|  |  |  |
| --- | --- | --- |
| **Docente** | **Reperibilità** | **Argomento tesi – numero studenti**  |
| Maria Pia Conte | mariapia.conte@uniroma1.it | 1 studente, da gennaio 2019Linee di ricerca:-The pangenome structure of Escherichia coli patotypes: comparative genomic analysis and phylogenetic organisation.-Integrative metagenomic and culturomic approach to study temporal dynamics of skin microbiota inatopic dermatitis-Pangenome analysis of Burkholderia cepacia complex involved in healthcare-associated infections (HAIs) |
| Giuseppe Biondi Zoccai | gbiondizoccai@gmail.com | 1 studenteTitolo orientativo: Integration of preclinical and clinical evidence with umbrella reviews.In pratica il progetto si concentrerebbe sulla possibilità di condurre analisi semiautomatiche che permettano la sintesi quantitative di diverse fonti di dati, dagli studi clinici nell'uomo, agli studi negli animali, a quelli in vitro. |
| Serena Schippa | serena.schippa@uniroma1.it | Uno studente per volta:" Caratterizzazione del batterioma intestinale (fecale o mucosa-adeso) nell’ambito di diversi contesti patologici”. |
| Luigi Faino | luigi.faino@uniroma1.it | 3 studentiLinee diricerca:Nanopore Real-Time sequencing application in plant pathology.Real-Time sequencing is a cutting-edge technology developed by Nanopore to sequence nucleic acids in a fast and portable manner. To date, several applications have been developed to use this technology in fields like health care, molecular biology and ecology. Plant pathology is another field that can benefit by such a technology to speed-up pathogen detection and determination. In our lab, we have currently running three projects that make use of the MinION (the smallest sequencing device to date) from Nanopore:1. Sequencing and assembly of different strains of *Pantoea stewartii* subsp. *stewartii*. This pathogen is a quarantine pathogen of maize but endemic on other cereals. The aim of the project is to assemble these strains and find genomic features that would correlate with the different virulence levels on maize and identify regions that can distinguish *P. stewartii* subsp. *stewartii* isolated on maize and rice.
2. Build a web service that would help researcher to identify batch of maize contaminated by *P. stewartii*. The project aims to the setup of a protocol that would identify infections of *P. stewartii* by PCR and DNA-sequencing. The webtools should give as output the probability of contamination. This project will be conducted in collaboration with another student that will work in the laboratory.
3. Assembly of two strains of *Fusarium verticillioides*. Currently, we are sequencing two strains of *F. verticillioides* by Nanopore. The next step will be the assembly of the data and the polishing of the assembly using Illumina data (we have already in house) to correct assembly error. At the end, we will annotate the genome of both strains using RNA-seq data.
 |
| Maria Elisa Crestoni | mariaelisa.crestoni@uniroma1.it | Uno-due studenti"Analisi statistica e bioinformatica per lo studio di metaboliti in matrici alimentari complesse con il supporto di banche dati dedicate" |
| Walter Adriani | walter.adriani@iss.it | 2 progetti di ricerca:"Ritmi circadiani e fenotipo comportamentale del ratto DAT ko" Su modello animale (ratto) di sintomi riconducibili al disturbo ossessivo-compulsivo, alla ADHD ed alle addictions - anche di natura puramente comportamentale (es. gioco d'azzardo patologico) - si prevede di svolgere letture ed analisi dati ottenuti con sistemi automatizzati, come [p.es](http://p.es). sensori di attività 24h/24 7gg/7 oppure test di comportamento operante a valenza neuro-economica ("gambling tasks" e simili). OPPURE "Sviluppo di una App per implementare il Public Goods Game"Il PGG è un 'gioco' che contempera la propensione a donare risorse economiche per il beneficio di tutti i partecipanti versus la tentazione di trattenere per sè queste stesse risorse (altruismo vs egoismo): si prevede di sviluppare una App online che possa permettere il reclutamento di volontari sani, al fine di monitorare e valutare le loro strategie di gioco.  |
| Valerio Fulci | valerio.fulci@uniroma1.it | 1 studente:Development of a bisulfite RNA-seq analysis pipeline to assess the 5mC profile in mammalian small RNAs.Recently, a great deal of attention has been paid to so called "epitranscriptomic" modifications, i.e. covalent modifications of cell RNA molecules. It has been shown that mammalian mRNAs and non-coding RNAs harbour several modifications such as 1-methyl-Adenosine (m1A), 6-methyl-Adenosine (m6A) and 5-methyl-cytosine (m5C). m5C can be easily detected by bisulfite treatment which results in the conversion of non-methylated C residues into U residues, which are subsequently sequenced as T residues. 5mC residues are not converted. The pipeline we are currently developing aims to focus on the small RNA fraction in which 5mC has been poorly investigated thus far. The pipeline will consist of a dedicated alignment strategy taking advantage of custom PERL/python scripts for parsing and processing of aligned reads. R statistical software will be used to generate publication quality pictures and statistical testing to call significantly differentially 5-methylated C residues across multiple samples. |
| Paola Bertolazzi | paola.bertolazzi@iasi.cnr.it | **PROPOSTE DI TESI SU METODI E STRUMENTI PER L’ANALISI DI DATI BIOLOGICI****Istituto di Analisi dei Sistemi e Informatica****Consiglio Nazionale delle Ricerche****Gruppo di Bioinformatica e Systems Biology Roma**Negli ultimi tre decenni la ricerca in biologia è stata man mano affiancata da tecniche informatiche che sono state determinanti per la comprensione del materiale e dei fenomeni biologici.Lo IASI lavora su molti di questi argomenti sin dal 2005. Presso lo IASI posso essere effettuate tesi sui seguenti argomenti:1. Studio e caratterizzazione di reti biologiche statiche
2. Studio e progettazione di modelli per la simulazione di processi dinamici in reti biologiche
3. Studio di nuovi metodi e applicazioni di metodi noti per l’analisi di dati, da data set di letteratura o prodotti da laboratori di ricerca, fra cui dati prodotti da sequenziatori di nuova generazione (NGS)
4. Metodi e tecnologie per l’integrazione di dati, anche distribuiti
5. Metodi e tecnologie per il trattamento di Big Data

Si propongono le seguenti tesiARGOMENTO 1Nuovi metodi per la identificazione di Geni Differenzialmente Espressi (DEG)Sperimentazione e valutazione di nuovi metodi per la ricerca di DEG con un comportamento uniforme nelle due classi di esperimenti. La tesi riguarda l’analisi, attraverso tecniche di Gene set enrichment analysis (GSEA), dei risultati ottenuti dall’applicazione di diversi metodi per il calcolo dei DEG su data set estratti da banche dati pubbliche (GEO, TCGA). Uno dei metodi è un’euristica per la soluzione del problema di feature selection con tecniche combinatorie, un altro è un’evoluzione del classico metodo Limma.ARGOMENTO 2Metodi e strumenti per l’identificazione di processi biologici patologici La tesi riguarda l’analisi dello stato dell’arte sui metodi per l’individuazione di processi biologici patologici a partire da dati di espressione genica, sia da microarray sia da NGS, e dalla conoscenza della topologia di reti di tipo Protein-Protein Interaction (PPI). |
| Ettore Meccia | ettore.meccia@iss.it | 1. Tempo fa abbiamo fatto il sequenziamento dell¹esoma di 7 tumori delcolon-retto in pazienti affetti da MAP (MUTYH associated polyposis) inpratica persone che hanno una mutazione ad una proteina che rimuove unaparticolare forma di danno al DNA. I sequenziamenti sono stati giàanalizzati ed alcuni dati pubblicati, ma ci sarebbe sicuramente qualchealtro da ³tirar fuori² da quei dati.2. Abbiamo fatto l¹analisi del metiloma sul sangue del cordone di 128bambini di una coorte polacca, nell¹ambito di un progetto europeo FP7. Idati sono in corso di analisi, ma sicuramente ci sarebbe molto da fare inuno stage di due mesi. |
| Piera Valenti,Maria Pia Conte eMassimiliano Marazzato | piera.valenti@uniroma1.itmariapia.conte@uniroma1.itmassimiliano.marazzato@uniroma1.it | **Comparative genomic analysis of strains belonging to the Burkholderia cepacia complex (Bcc) and determination of markers for fast and accurate identification of Bcc species.**The study aims to clarify the pangenome structure of bacterial species belonging to the Burkholderia cepacia complex (Bcc) as well as to identify suitable genetic markers for the rapid identification of Bcc species. For this purpose, the genome of clinical Bcc strains will be sequenced by next-generation sequencing and assembled by using specific software packages (Velvet, SSPACE, SOAPdenovo, Medusa, GapFiller). The newly assembled genomes will be functionally annotated and compared with other publicly available genomes in order to perform a pangenomic analysis by using different pipelines (BPGA, PGAP). Genetic markers, as iron transport systems, suitable for pathogenicity or other markers appropriate to identify Bcc species will be investigated and further evaluated by *in-silico* analysis.**Genotypic characterization of persistent Uropathogenic Escherichia coli (UPEC) strains.**The aim of this study is to genotypically characterize Uropathogenic Escherichia coli (UPEC) strains able to persist intracellularly in the human urinary tract. For this purpose, the clinical origin UPEC strains EC73 isolated from a subject with recurrent urinary tract infection will be sequenced and assembled by using specific software packages (Velvet, SSPACE, SOAPdenovo, Medusa, GapFiller). The newly assembled genome will be functionally annotated and compared with other publicly available genomes relative to persistent UPEC strains and other Escherichia coli strains isolated from acute urinary infections.**Analysis of the skin microbial community in patients with atopic dermatitis. Identification of keystone microorganisms involved in health and disease.**The project on patients suffering from atopic dermatitis has an overall aim consisting in analyze the composition of the dominant skin-associated microbiota through metagenomics and culturomics approaches. Analyses of microbiome sequencing data from patients compared with healthy controls will be used to generate hypotheses about putative disease-causing microorganisms. All microbiological and molecular data will be correlated to different phases of the disease (active, remission). |
| Matteo Pallocca | matteo.pallocca@gmail.com |

|  |  |
| --- | --- |
| Project 1. A coverage check-package for clinical-based Whole Exome Sequencing | **Background**Wide-target genomic NGS experiments are becoming a de facto clinical tool for some specific situations, e.g. when the histology of the tumor is unknown and/or when there’s need to unveil possible alternative drugs for the patient.**Task**Given a BAM file of a WES experiment and a specific list of genes of interest (e.g. Actionable genes), provide a minimum/median coverage for each one and represent it in a graphical way. This process is extremely computation-intensive so it requires some optimization. **Tools**Bash, Python/R, Web technologies (extra) |
| Project 2. Automated Ontologies (AUTO-GO) | **Background**After a typical RNA-seq experiment, we want to annotate for Ontology and functional enrichment a list of genes. This is a process extremely useful but repetive that can use some automatization. **Task**We did develop our own automated package for ontology-functional annotation of RNA-seq DE genes list, Auto-Go. The student should expand this package with other functional classes, create automated plots, impute a “transcriptomic background” in order to normalize results.**Tools**Python, Bash, Web technologies (extra). |
| Project 3. DRIP-seq | **Background****DRIP-seq (DRIP-sequencing)** is a technology for genome-wide profiling of a type of DNA-RNA hybrid called an "[R-loop](https://en.wikipedia.org/wiki/R-loop)".[[1]](https://en.wikipedia.org/wiki/DRIP-seq) DRIP-seq utilizes a sequence-independent but structure-specific [antibody](https://en.wikipedia.org/wiki/Antibody) for DNA-RNA [immunoprecipitation](https://en.wikipedia.org/wiki/Immunoprecipitation) (DRIP) to capture R-loops for massively parallel [DNA sequencing](https://en.wikipedia.org/wiki/DNA_sequencing). (wiki)**Task**Develop and test a simple pipeline for DRIP-seq data analysis on public and in-house data (QC, align, peak calling). Compare and discuss the R-loop signal with other epigenomic profiles (e.g. ChIP-seq/ATAC).**Tools**Python, Bash, Web technologies (extra). |
| Project 4. Automated Variant Annotation | **Background**Variant/Mutation annotation is a process that involves the integration of several databases, for clinical significance (e..g ClinVar), somatic tumor evidence (COSMIC), pathogenicity prediction (e.g PolyPhen). Can be cumbersome and, as usual, could use some automatization. **Task**Starting from Basic Scripts calling the ANNOVAR package, create a web page that takes as an input a VCF, calls Annotation with ANNOVAR and displays a series of plots or statistics, and enables the download of the output file in TSV/XLSX format **Tools**Web technologies, Bash, Python.  |

 |
|  |  |  |
|  |  |  |
|  |  |  |