



AL MAGNIFICO RETTORE
DELL'UNIVERSITA' DEGLI STUDI DI MILANO

COD. ID: 4318

Il sottoscritto chiede di essere ammesso a partecipare alla selezione pubblica, per titoli ed esami, per il conferimento di un assegno di ricerca presso il Dipartimento di Oncologia ed Emato-Oncologia, responsabile scientifico il **Prof. Giuseppe Testa**

DANIELE CAPOCEFALO
CURRICULUM VITAE

INFORMAZIONI PERSONALI

Cognome	Capocefalo
Nome	Daniele
Data Di Nascita	7 Maggio 1987

OCCUPAZIONE ATTUALE

Incarico	Struttura

ISTRUZIONE E FORMAZIONE

Titolo	Corso di studi	Università	anno conseguimento titolo
Laurea Magistrale o equivalente	Bioinformatica	Università degli studi di Roma "Tor Vergata"	2014
Dottorato Di Ricerca	Scienze della Vita	Università degli studi di Roma "La Sapienza"	2019
Altro			

ISCRIZIONE AD ORDINI PROFESSIONALI

Data iscrizione	Ordine	Città

LINGUE STRANIERE CONOSCIUTE

lingue	livello di conoscenza
Inglese	C1
Francese	B2



PREMI, RICONOSCIMENTI E BORSE DI STUDIO

anno	Descrizione premio
2017	Borsa "Fondazione Primoli" per organizzazioni di convegni con istituzioni italiane e francesi
2015	Travel Grant per partecipazione a Summer School "Statistics and Computing in Genome Data Science - CSAMA 2015"

ATTIVITÀ DI FORMAZIONE O DI RICERCA

descrizione dell'attività

La mia attività di ricerca si articola nella ricostruzione di reti biologiche partendo da dati *omici* ottenuti da tecnologie high throughput (microarray, Next Generation Sequencing) sia dirette (es: correlazioni basate su profili di espressione genica, vedi <https://doi.org/10.1093/nar/gkw245> e <https://doi.org/10.18632/oncotarget.22184>) che indirette (es: reti di interazione miRNA-RNA differenzialmente espressi in tessuti tumorali e ricostruite attraverso data mining da database pubblici, vedi <https://doi.org/10.18632/oncotarget.9989>). Queste reti sono state da me analizzate al fine di caratterizzarne la topologia (i.e. definire se la rete sia scale-free) e definire le proprietà generali (es: *network density*) e inferire quali siano i nodi più rilevanti (hubs) all'interno della rete tramite misure di centralità locali (*degree*, *clustering coefficient*, *closeness*). Gli approcci basati su network hanno fatto da apripista in diversi ambiti, ad es. per studiare il ruolo dei geni del ritmo circadiano all'interno del cancro del colon retto, o per identificare *master regulators* epigenetici (miRNA) che coordinino a monte l'attivazione o il silenziamento di geni chiave nella progressione tumorale.

Durante il mio progetto di dottorato, ho esteso la caratterizzazione delle reti molecolari utilizzando indici che vengono usati in altri contesti dove la teoria dei grafi si applica per l'analisi di sistemi complessi, come l'ecologia e l'analisi di reti sociali (Social Network Analysis). A tal riguardo, ho studiato e implementato sette indici di centralità di gruppo, che valutano l'impatto che gruppi di nodi, non necessariamente adiacenti) in una rete esercitano nei confronti del flusso di comunicazioni e nella coesione delle reti, altresì definiti key players. Ho racchiuso questi metodi in un tool bioinformatico e open source, Pyntacle <http://pyntacle.css-mendel.it/> (pubblicazione in stesura). Pyntacle è uno strumento da linea di comando e una libreria Python, installabile su diversi sistemi operativi, che racchiude tutta una serie di strumenti e automatizzazioni per effettuare una varietà di analisi di reti, anche non biologiche. Il fulcro di Pyntacle è la valutazione e la ricerca di gruppi di key player di una grandezza definita dall'utente. Il programma contiene due algoritmi di ricerca, uno *greedy* e uno *bruteforce*, che valutano quale siano il o i gruppi di nodi che massimizzano una o più metriche di gruppo. Pyntacle è studiato per analizzare reti di grandi dimensioni fino a 10000 nodi) e implementa una serie di strategie per effettuare la caratterizzazione di gruppi di nodi utilizzando elaborazione in parallelo su più CPU, rendendolo adattato ad essere usato in strutture HPC, oltre che su desktop.

Oltre alla centralità di gruppo, Pyntacle supporta tutta una serie di analisi per la rete di network, comprese la ricerca di gruppi di nodi densamente connessi rispetto al resto della rete (communities), indici di topologia globali e locali utilizzati in diversi contesti e informativi su proprietà specifiche della rete, operazioni booleane tra reti (unione, intersezione, differenza). Lo strumento produce una serie di report in diversi formati e contiene un visualizzatore interno, sviluppato in JavaScript, che permette l'esplorazione visiva delle reti. Infine, Pyntacle supporta una serie di formati di file standard per la rappresentazione di network e permette la rapida conversione da un formato all'altro, per proseguire l'analisi delle reti su altri strumenti.

Pyntacle è stato dapprima testato su reti trofiche in ambito ecologico per identificare proprietà strutturali di ecosistemi marini e definire quali sono le catene alimentari più soggette a stress esogeni (es: pesca non controllata, cambiamenti climatici), vedi <https://doi.org/10.1155/2018/1979214>.

In studi di biologia molecolare, ho impiegato Pyntacle per studiare la dinamicità del trascrittoma di *Drosophila melanogaster*, dallo sviluppo embrionale alla maturazione adulta, attraverso la ricostruzione di reti di correlazione pesate ottenute attraverso *Weighted Gene Co-Expression Analysis* (WGCNA)



ricostruite a partire da dati pubblici (trascrittomici). Il fine ultimo del progetto è quello di definire se esistono gruppi di geni all'interno del trascrittoma che coordinino a monte importanti processi legati all'invecchiamento, e di caratterizzare il loro ruolo durante lo sviluppo del moscerino.

Nella mia esperienza di ricerca ho inoltre fornito supporto a diverse unità di ricerca sperimentale nell'ambito della caratterizzazione di diversi tumori (cancro del colon retto, tumore al pancreas) e malattie rare (es: Diabete di tipo 2 multigenerazionale). Ho personalmente implementato pipeline per l'analisi automatica di dati provenienti da genomici (*Whole Exome Sequencing*, Pannelli genici) e trascrittomici (RNAseq) ed effettuato analisi ad hoc per la caratterizzazione dei dataset. Il successo delle mie attività di analisi in questo ambito è riconosciuto da diverse pubblicazioni su riviste scientifiche internazionali, ad es.

ATTIVITÀ PROGETTUALE

Anno	Progetto
2014 - 2016	<p>Studio di reti di interazione mRNA-miRNA in cancro coloretale con approcci di teoria dei grafi partendo da dati <i>omici</i> (microarrays)</p> <p>Pubblicazioni:</p> <ul style="list-style-type: none">Mazza, T., Mazzoccoli, G., Fusilli, C., Capocefalo, D., Panza, A., Biagini, T., Castellana, S., Gentile, A., De Cata, A., Palumbo, O., et al. (2016). Multifaceted enrichment analysis of RNA-RNA crosstalk reveals cooperating micro-societies in human colorectal cancer. <i>Nucleic Acids Res.</i> 44, 4025-4036. https://doi.org/10.1093/nar/gkw245Mazzoccoli, G., Colangelo, T., Panza, A., Rubino, R., Tiberio, C., Palumbo, O., Carella, M., Trombetta, D., Gentile, A., Tavano, F., et al. (2016). Analysis of clock gene-miRNA correlation networks reveals candidate drivers in colorectal cancer. <i>Oncotarget</i> 7, 45444-45461. https://doi.org/10.18632/oncotarget.9989
2016-2019	<p>Analisi di oscillazione ultradiane in reti circadiane per capire presenza e grado di conservazione "circadian like" a partire da dati pubblici di espressione su serie temporali (time series).</p> <p>Pubblicazioni:</p> <p>Castellana, S., Mazza, T., Capocefalo, D., Genov, N., Biagini, T., Fusilli, C., Scholkmann, F., Relógio, A., Hogenesch, J.B., and Mazzoccoli, G. (2018). Systematic Analysis of Mouse Genome Reveals Distinct Evolutionary and Functional Properties Among Circadian and Ultradian Genes. <i>Front. Physiol.</i> 9, 1178. https://doi.org/10.3389/fphys.2018.01178</p> <p>Genov, N., Castellana, S., Scholkmann, F., Capocefalo, D., Truglio, M., Rosati, J., Turco, E.M., Biagini, T., Carbone, A., Mazza, T., Mazzoccoli G., Relógio, A..A multi-layered study on harmonic oscillations in mammalian genomics and proteomics IJMS (in revisione)</p>
2015-2019	<p>Network Approaches for the characterization of Ageing in <i>Drosophila</i>. Il progetto, svolto in collaborazione con l'Universidad Mayor del Chile, prevede l'utilizzo di reti basate su correlazioni pesate su dati pubblici di trascrittomici che catturano i profili di espressione di <i>D. melanogaster</i> dallo sviluppo embrionale fino alla maturazione adulta. Scopo del progetto è di identificare gruppi di geni nelle reti di co-espressione che hanno pattern ricorrenti all'interno delle reti. Inoltre, scopo del progetto è quello di caratterizzare i gruppi di geni (comunità) associate ai processi di invecchiamento e senescenza e di tracciare i loro partner di co-espressione dall'embrione fino all'età adulta.</p> <p>Pubblicazioni:</p> <p>Caris-Maldonado. J.C., Capocefalo D., Lopez-Quilodran N., Molina-Fernandez C., Arias-Carrasco R., Sepúlveda-Hermosilla G., Slater A., Martinez P., van Zundert B., Vargas R.,</p>



	Mazza T., Maracaja-Coutinho V., Tevy, M.F. The dynamic aging transcriptome of <i>Drosophila Melanogaster</i> (in elaborazione)
2016-2019	<p>Pyntacle: elaborazione di uno strumento open-source per la caratterizzazione di reti biologiche utilizzando approcci di social network analisi ed ecologia.</p> <p>Pubblicazioni:</p> <p>Capocefalo, D., Truglio, M., Ferénc, J., Mazza, T. Unravelling team play effects with Pyntacle, (in elaborazione)</p> <p>Sito Ufficiale:</p> <p>http://pyntacle.css-mendel.it/ (Sito Principale)</p> <p>https://github.com/mazzalab/pyntacle (Repository Git)</p> <p>https://hub.docker.com/u/mazzalab (Immagine Docker)</p>
2015-2016	<p>Caratterizzazione e prioritizzazione di polimorfismi caratterizzanti il Diabete di Tipo 2 multigenerazionale.</p> <p>Pubblicazioni:</p> <p>Pezzilli, S., Ludovico, O., Biagini, T., Mercuri, L., Alberico, F., Lauricella, E., Dallali, H., Capocefalo, D., Carella, M., Miccinilli, E., et al. (2018). Insights From Molecular Characterization of Adult Patients of Families With Multigenerational Diabetes. <i>Diabetes</i> 67, 137-145. https://doi.org/10.2337/db17-0867</p>
2015-2017	<p>APOGEE: un nuovo indice per determinare l'impatto di varianti a singolo nucleotide mitocondriali basato su Machine Learning.</p> <p>Pubblicazioni:</p> <p>Castellana, S., Fusilli, C., Mazzoccoli, G., Biagini, T., Capocefalo, D., Carella, M., Vescovi, A.L., and Mazza, T. (2017). High-confidence assessment of functional impact of human mitochondrial non-synonymous genome variations by APOGEE. <i>PLoS Comput. Biol.</i> 13, e1005628. https://doi.org/10.1371/journal.pcbi.1005628</p> <p>Sito Ufficiale:</p> <p>http://mitimpact.css-mendel.it/</p>
2014-2016	<p>Benchmarking di diversi strumenti di Dinamica Molecolare per determinarne i maggiori costi/benefici in termini di risorse hardware e risorse computazionali per piattaforme High Performance Computing (HPC).</p> <p>Pubblicazioni:</p> <p>Biagini, T., Chillemi, G., Mazzoccoli, G., Grottesi, A., Fusilli, C., Capocefalo, D., Castellana, S., Vescovi, A.L., and Mazza, T. (2016). Molecular dynamics recipes for genome research. <i>Brief. Bioinform.</i> 19, 853-862. https://doi.org/10.1093/bib/bbx006</p>

TITOLARITÀ DI BREVETTI

Brevetto

CONGRESSI, CONVEGNI E SEMINARI

Data	Titolo	Sede
------	--------	------



27-29 Giugno 2018	15th Annual Meeting of the Bioinformatics Italian Society	Università degli studi di Torino, IT
9-14 Luglio 2017	In Silico Systems Biology	European Bioinformatics Institute (EBI), Hinxton, Cambridge, UK
15-17 Luglio 2016	13th Annual Meeting of the Bioinformatics Italian Society	Università degli Studi di Salerno, IT
4-6 Febbraio 2016	High Performance Molecular Dynamics	CINECA, Via dei Tizii 6B, Roma, IT

PUBBLICAZIONI

Libri

Tesi di Dottorato: "Reverse Engineering of Natural Systems by Graph Theory", Capocéfalo, D., Università degli studi di Roma "La Sapienza", Gennaio 2019
<https://iris.uniroma1.it/handle/11573/1271802#.XXaEmigzY2w>

Articoli su riviste

Biagini, T., Petrizzelli, F., Truglio, M., Cespa, R., Barbieri, A., **Capocéfalo, D.**, Castellana, S., Tevy, M.F., Carella, M., and Mazza, T. (2019). Are Gaming-Enabled Graphic Processing Unit Cards Convenient for Molecular Dynamics Simulation? *Evol. Bioinform. Online* 15, 1176934319850144.

<https://doi.org/10.1177/1176934319850144>

Capocéfalo, D., Pereira, J., Mazza, T., and Jordán, F. (2018). Food Web Topology and Nested Keystone Species Complexes. *Complexity* 2018.

<https://doi.org/10.1155/2018/1979214>

Castellana, S., Mazza, T., **Capocéfalo, D.**, Genov, N., Biagini, T., Fusilli, C., Scholkmann, F., Relógio, A., Hogenesch, J.B., and Mazzoccoli, G. (2018). Systematic Analysis of Mouse Genome Reveals Distinct Evolutionary and Functional Properties Among Circadian and Ultradian Genes. *Front. Physiol.* 9, 1178.

<https://doi.org/10.3389/fphys.2018.01178>

Castellana, S., Fusilli, C., Mazzoccoli, G., Biagini, T., **Capocéfalo, D.**, Carella, M., Vescovi, A.L., and Mazza, T. (2017). High-confidence assessment of functional impact of human mitochondrial non-synonymous genome variations by APOGEE. *PLoS Comput. Biol.* 13, e1005628.

<https://doi.org/10.1371/journal.pcbi.1005628>

Pezzilli, S., Ludovico, O., Biagini, T., Mercuri, L., Alberico, F., Lauricella, E., Dallali, H., **Capocéfalo, D.**, Carella, M., Miccinilli, E., et al. (2018). Insights From Molecular Characterization of Adult Patients of Families With Multigenerational Diabetes. *Diabetes* 67, 137-145.

<https://doi.org/10.2337/db17-0867>

Mazza, T., Copetti, M., **Capocéfalo, D.**, Fusilli, C., Biagini, T., Carella, M., De Bonis, A., Mastrodonato, N., Piepoli, A., Paziienza, V., et al. (2017). MicroRNA co-expression networks exhibit increased complexity in pancreatic ductal compared to Vater's papilla adenocarcinoma. *Oncotarget* 8, 105320-105339.

<https://doi.org/10.18632/oncotarget.22184>

Mazzoccoli, G., Castellana, S., Carella, M., Palumbo, O., Tiberio, C., Fusilli, C., **Capocéfalo, D.**, Biagini, T., Mazza, T., and Lo Muzio, L. (2017). A primary tumor gene expression signature identifies a crucial role played by tumor stroma myofibroblasts in lymph node involvement in oral squamous cell carcinoma.



Oncotarget 8, 104913-104927. https://doi.org/10.18632/oncotarget.20645
Biagini, T., Chillemi, G., Mazzoccoli, G., Grottesi, A., Fusilli, C., Capocefalo, D. , Castellana, S., Vescovi, A.L., and Mazza, T. (2016). Molecular dynamics recipes for genome research. <i>Brief. Bioinform.</i> 19, 853-862. https://doi.org/10.1093/bib/bbx006
Mazza, T., Mazzoccoli, G., Fusilli, C., Capocefalo, D. , Panza, A., Biagini, T., Castellana, S., Gentile, A., De Cata, A., Palumbo, O., et al. (2016). Multifaceted enrichment analysis of RNA-RNA crosstalk reveals cooperating micro-societies in human colorectal cancer. <i>Nucleic Acids Res.</i> 44, 4025-4036. https://doi.org/10.1093/nar/gkw245
Mazzoccoli, G., Colangelo, T., Panza, A., Rubino, R., Tiberio, C., Palumbo, O., Carella, M., Trombetta, D., Gentile, A., Tavano, F., et al. (2016). Analysis of clock gene-miRNA correlation networks reveals candidate drivers in colorectal cancer. <i>Oncotarget</i> 7, 45444-45461. https://doi.org/10.18632/oncotarget.9989

Atti di convegni
Poster: "Unraveling the complexity of molecular networks with Pyntacle" Presentato alla "17 European Conference on Computational Biology (ECCB)", Atene, Grecia, Settembre 2018
Poster: "Dedalus: A Python package for mesoscale analysis of biological network" Presentato alla Summer School "In Silico Systems Biology", European Bioinformatics Institute (EBI), Hinxton, Cambridge, UK, Luglio 2017
Poster: "Pyntacle-A high-performance framework for the analysis team-play effects in biological networks" Presentato al 15esimo raduno della Società Italiana di Bioinformatica (BITS), Torino, IT, Giugno 2016
Talk: "Exploring the development and aging-dependent patterns in <i>D.melanogaster</i> through network-based Approaches" Presentato al seminario: "The Pendulum of Life: a multidisciplinary workshop on the role of oscillatory patterns in post-contemporary biology", tenuto presso Istituto MENDEL, Roma, IT, 11 Maggio 2017

ALTRE INFORMAZIONI

Le dichiarazioni rese nel presente curriculum sono da ritenersi rilasciate ai sensi degli artt. 46 e 47 del DPR n. 445/2000.

Il presente curriculum non contiene dati sensibili e dati giudiziari di cui all'art. 4, comma 1, lettere d) ed e) del D.Lgs. 30.6.2003 n. 196.

Luogo e data: Roma, 10/09/2019

FIRMA