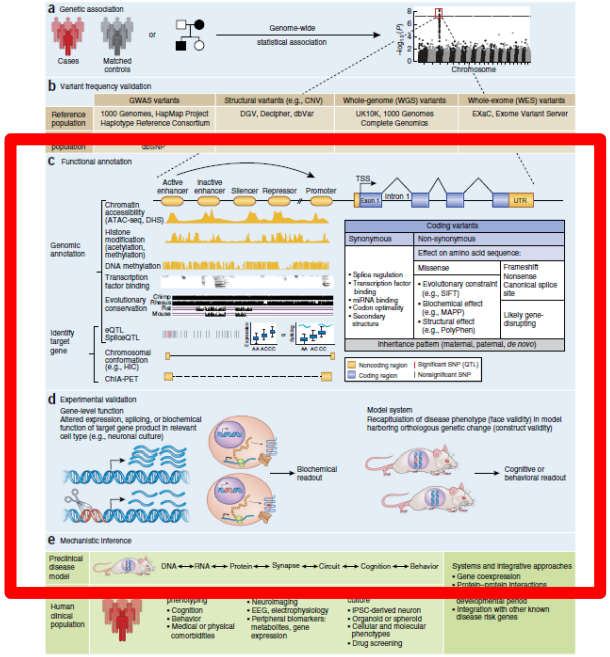
- Functional annotation: to define genomic regulatory regions by genome-wide integration data

- Experimental validation

- Disease Animal models

- Correlation between molecular mechanisms and disease symptoms

- Drug Discovery



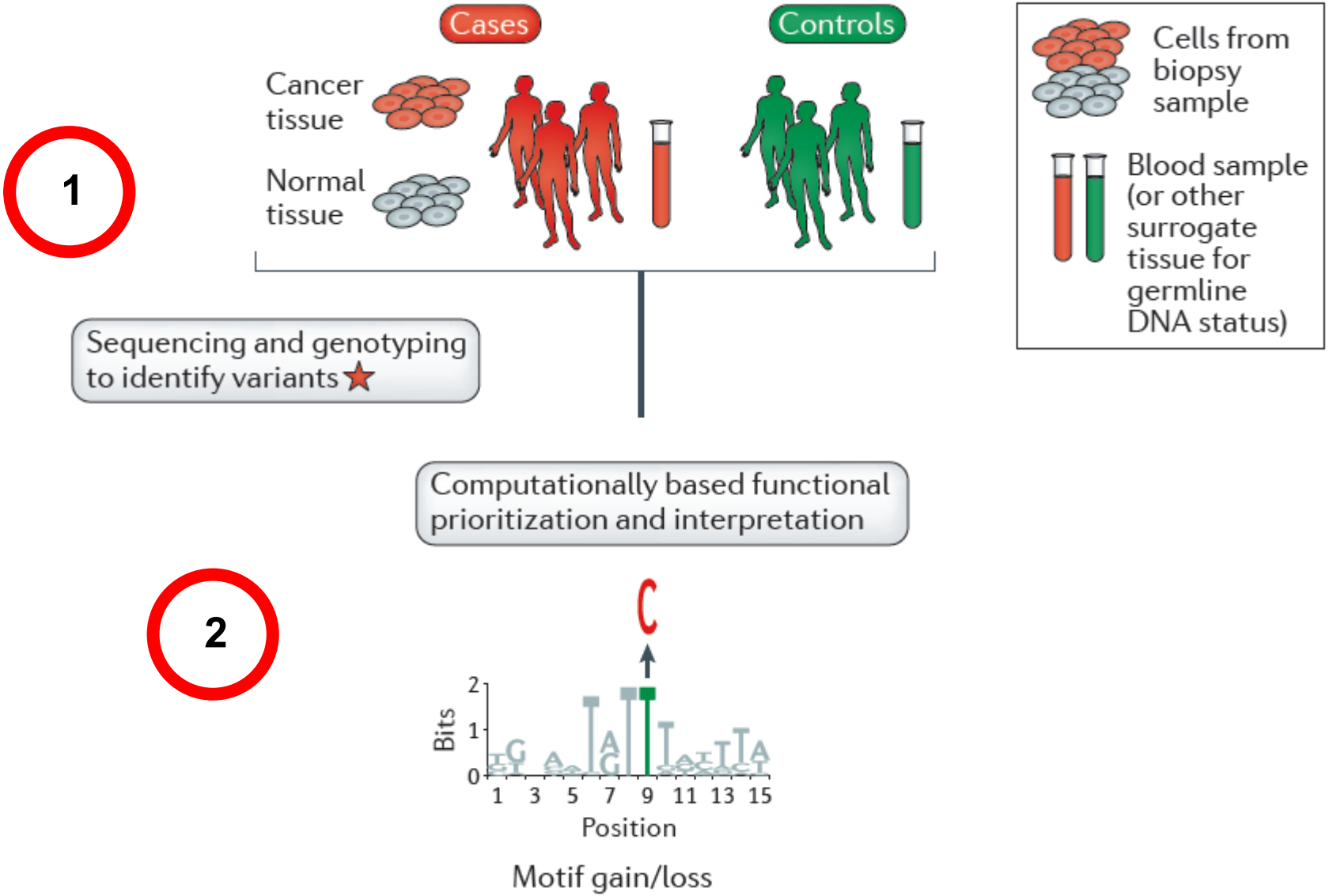
Role of non-coding sequence variants in cancer

Ekta Khurana[1–4], Yao Fu[5], Dimple Chakravarty[2,6], Francesca Demichelis[2,3,7], Mark A. Rubin[1,2,6] and Mark Gerstein[8–10]

Abstract | Patients with cancer carry somatic sequence variants in their tumour in addition to the germline variants in their inherited genome. Although variants in protein-coding regions have received the most attention, numerous studies have noted the importance of non-coding variants in cancer. Moreover, the overwhelming majority of variants, both somatic and germline, occur in non-coding portions of the genome. We review the current understanding of non-coding variants in cancer, including the great diversity of the mutation types — from single nucleotide variants to large genomic rearrangements — and the wide range of mechanisms by which they affect gene expression to promote tumorigenesis, such as disrupting transcription factor-binding sites or functions of non-coding RNAs. We highlight specific case studies of somatic and germline variants, and discuss how non-coding variants can be interpreted on a large-scale through computational and experimental methods.

SNPs with an impact in tumorigenesis

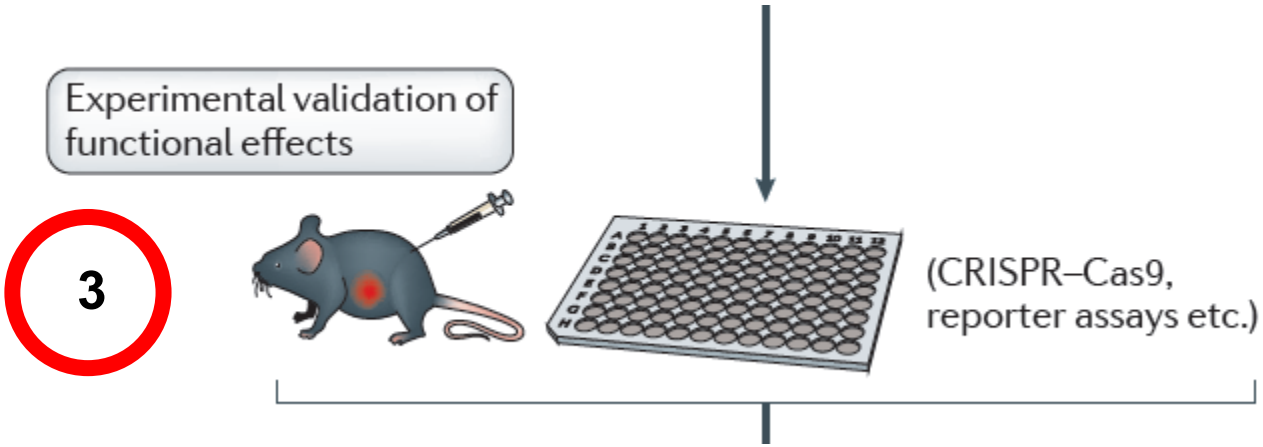
Steps for studying the role of SNP



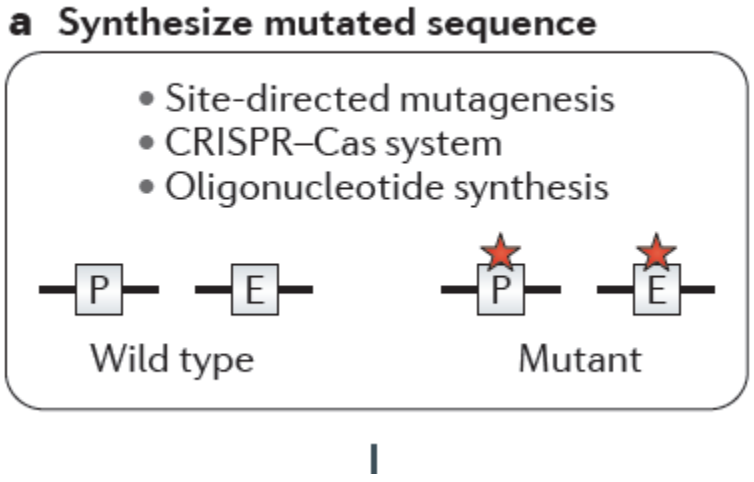
SNPs may have an impact in tumorigenesis

2

FUNCTIONAL ANNOTATION OF SNPs



SNPs ESPERIMENTAL VALIDATIONS



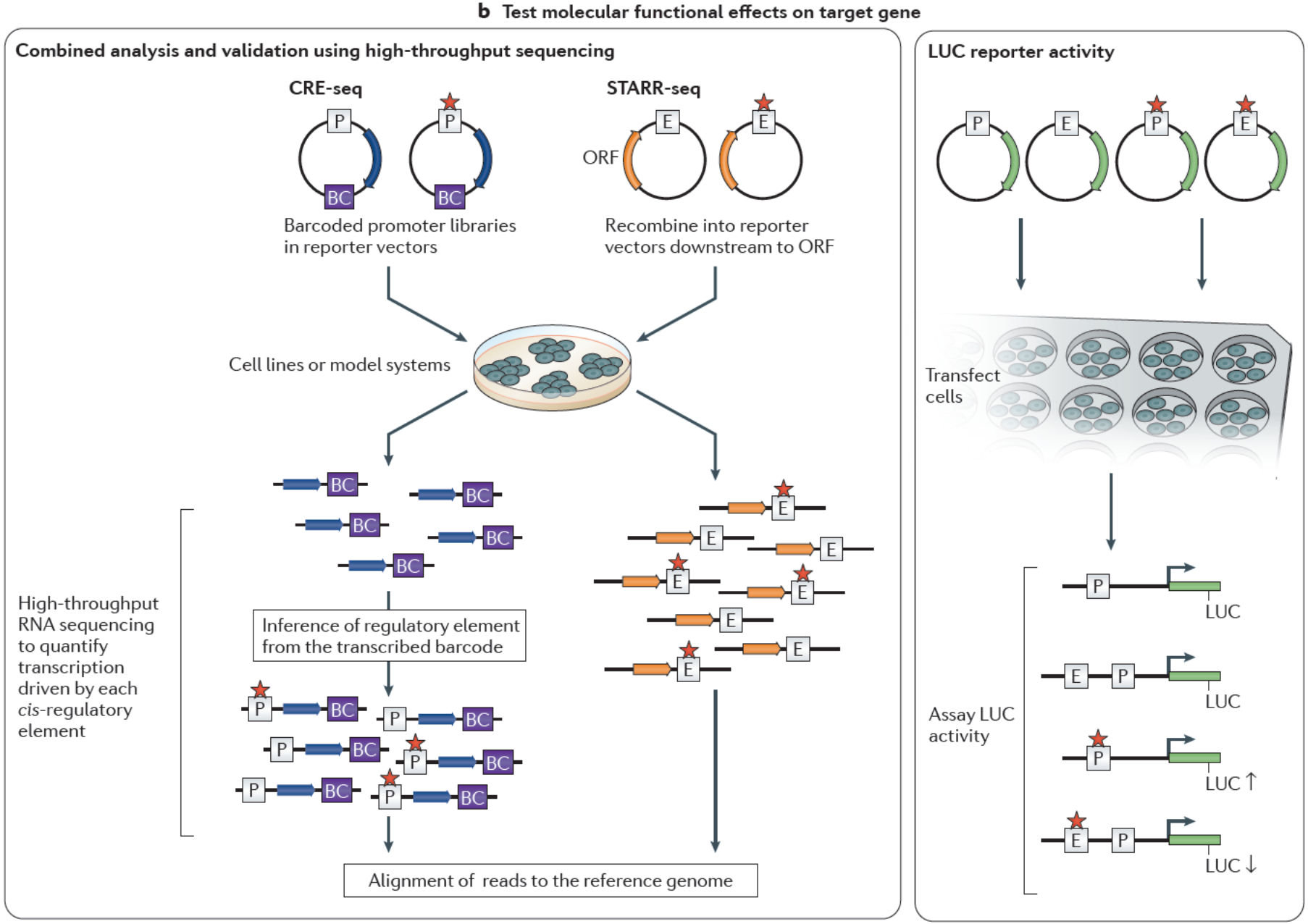
Task 5- Design an experiment by using plasmid with luciferase reporter

- How you create the mutation in the plasmid

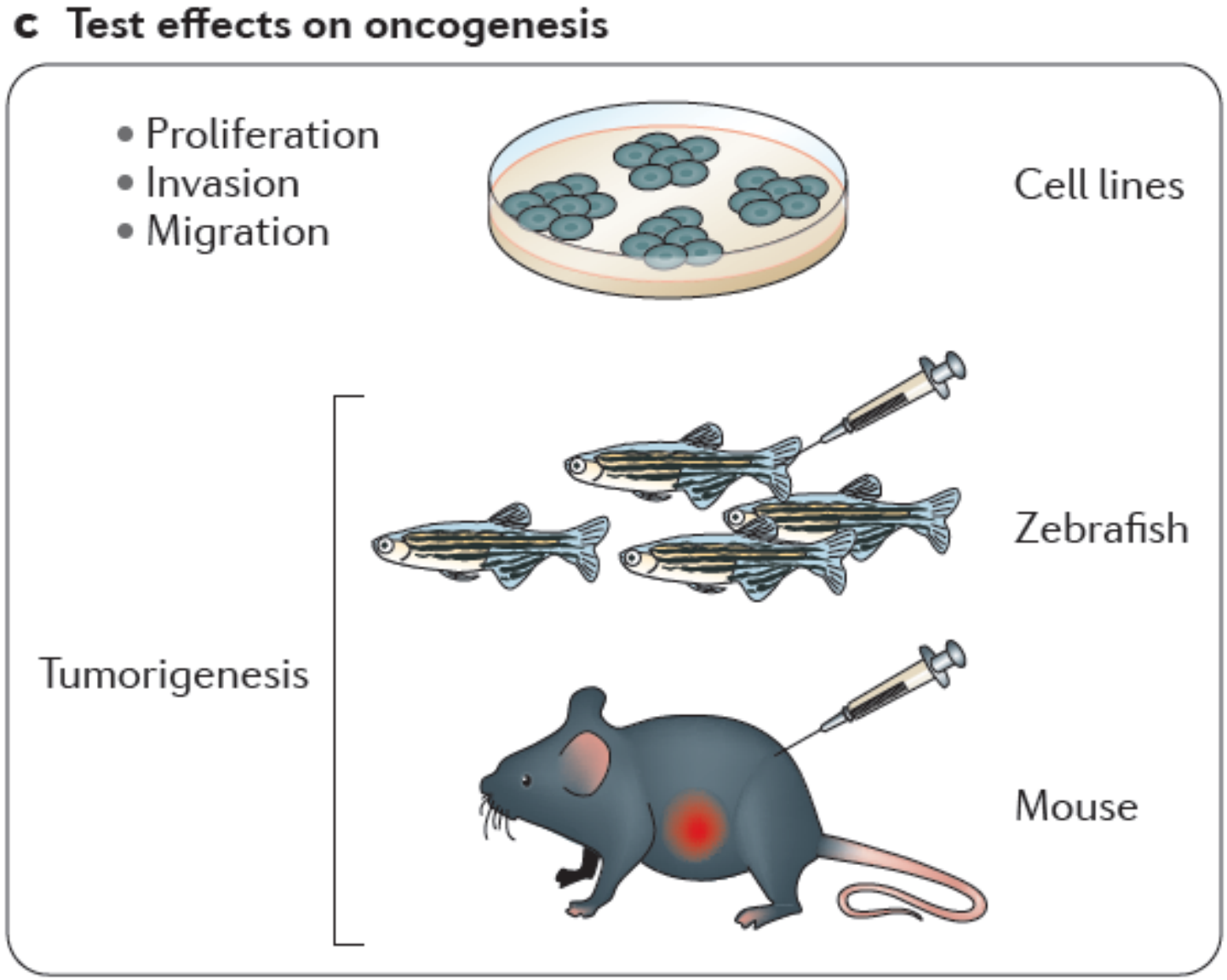
- Which are the samples of your experiment? Positive control and negative control

- Data interpretation

MOLECULAR FUNCTIONAL EFFECTS

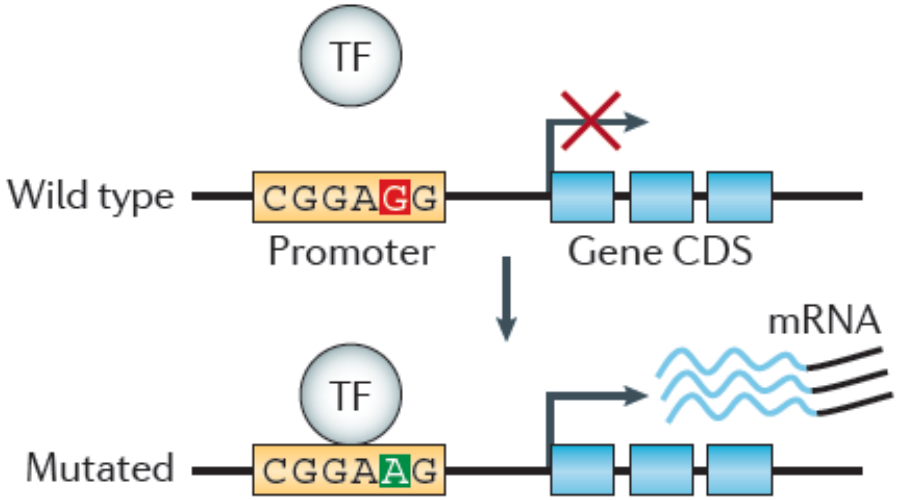


BIOLOGICAL FUNCTION TESTS

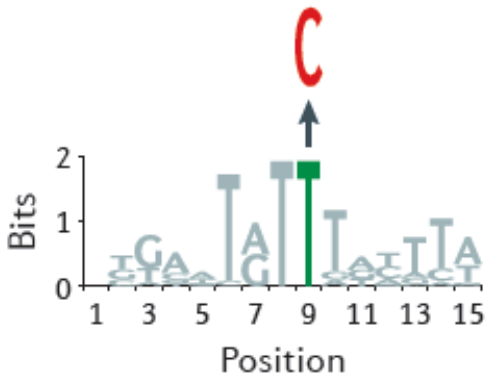


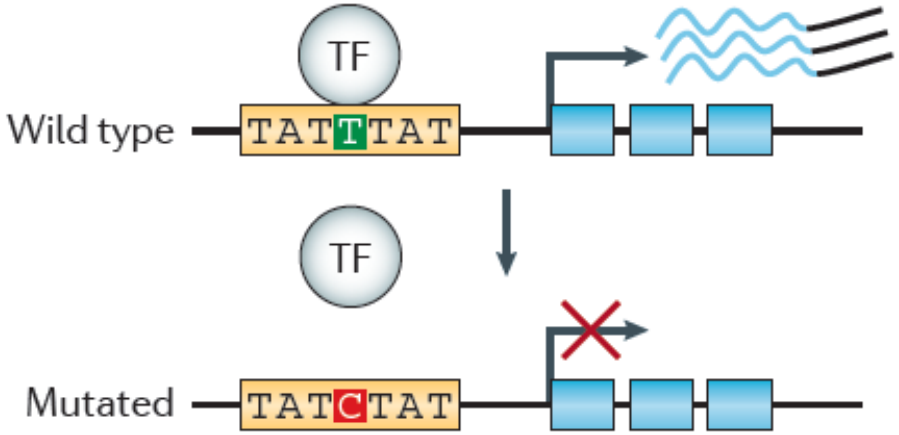
SNPs types functions:

Ba Gain of motif



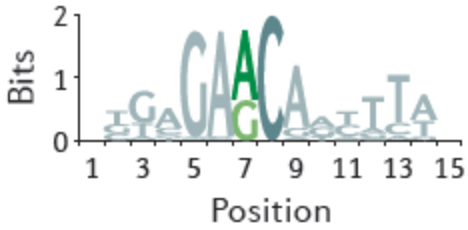
Bb Loss of motif



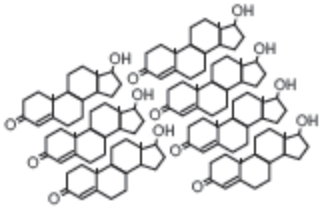


SNPs types functions:

Bc Altered binding effects in hormonal cancers

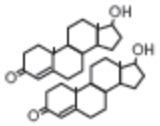


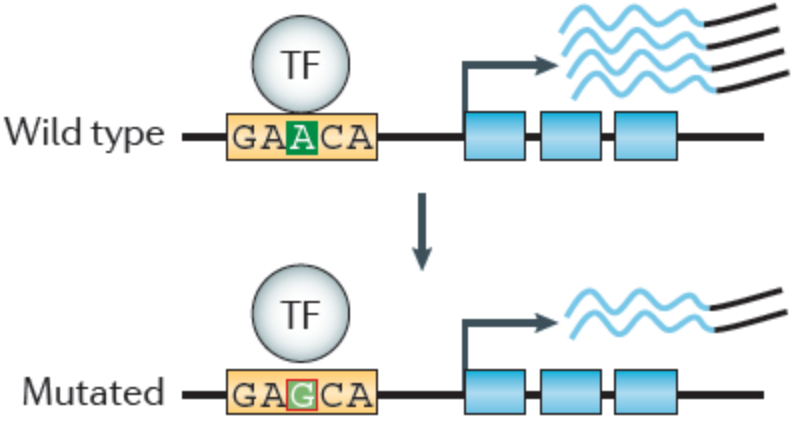
Hormones

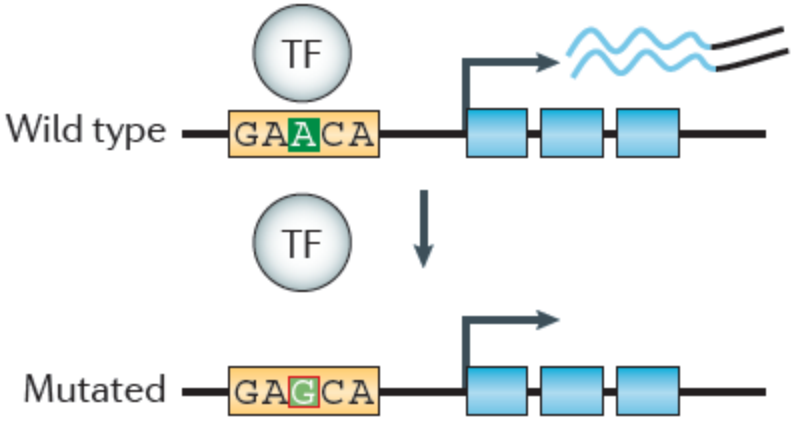


High

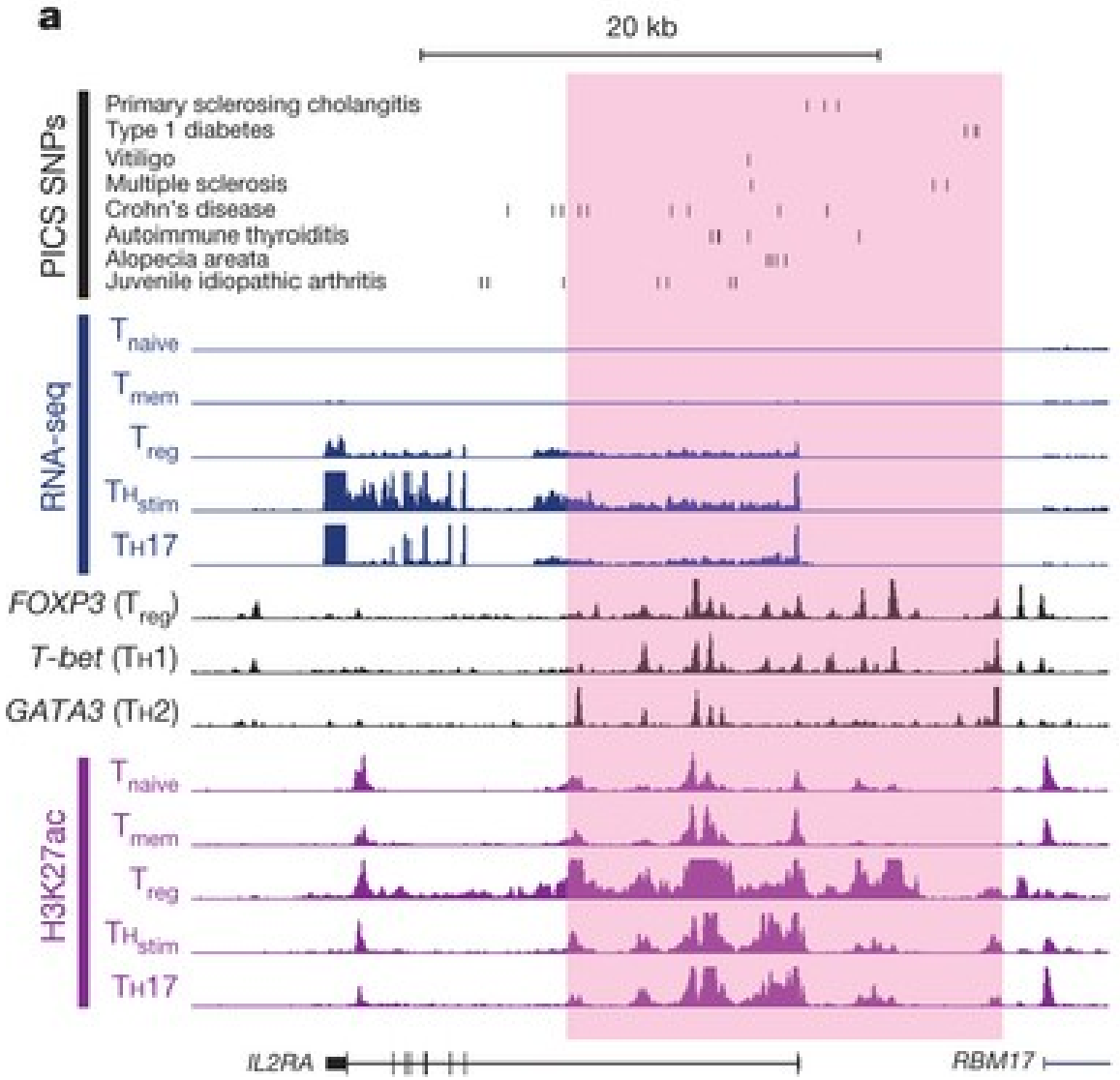
Low



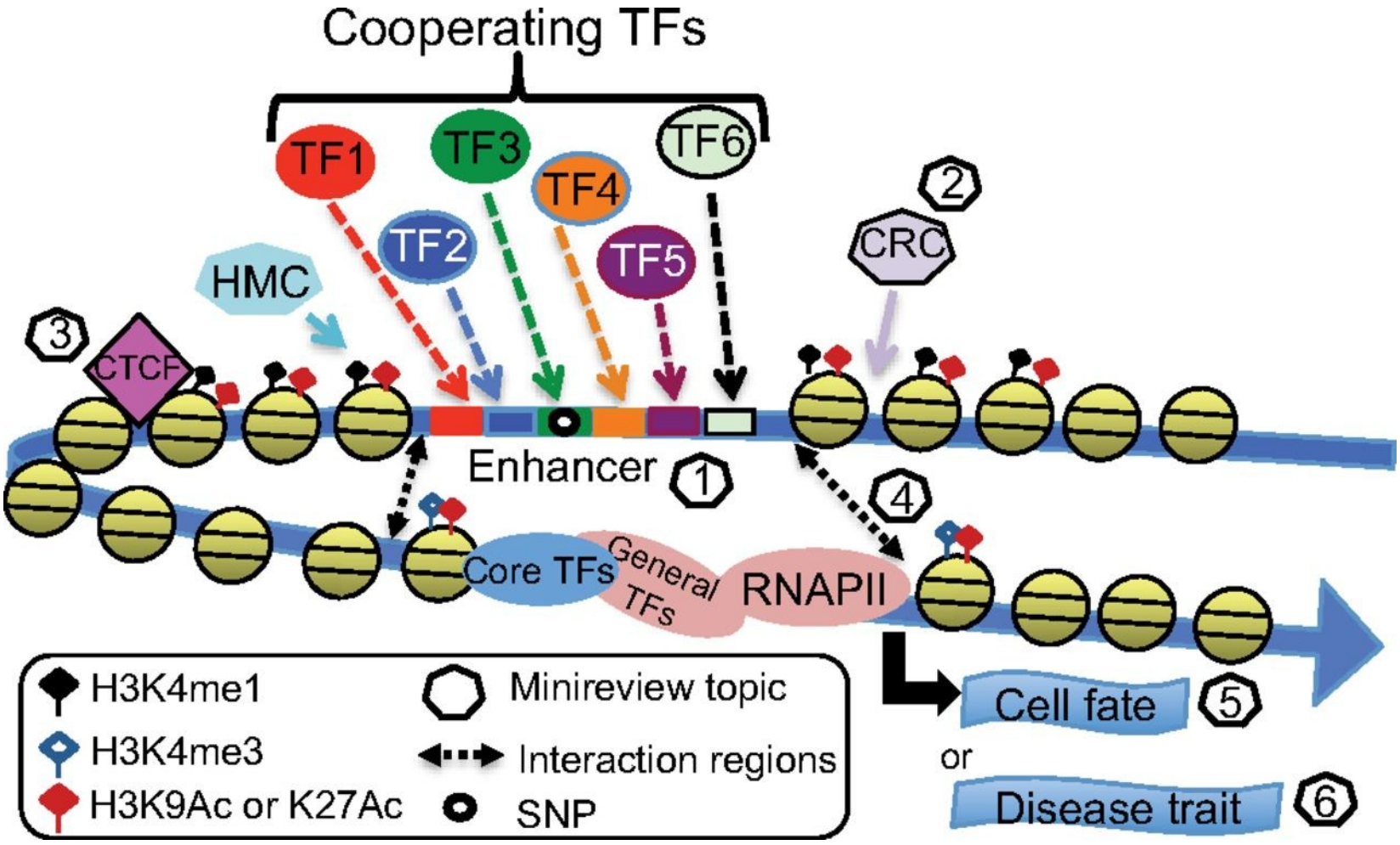




Genome-wide data describe activation state of specific gene locus and the correlation of these features with disease open the way to understand disease outcome



Genome-wide characterizations of regulatory regions.



Peggy J. Farnham J. Biol. Chem. 2012;287:30885-30887

Framework for interpretation of individual disease-associated variants

- Single nucleotide polymorphisms (SNPs) is the nucleotide variations associated with disease

- Genome-wide association studies (GWAS) have successfully identified thousands of common genetic variants associated with complex diseases (http://www.ebi.ac.uk/gwas/)

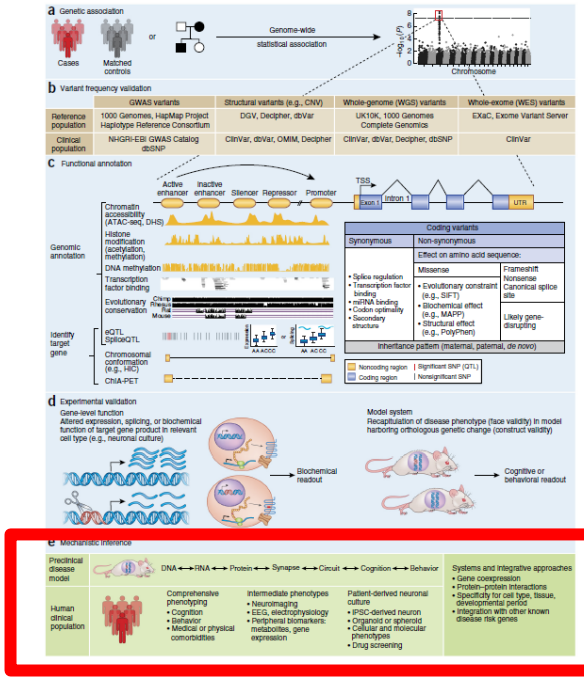
- Functional annotation: to define genomic regulatory regions by genome-wide integration data

- Experimental validation

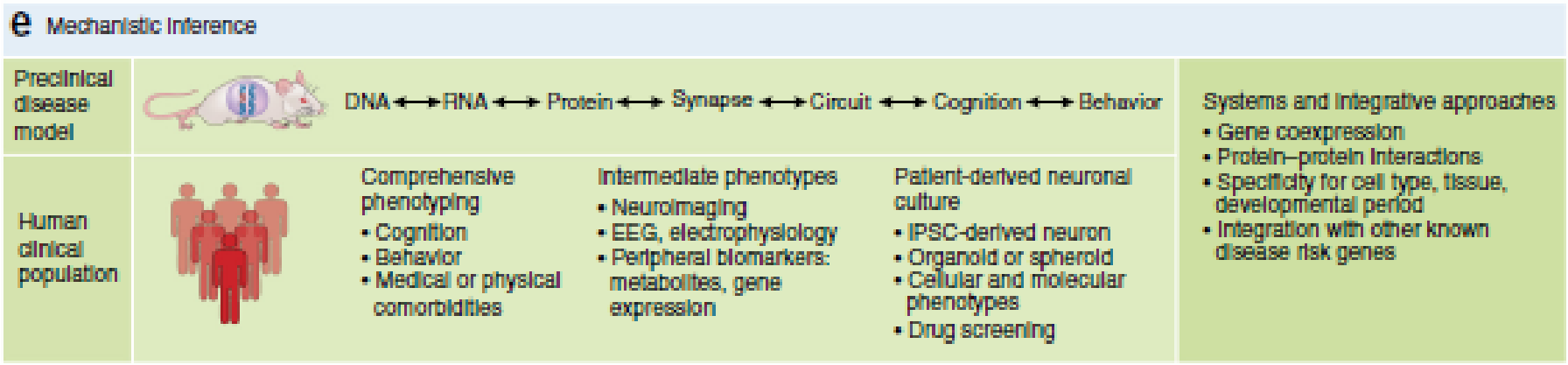
- Disease Animal models

- Correlation between molecular mechanisms and disease symptoms

- Drug Discovery



Correlation of SNP/functions with several clinical analysis

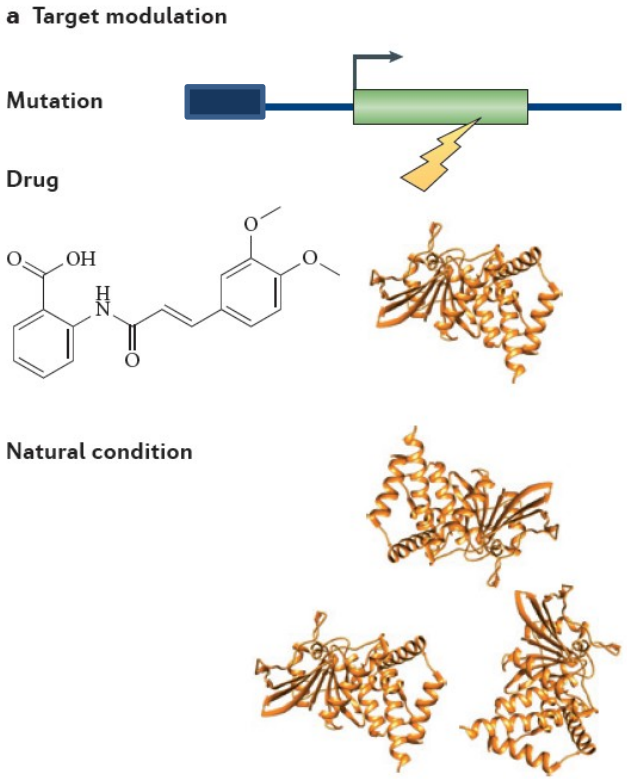


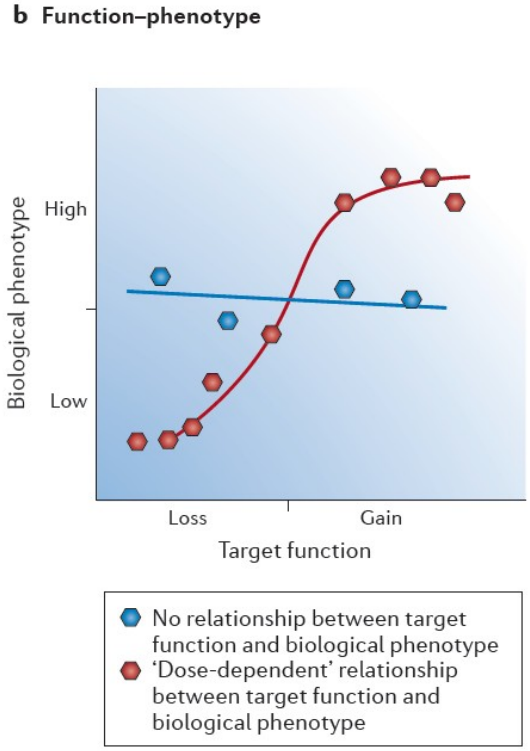
How can we use these knowledge?

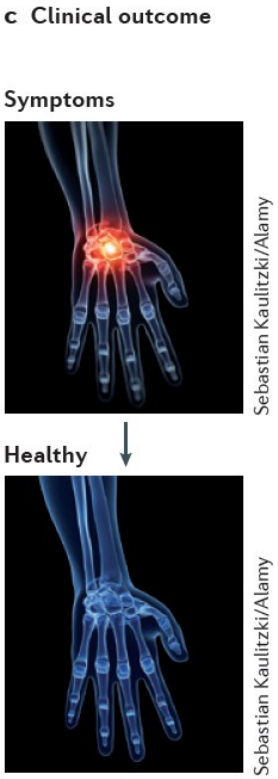
EXAMPLE

Gene expression alteration in disease May be used as BIOMARKERS (molecules acting as sensor of disease)

Gene expression alteration in disease May be used as DRUG TARGET (drug discovery to stop disease and restore health)







In Summary:

- **Functional genomics is a field of molecular biology based on genome-wide sequencing data.**
- **Genome-wide sequencing data describe genomic regulatory regions that control gene expression**
- **Gene expression disregulation may be linked to the disease**
- **Understanding molecular mechanisms of disease outcome opens the way to discovery drug and identify biomarkers**

http://biologia.i-learn.unito.it/:

1. Lecture PDFs: the slides we used during the class
2. Textbook: *reviews* that will give the necessary background and lessons first part
3. Research Papers: articles that we will analyze
4. Bibliography: scientific literature concerning the subject
5. Audio and Main Concept Lessons

EXAM

Students are expected to demonstrate:

1. Knowledge of **basic** concepts
2. Understanding of **specific** concepts
3. Comprehension of experimental **methodology**
4. **Solving problem** that we have discuss during lesson

Evaluation:

EXAMS is based on lessons and is composed to multiple choice questions and two open questions.