



Corso di Laurea Magistrale in
INFORMATICA

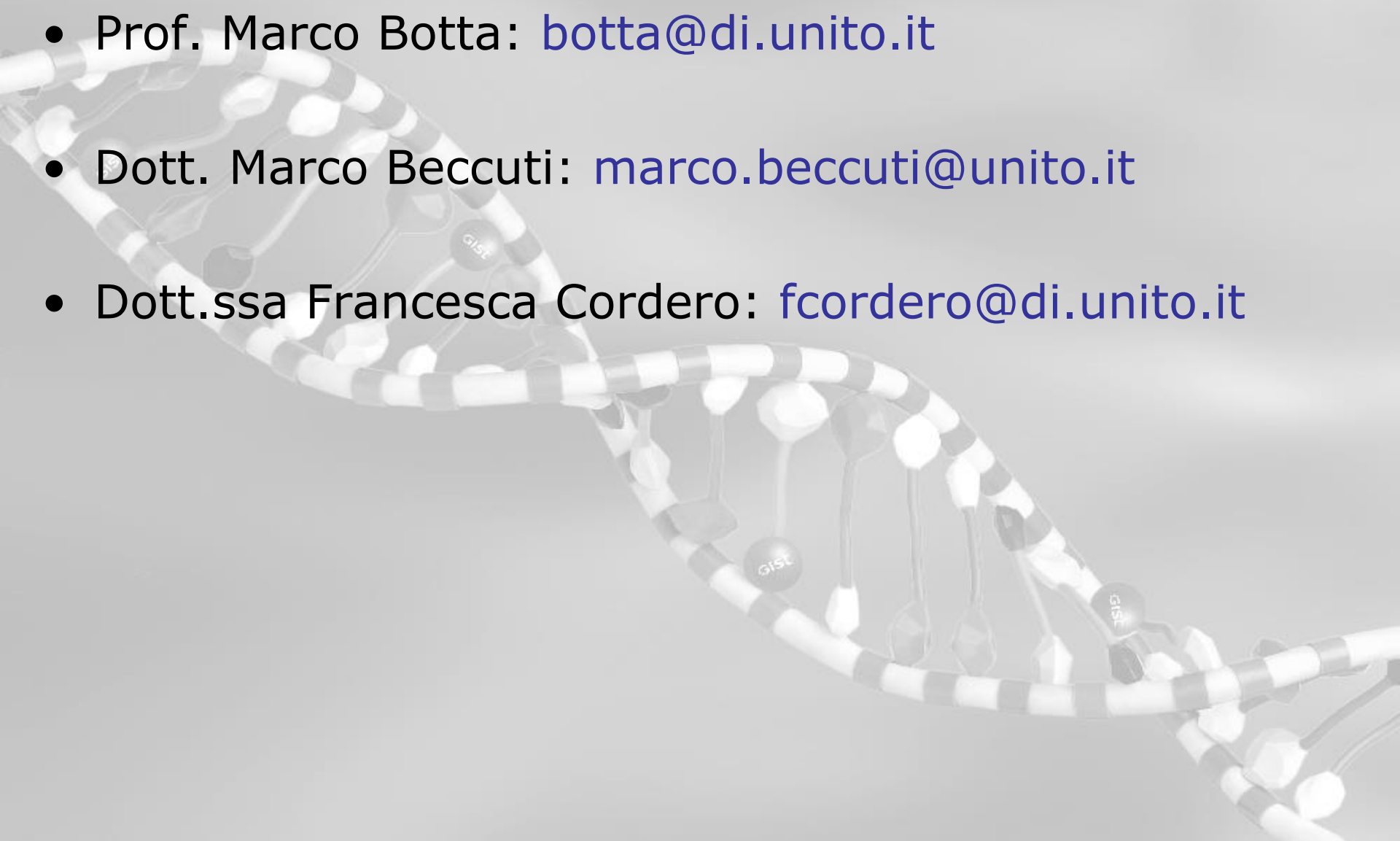
Bioinformatica

A.A. 2024/2025

Prof. Marco Botta
Prof. Marco Beccuti
Prof.ssa Francesca Cordero

Contatti

- Prof. Marco Botta: botta@di.unito.it
- Dott. Marco Beccuti: marco.beccuti@unito.it
- Dott.ssa Francesca Cordero: fcordero@di.unito.it



Orari di Ricevimento (per appuntamento)

- Prof. Marco Botta

Tel. 011 6706721

botta@di.unito.it

- Dott.ssa Francesca Cordero

Tel. 011 6706773

fcordero@di.unito.it

- Dott. Marco Beccuti

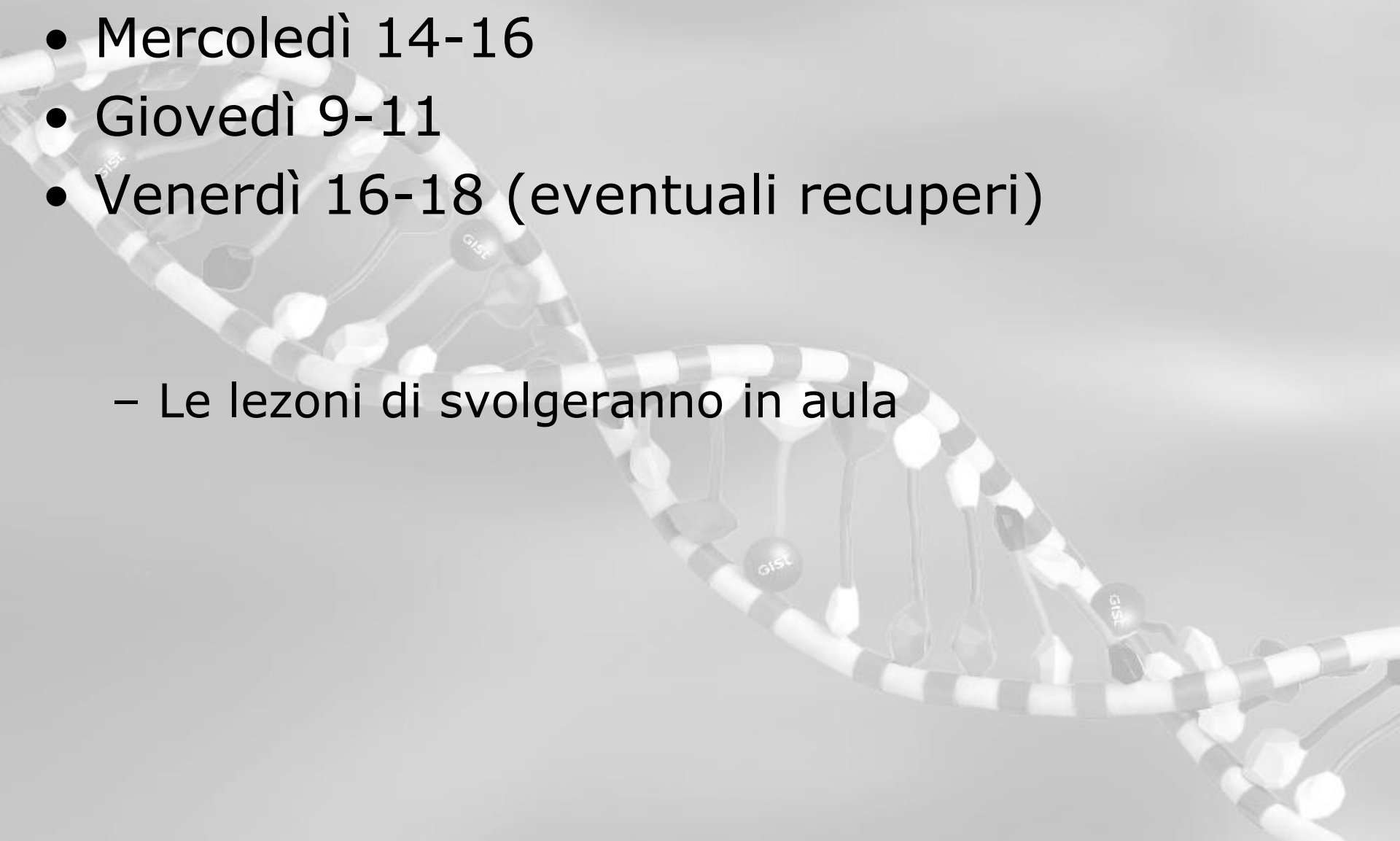
Tel. 011 6706711

marco.beccuti@unito.it

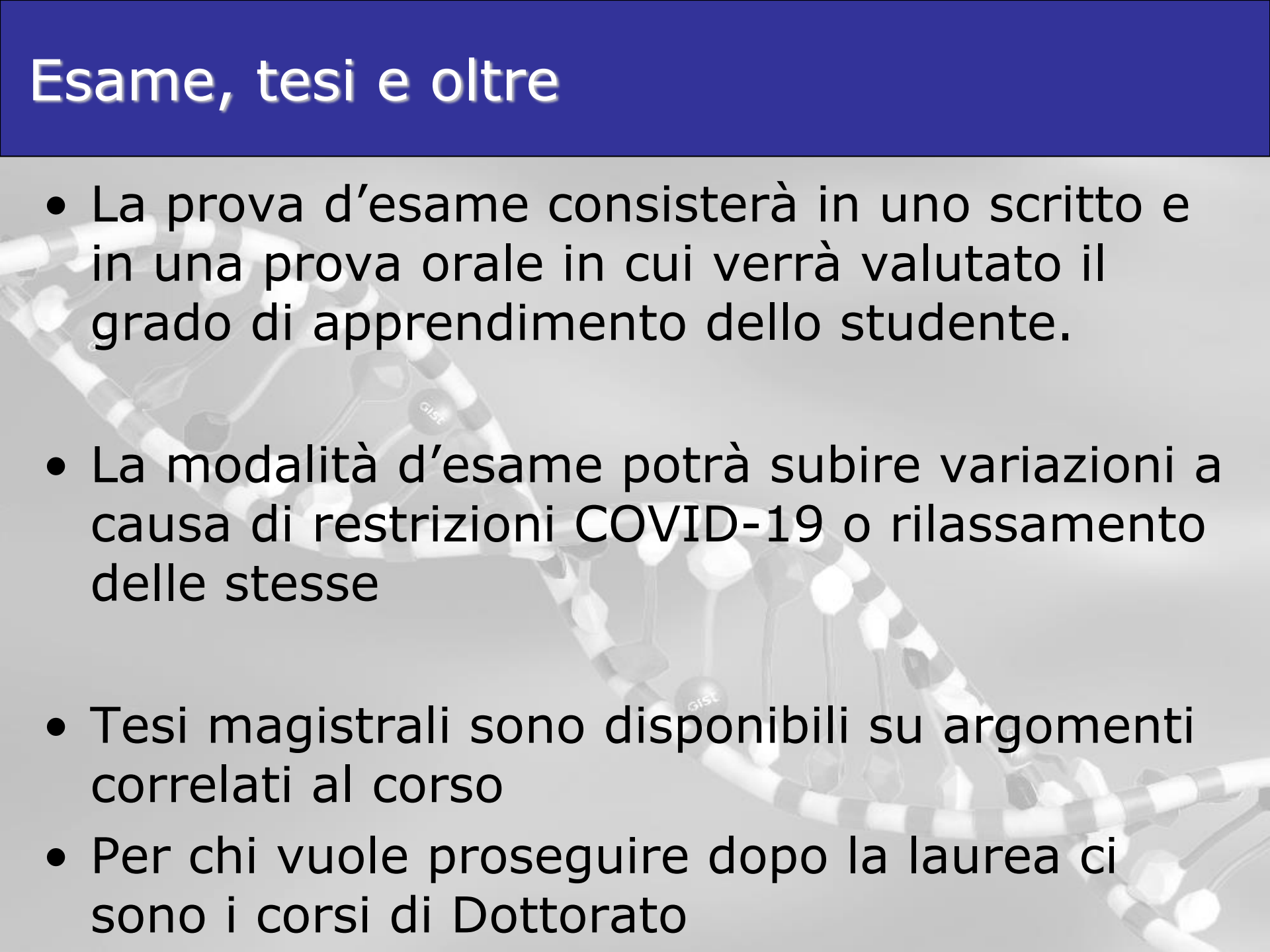


Orario lezioni

- Mercoledì 14-16
- Giovedì 9-11
- Venerdì 16-18 (eventuali recuperi)
 - Le lezioni di svolgeranno in aula



Esame, tesi e oltre

- La prova d'esame consisterà in uno scritto e in una prova orale in cui verrà valutato il grado di apprendimento dello studente.
 - La modalità d'esame potrà subire variazioni a causa di restrizioni COVID-19 o rilassamento delle stesse
 - Tesi magistrali sono disponibili su argomenti correlati al corso
 - Per chi vuole proseguire dopo la laurea ci sono i corsi di Dottorato
- 

Materiale didattico

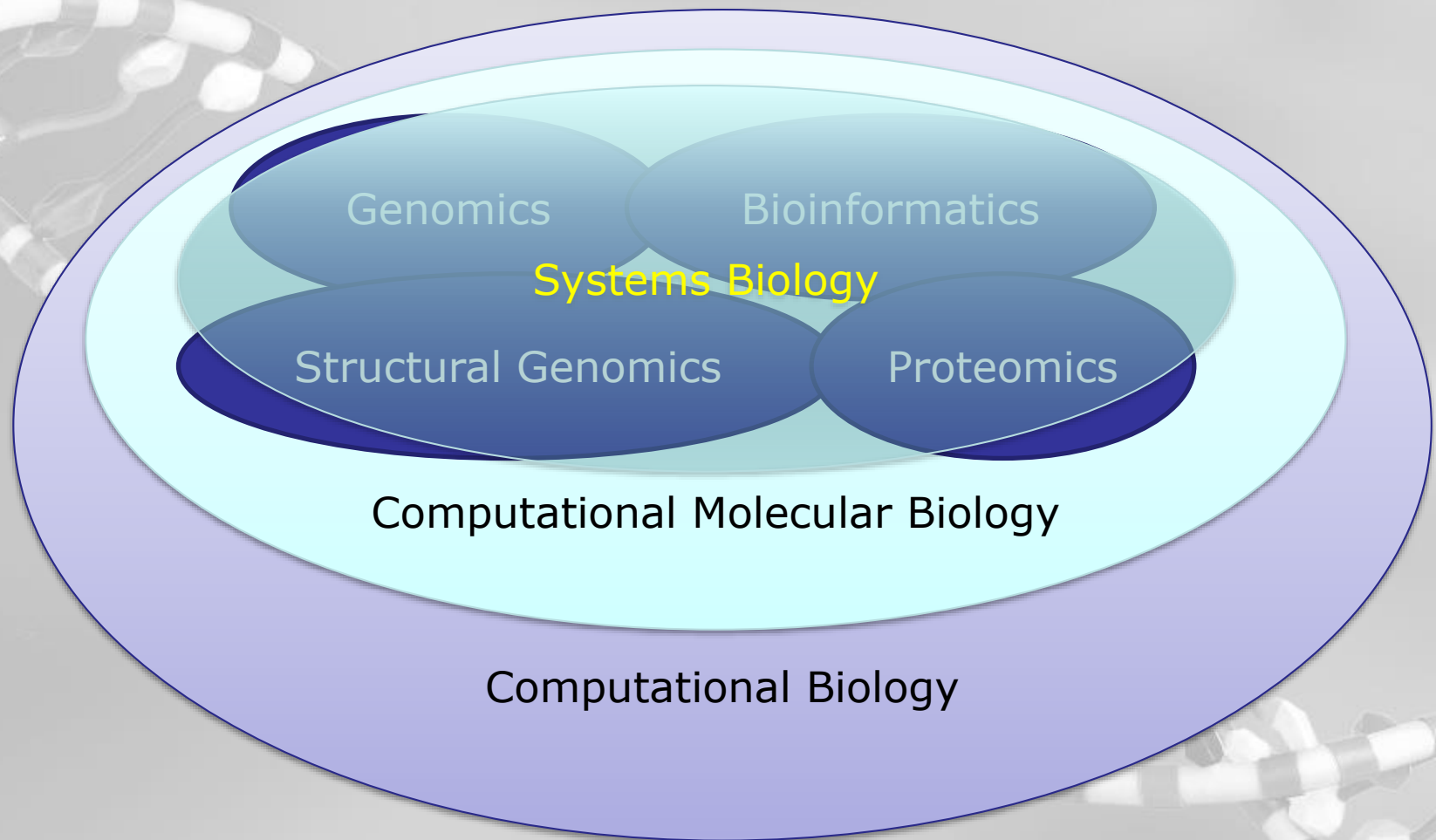
- Slide usate a lezione
- Sito Moodle del corso di laurea
 - <https://informatica.i-learn.unito.it/course/view.php?id=3018>
- [Libri di Bioinformatica](#)



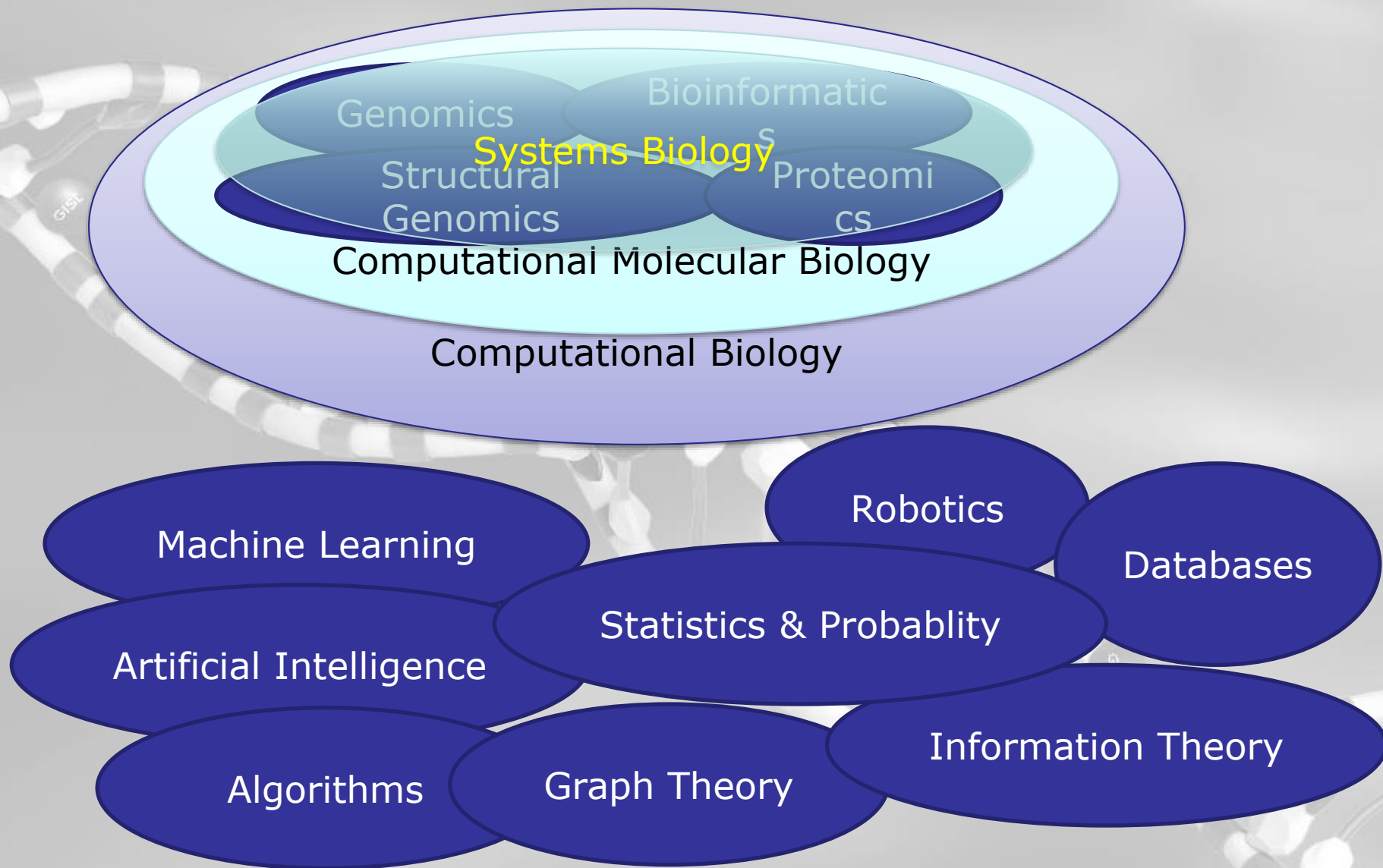
Programma del corso

- Introduzione al corso (oggi)
- Pattern Matching esatto
- Allineamento di sequenze
- Allineamento multiplo e profili
- Algoritmi di Clustering e filogenesi
- Algoritmi di predizione e classificazione
- (Introduzione alla teoria dei grafi (reti biologiche))
- Microarray data analysis
- Tecniche di rappresentazione e individuazione di TFBS
- Algoritmi di pattern matching euristico
- ...

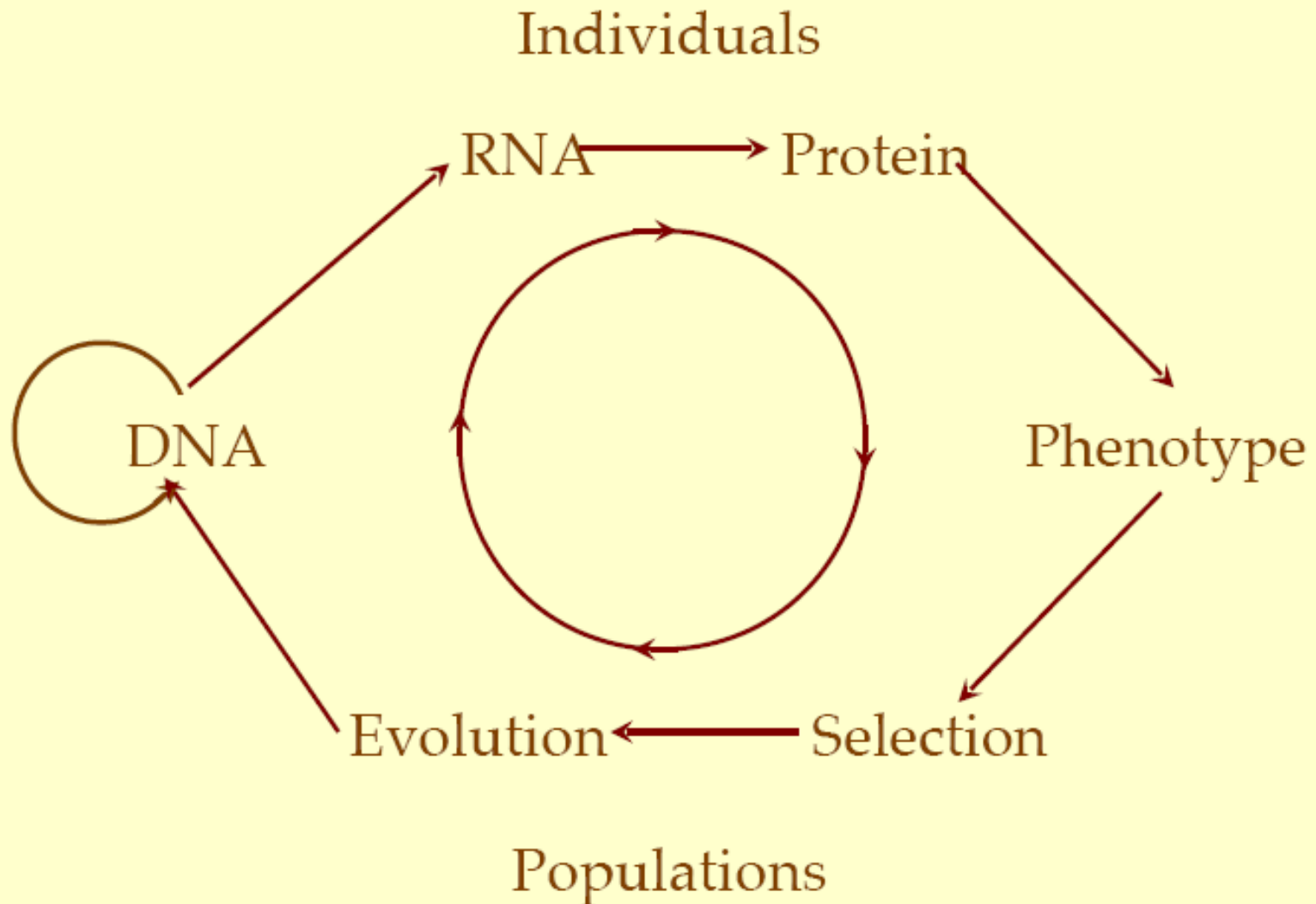
Genomics, Bioinformatics & Computational Biology



Computational Biology & Computer Science



Central Paradigm of Molecular Biology



Central Paradigm of Bioinformatics

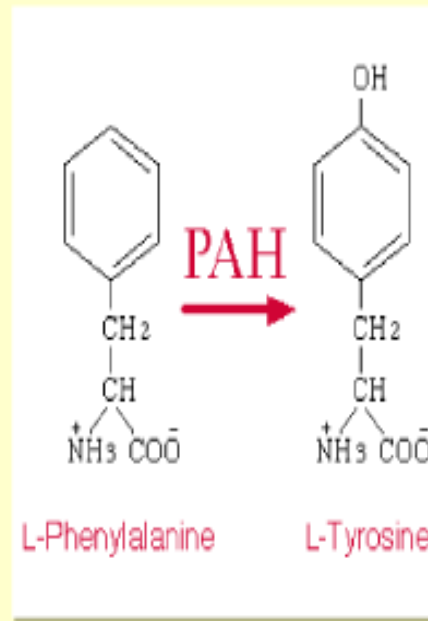
Genetic
Information

MVHLTPEEKT
AVNALWGKVN
VDAVGGEALG
RLLVVYPWTQ
RFFESFGDLS
SPDAVMGNPK
VKAHGKKVLG
AFSDGLAHL
DLKGTFSQLS
ELHCDKLHVD
PENFRLLGNV
LVCVLARNFG
KEFTPQMCAA
YQKVVAGVAN
ALAHKYH

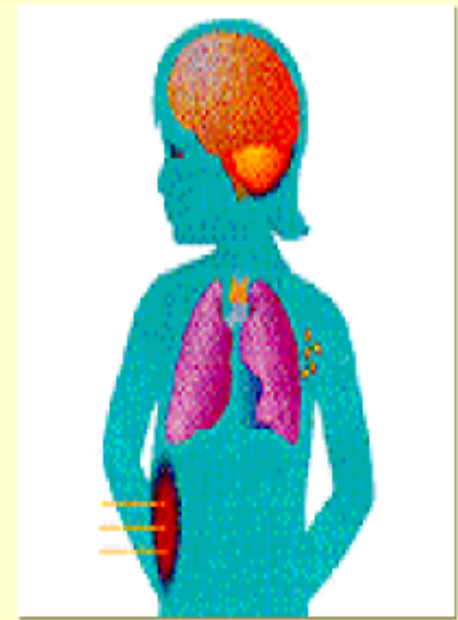


Molecular
Structure

Biochemical
Function



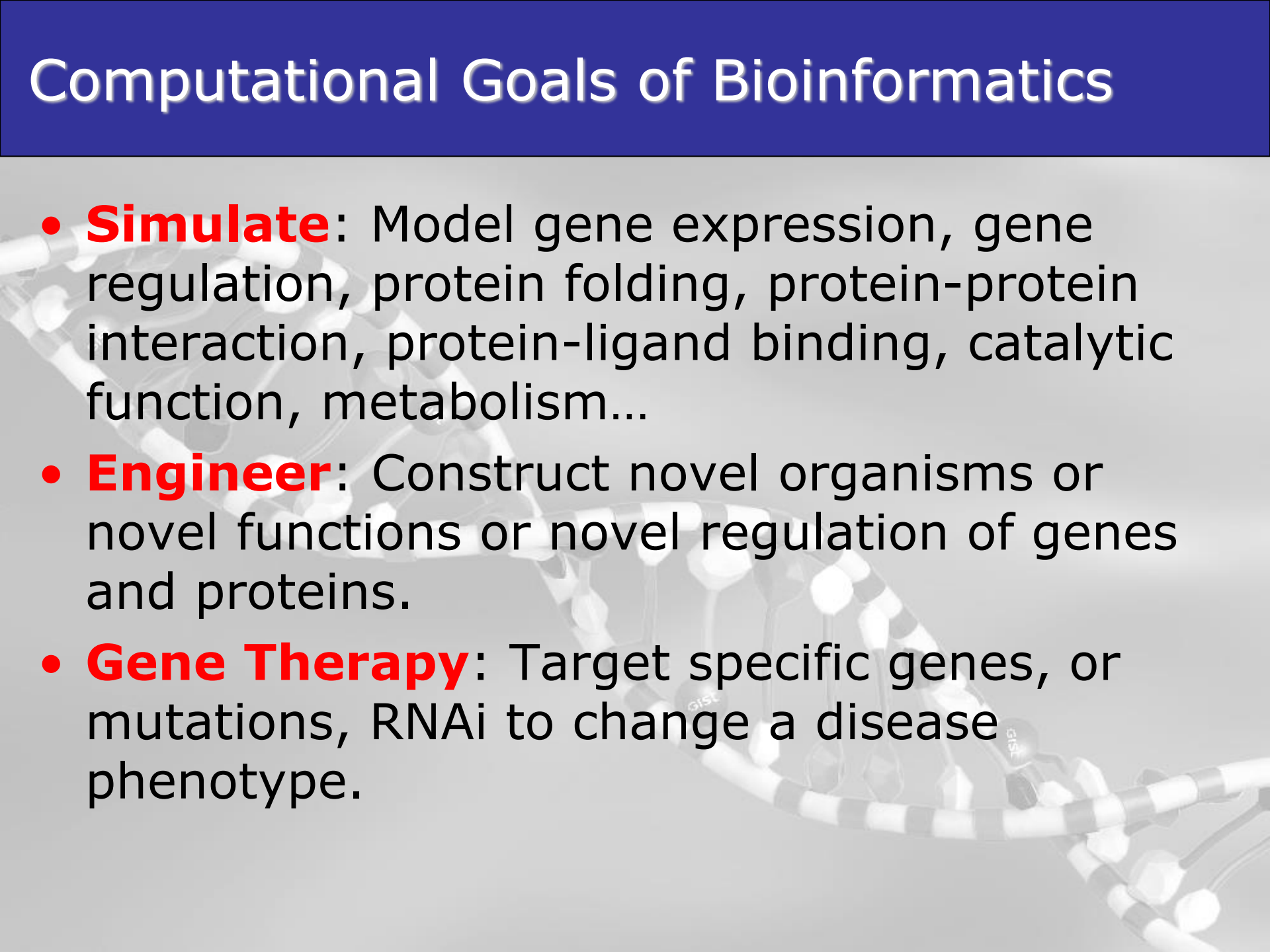
Phenotype
(Symptoms)



Computational Goals of Bioinformatics

- **Learn & Generalize**: Discover conserved patterns (models) of sequences, structures, interactions, metabolism & chemistries from well-studied examples.
- **Prediction**: Infer function or structure of newly sequenced genes, genomes, proteins or proteomes from these generalizations.
- **Organize & Integrate**: Develop a systematic and genomic approach to molecular interactions, metabolism, cell signaling, gene expression...

Computational Goals of Bioinformatics

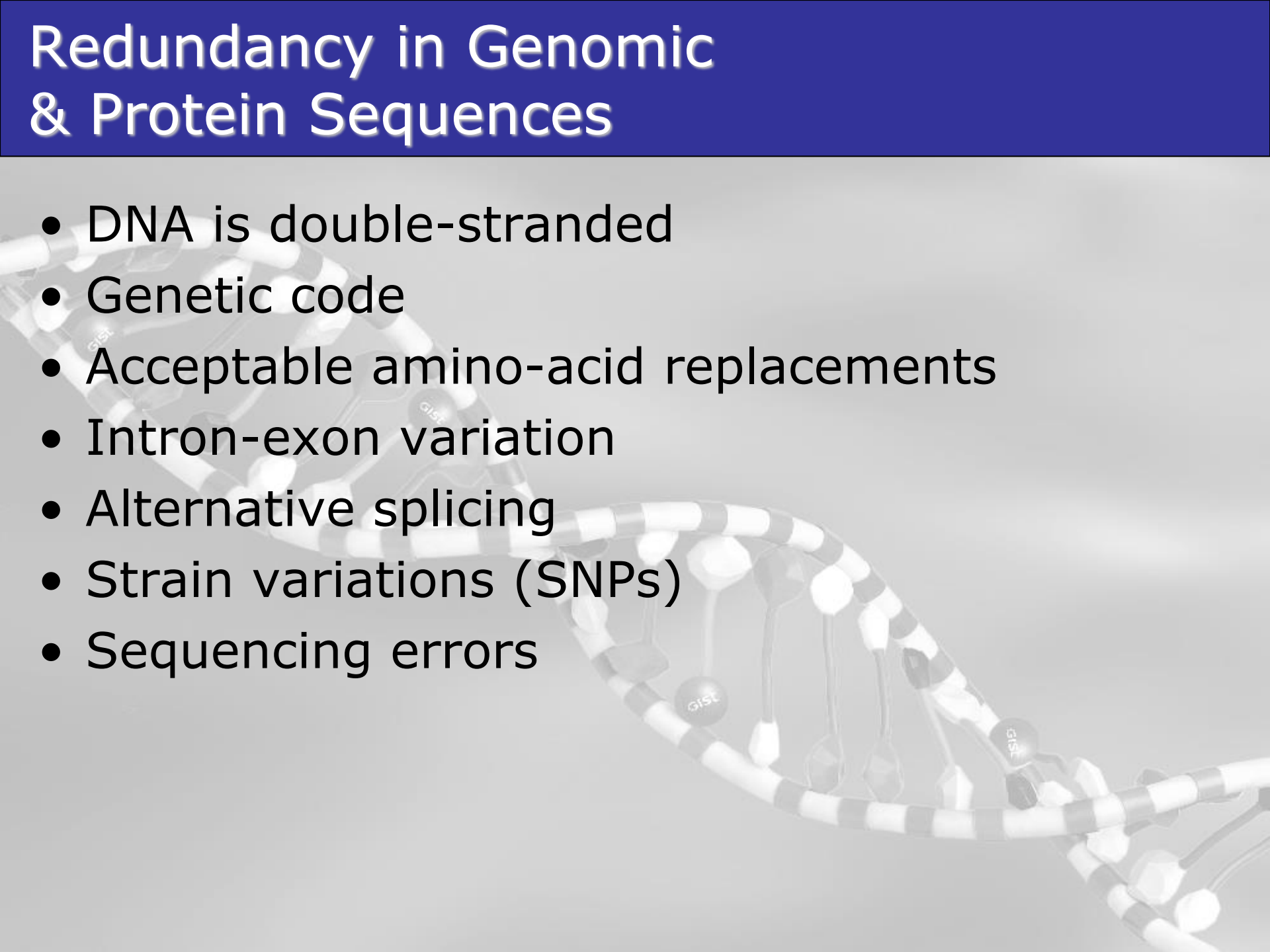
- **Simulate**: Model gene expression, gene regulation, protein folding, protein-protein interaction, protein-ligand binding, catalytic function, metabolism...
 - **Engineer**: Construct novel organisms or novel functions or novel regulation of genes and proteins.
 - **Gene Therapy**: Target specific genes, or mutations, RNAi to change a disease phenotype.
- 

Challenges Understanding Genetic Information



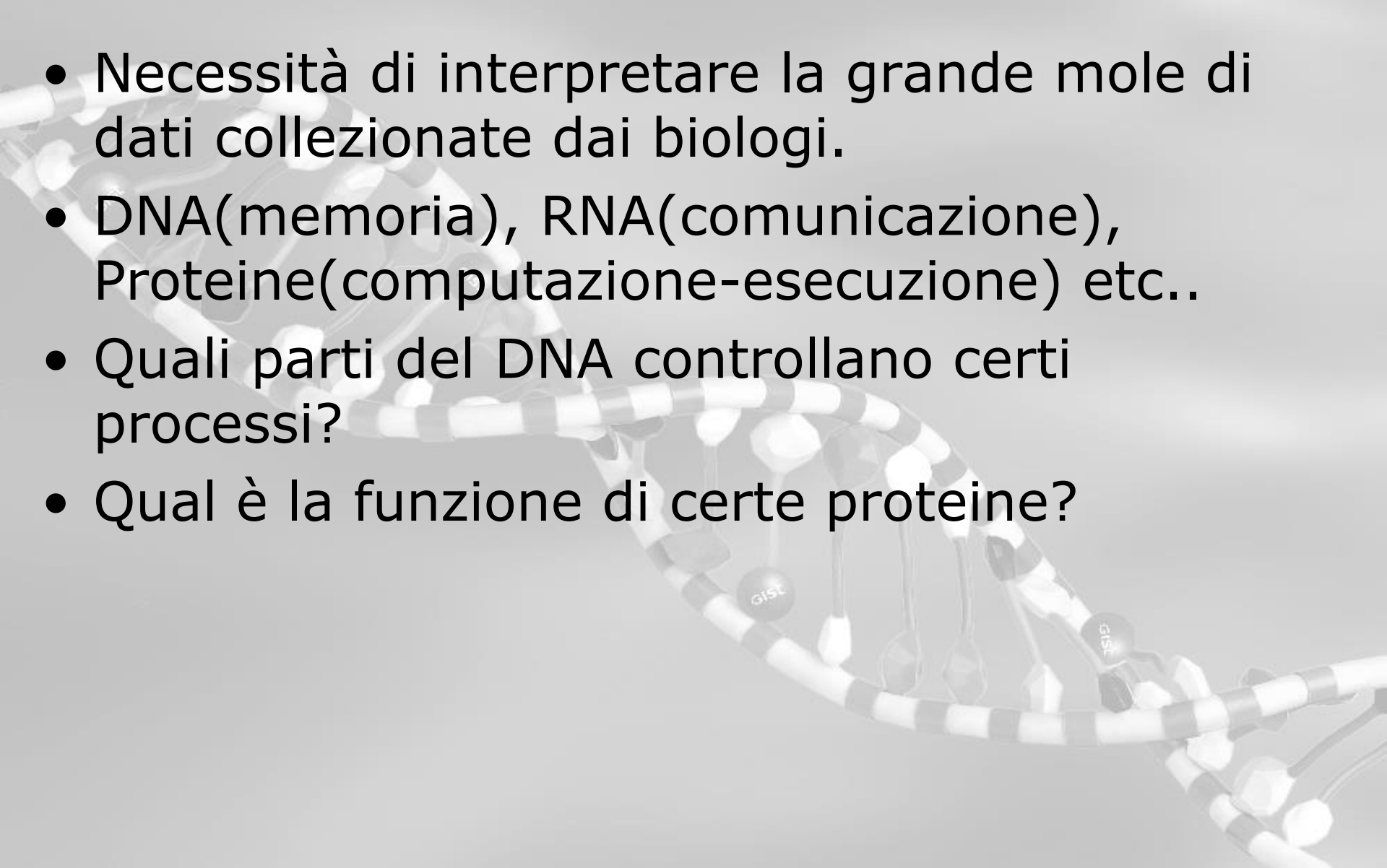
- Genetic information is **redundant**
- Structural information is **redundant**
- Genes and proteins are **meta-stable**
- Single genes have **multiple functions**
- Genes are **one dimensional** but function depends on **three-dimensional** structure

Redundancy in Genomic & Protein Sequences

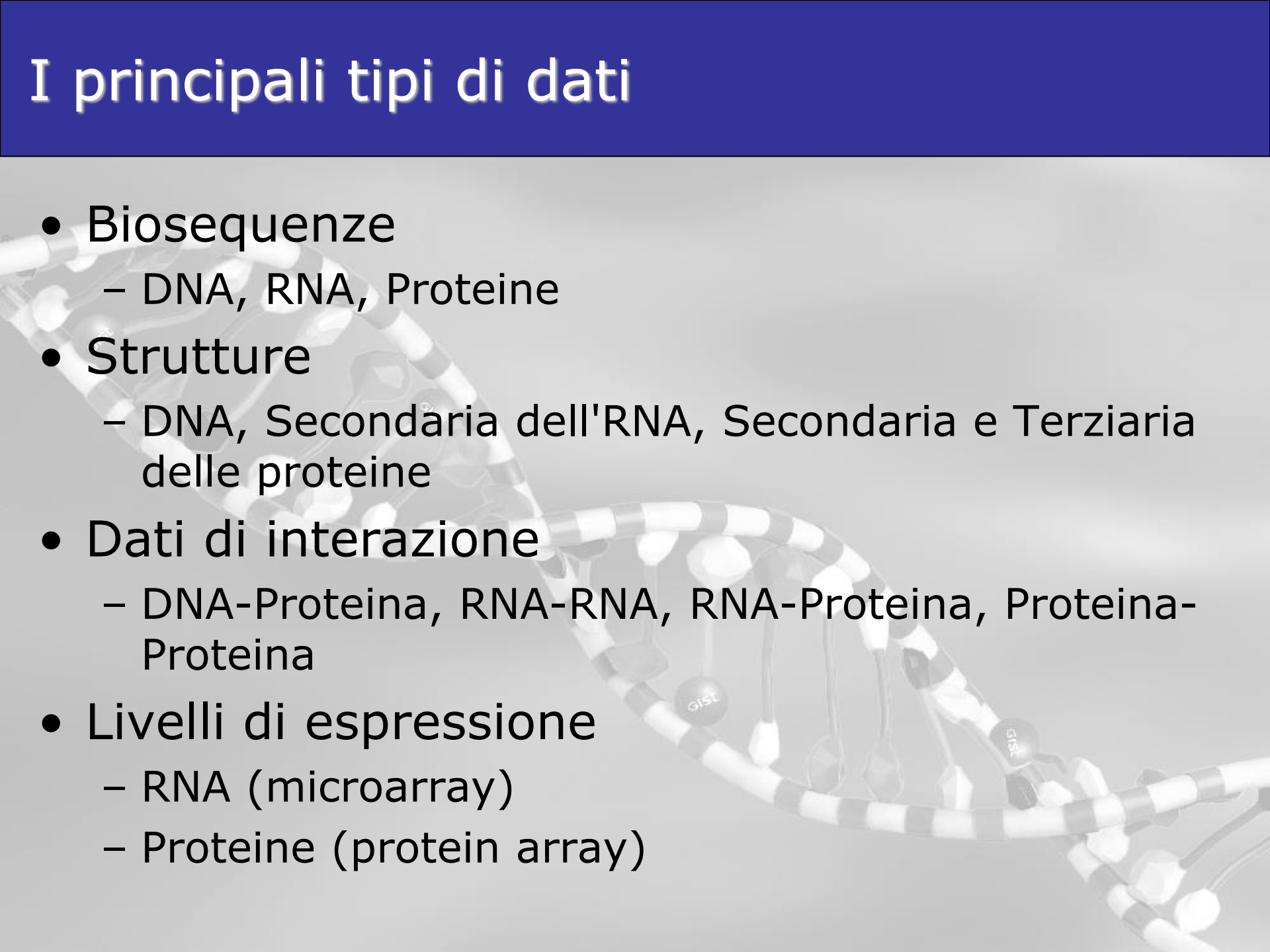
- DNA is double-stranded
 - Genetic code
 - Acceptable amino-acid replacements
 - Intron-exon variation
 - Alternative splicing
 - Strain variations (SNPs)
 - Sequencing errors
- 

La Bioinformatica

- Necessità di interpretare la grande mole di dati collezionate dai biologi.
- DNA(memoria), RNA(comunicazione), Proteine(computazione-esecuzione) etc..
- Quali parti del DNA controllano certi processi?
- Qual è la funzione di certe proteine?



I principali tipi di dati

- Biosequenze
 - DNA, RNA, Proteine
 - Strutture
 - DNA, Secondaria dell'RNA, Secondaria e Terziaria delle proteine
 - Dati di interazione
 - DNA-Proteina, RNA-RNA, RNA-Proteina, Proteina-Proteina
 - Livelli di espressione
 - RNA (microarray)
 - Proteine (protein array)
- 

Esempio 1

- In una sequenza proteica è possibile individuare regioni funzionalmente importanti.
- Ogni sequenza proteica è codificata da una sequenza genomica.
- Supponiamo che la regione X nel moscerino sia cruciale in una certa funzione.
- Domanda: esiste un analogo nell'uomo?
- Risposta: effettuando una ricerca per similarità della regione X nel genoma umano è possibile individuare dei geni candidati.

Esempio 2

- Tutte le cellule di un individuo contengono lo stesso DNA.
- Eppure un neurone è molto diverso da un globulo bianco!
- Che cosa li rende così diversi nella forma e nella funzione?
- Sebbene il DNA sia lo stesso, esso contiene delle regioni importanti in tutte le cellule ed altre specifiche per alcune di esse.
- Mediante un'analisi del trascrittoma (microarray) è possibile stabilire quali regioni del DNA contengono informazioni relative al funzionamento di ognuna delle due cellule.

