BIOGRAPHICAL SKETCH

NAME: Francesco Ferrari

eRA COMMONS USER NAME: FRANCESCO.FERRARI

POSITION TITLE: Principal Investigator,
Computational Genomics Laboratory,
IFOM, the FIRC Institute of Molecular Oncology

EDUCATION/TRAINING

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Completion Date MM/YYYY	FIELD OF STUDY
University of Modena and Reggio Emilia, Modena, Italy	Ms	09/2004	Medical Biotechnologies
University of Modena and Reggio Emilia, Modena, Italy	PhD	02/2008	Biotechnology and Molecular Medicine
University of Padova, Padova, Italy	Postdoctoral res. fellow	04/2010	Computational biology
Harvard Medical School, Boston, MA	Postdoctoral res. assoc.	11/2014	Computational biology

A. Personal Statement

I established my research group as independent PI at IFOM in Milan in January 2015. My group is interested in studying distal regulatory elements (enhancers) and their epigenetic or genetic alterations in genetics diseases and cancer. Our expertise is particularly focused on the use of 3D chromatin organization data obtained by Hi-C and other high throughput techniques derived from chromosome conformation capture (3C). We also use functional genomics data, mainly derived from transcriptomics and epigenomics techniques based on next generation sequencing (NGS). We adopt these omics data to gain mechanistic insights into transcription regulation. In particular, we characterize the role of enhancers in epigenetics and gene expression regulation, within the broader gene regulatory network. My group is currently composed only of bioinformaticians, but we have plans (and funding support) to start wet lab experiments as well.

B. Positions and Honors

2005-2007 – University of Modena and Reggio Emilia, Italy
Graduate student in the Biotechnology and Molecular Medicine program.
Supervisor prof. Sergio Ferrari

January 2008-April 2010 – University of Padova, Italy
Postdoctoral research fellow.
Supervisors prof. Silvio Bicciato and Stefania Bortoluzzi.

May 2010 - November 2014 – Harvard Medical School, Boston MA, USA
Postdoctoral research associate at the Center for Biomedical Informatics.
Supervisor prof. Peter J. Park.

January 2015 - present – IFOM, the FIRC Institute of Molecular Oncology, Milan, Italy
Principal Investigator, Computational Genomics Laboratory.

OTHER RESEARCH EXPERIENCES

November-December 2007 – Weizmann Institute of Science, Rehovot, Israel
Visiting graduate student – Sergio Lombroso program for cancer research fellowship.

September-October 2009 – Leiden University Medical Center, Leiden, The Netherlands
Visiting postdoctoral fellow – HPC Europa transnational access grant.

FELLOWSHIPS

2005-2007 – Doctoral degree fellowship of the Italian Ministry of Research
2008-2010 – Postdoctoral fellowship ("assegno di ricerca") of the University of Padova

SHORT TERM FELLOWSHIPS

2007 - Visiting PhD student Fellowship of Sergio Lombroso Foundation for Cancer Research
2009 - HPC Europe transnational access program grant.

C. Contributions to Science

1) Chromatin architecture, epigenetics and transcriptional regulation of pluripotency

As part of my postdoctoral work, in collaboration with Konrad Hochedlinger's group (MGH and Harvard Stem Cell Institute), I've studied 3D-organization of the genome around a key pluripotency gene (*Nanog*) during reprogramming of somatic cells to induced pluripotent stem cells *(Apostolou*, Ferrari* et al., Cell Stem Cell, 2013)*. We used a modified 4C (Circular Chromosome Conformation Capture) protocol coupled with deep sequencing (m4C-seq) to identify genome-wide chromatin interactions of the *Nanog* locus. We identified a pluripotency-specific interactome for *Nanog*, which is lost upon differentiation, and re-established during reprogramming. We also found that the genomic rearrangements precede gene expression changes associated to pluripotency, thus being pivotal for cell identity change. As later confirmed also by others the re-organization of chromatin is a key molecular event associated to cell identity transition *(Ferrari*, Apostolou* et al., Cell Cycle 2014)*.

I have collaborated with Konrad Hochedlinger's lab (MGH and Harvard Stem Cell Institute) for other projects focusing on the epigenetics and transcriptional events associated to reprogramming of somatic cells to pluripotent stem cells (*Cheloufi et al., Nature 2015; Choi et al., Nature Biotechnology 2015; Stadtfeld et al., Nature Genetics 2012*).

Also connected to this line of research, in collaboration with George Daley's group (Harvard Medical School), I defined computational genomics approaches to verify the origin of pluripotent stem cells *(De Los Angeles*, Ferrari* et al., Nature, 2015)*. This addressed a relevant problem in the field of stem cells research as a number of controversial claims have arisen during the years.

2) Epigenetics regulation of transcription on whole chromosomes

During my postdoc, in collaboration with Mitzi Kuroda's group (Harvard Medical School) I've investigated the epigenetic regulation of transcription for compensating copy number differences between sex chromosomes in *Drosophila melanogaster*. Dosage compensation is an interesting model of coordinated epigenetic regulation of an entire chromosome. Different species have evolved distinct mechanisms, yet all are based on various level of epigenetic regulation, eventually resulting in differential transcriptional activity of sex chromosomes genes, as reviewed in (*Ferrari et al., Nat Struct Mol Biol, 2014*). In *D. melanogaster* the single copy of X in male is up-regulated, yet it has been debated which stages of the RNA Polymerase II (Pol II) transcription cycle are involved affected in the boost of gene expression associated to dosage compensation: Pol II recruitment, initiation, pausing release or elongation. Indeed we challenged a previously proposed model suggesting that the two-fold up-regulation of chromosome X is achieved at the initiation stage (*Ferrari*, Jung* et al., Science 2013*). In our work we combined transcription dynamics data obtained from several genome-wide techniques, to gather information on distinct stages of transcription (*Ferrari*, Plachetka*, Alekseyenko* et al., Cell Reports, 2013*). Specifically we used proprietary data on nascent RNAs sequencing, along with public datasets for two Global Run-On sequencing (GRO-seq) protocol variants (to map paused and actively elongating Pol II), 5'-short RNAs sequencing (a proxy for 5'-end proximal paused Pol II) and Chromatin Immunoprecipitation coupled with sequencing (ChIP-seq) of Pol II with Serine 2phosphorylation (marking elongating Pol II). By combining information from these different genome wide techniques, we were able to

support a model where a combination of enhanced pausing release and facilitated elongation are involved in increased transcription of X-linked genes in males fruitflies.
Connected to this same line of research, more recently as part of my work as a PI I have been investigating the role of chromatin 3D organization in the regulation of sex chromosomes dosage compensation (Pal et al., manuscript in preparation).

3) Epigenetics alterations in diseases
During my postdoc I collaborated also with Marcy MacDonald's lab (MGH and Harvard Medical School) for the characterization of epigenetics alterations associated to Huntington disease (*Biagioli*, Ferrari* et al., Human Molecular Genetics, 2015*). In relation to this line of research I also worked on other collaborative projects in the field of cancer research, including a work on chromatin modifiers role in pancreatic ductal adenocarcinoma (PDAC) (a cancer type with extremely bad prognosis). In this work I integrated genome-wide gene expression and histone marks profiles from proprietary and public datasets, which was instrumental to define a network of epigenetic regulation involving KDM2B, KDM5A and PRC2 to drive the pathogenesis of an aggressive subset of PDAC (*Tzatsos et al., J. Clin. Invest. 2013*).
Related to this line of research I have now additional projects ongoing in my own group as a Principal investigator. In particular we are studying how the enhancer-gene regulatory network is altered in cancer (e.g. by non-coding mutations in enhancers) and in genetic diseases affecting chromatin modifiers (e.g. in Kabuki Syndrome, along with local collaborators)

4) Genomics determinants of gene expression regulation and bioinformatic tools development
As a graduate student I investigated the relationship between gene expression and genomic position in myeloid cells differentiation, by identifying genomic clusters of co-regulated genes (*Ferrari et al., BMC Genomics, 2007*). An important part of my work has been focused also on the development of bioinformatics tools (*Ferrari et al., BMC Bioinformatics, 2007*). In particular after my PhD I extended my previous work to produce a novel bioinformatics framework integrating sequence, structural and functional data to discover co-expressed, co-regulated and co-localized gene modules (*Coppe*, Ferrari* et al., Nucleic Acids Res, 2009*). In the same line of research, I also developed a new tool for integrative analysis of gene expression, copy number and potentially other functional genomics data to identify variation over large genomic regions (*Ferrari et al., Bioinformatics, 2011*). This tool had more than 1000 downloads (unique IPs) from Bioconductor over the past year (2016).
More recently as part of my work as a PI I've been supervising a collaboration to develop a web tool to implement this data analysis approach as a web tool that can be applied to bacterial genomics data (Puccio et al., manuscript submitted).
In terms of development of bioinformatics methods and tools, a large part of my group focus is currently on the development of methods for chromatin 3D organization data analysis, in particular data obtained with high-resolution Hi-C (Forcato et al., manuscript submitted).

PUBLICATIONS IN PEER REVIEWED INTERNATIONAL JOURNALS

(work as Principal Investigator)
RESEARCH ARTICLES SUBMITTED/UNDER REVISION

- Forcato M, Nicoletti C, Pal K, Livi CM, **Ferrari F***, Bicciato S*. Comparison of computational methods for the analysis of Hi-C data. *Article submitted, currently undergoing peer-review* (* co-corresponding/co-last authors)
- Puccio S, Grillo G, Liciulli F, Severgnini M, Liuni S, Bicciato S, De Bellis G, **Ferrari F*,** Peano C*. WoPPER: Webserver for Position Related data analysis of gene Expression in Prokaryotes. *Summary page article submitted and approved, invitation to submit full article received, full article manuscript in preparation* (* co-corresponding/co-last authors)

(postdoctoral and graduate research work)

RESEARCH ARTICLES - LEAD AUTHOR

1. De Los Angeles A*, **Ferrari F***, Fujiwara Y, Mathieu R, Lee S, Lee S, Tu H, Ross S, Chou S, Nguyen M, Wu Z, Theunissen TW, Powell BE, Imsoonnthornruksa S, Chen J, Borkent M, Krupalnik V, Lujan E, Wernig M, Hanna JH, Hochedlinger K, Pei D, Jaenisch R, Deng H, Orkin SH, Park PJ, Daley GQ. Failure to Replicate the STAP Cell Phenomenon. *Nature*, 2015 Sep 24;525(7570):E6-9. (* equal contribution)
2. Biagioli M*, **Ferrari F***, Mendenhall EM, Zhang Y, Erdin S, Vijayvargia R, Vallabh SM, Solomos N, Manavalan P, Ragavendran A, Ozsolak F, Lee JM, Talkowski ME, Gusella JF, MacDonald ME, Park PJ, Seong IS. Htt CAG repeat expansion confers pleiotropic gains of mutant huntingtin function in chromatin regulation. *Hum Mol Genet.*, 2015 May 1;24(9):2442-57. Epub 2015 Jan 8. (* equal contribution)
3. **Ferrari F***, Plachetka A*, Alekseyenko AA*, Jung YL, Ozsolak F, Kharchenko PV, Park PJ, Kuroda MI. "Jumpstart and gain" model for dosage compensation in Drosophila based on direct sequencing of nascent transcripts. *Cell Reports*, 2013 Nov 14;5(3):629-636 (* equal contribution)
4. Apostolou E*, **Ferrari F***, Walsh RM, Bar-Nur O, Stadtfeld M, Cheloufi S, Stuart HT, Polo JM, Ohsumi TK, Borowsky ML, Kharchenko PV, Park PJ, Hochedlinger K. Genome-wide interactions of the *Nanog* locus in pluripotency, differentiation and cellular reprogramming. *Cell Stem Cell*, 2013 Jun 6;12(6):699-712. (* equal contribution).
5. **Ferrari F***, Jung YL*, Kharchenko PV, Plachetka A, Alekseyenko AA, Kuroda MI, Park PJ. Comment on "*Drosophila* dosage compensation involves enhanced Pol II recruitment to male X-linked promoters". *Science*, 2013 Apr 19;340(6130):273. (* equal contribution) (technical comment article: i.e. reanalysis challenging previous publication)
6. **Ferrari F**, Solari A, Battaglia C, Bicciato S. PREDA: an R-package to identify regional variations in genomic data. *Bioinformatics.* 2011 Sep 1;27(17):2446-7.
7. Coppe A*, **Ferrari F***, Bisognin A, Danieli GA, Ferrari S, Bicciato S, Bortoluzzi S. Motif discovery in promoters of genes co-localized and co-expressed during myeloid cells differentiation. *Nucleic Acids Res*. 2009; 37(2):533-49. Epub 2008 Dec 5. (*equal contribution).
8. **Ferrari F**, Bortoluzzi S, Coppe A, Sirota A, Safran M, Shmoish M, Ferrari S, Lancet D, Danieli GA, Bicciato S. Novel definition files for human GeneChips based on GeneAnnot. *BMC Bioinformatics*. 2007 Nov 15;8:446
9. **Ferrari F**, Bortoluzzi S, Coppe A, Basso D, Bicciato S, Zini R, Gemelli C, Danieli GA, Ferrari S. Genomic expression during human myelopoiesis. *BMC Genomics*. 2007 Aug 3;8(1):264

REVIEW MATERIAL - LEAD AUTHOR

10. **Ferrari F**, Alekseyenko AA, Park PJ, Kuroda MI. Transcriptional control of a whole chromosome: emerging models for the molecular basis of dosage compensation. (Review article) *Nature Structural and Molecular Biology*, 2014 Feb;21(2):118-25.
11. **Ferrari F***, Apostolou E*, Park PJ, Hochedlinger K. Rearranging the chromatin for pluripotency. *Cell Cycle*, 2014 January 15; 13(2):167-168 Epub 2013, Nov 15 (editorial commentary) (* equal contribution)

OTHER SELECTED ARTICLES

12. Day DS, Zhang B, Stevens SM, **Ferrari F**, Larschan EN, Park PJ, Pu WT. Comprehensive analysis of promoter-proximal RNA polymerase II pausing across mammalian cell types. *Genome Biol.* 2016 Jun 3;17(1):120.
13. Cheloufi S, Elling U, Hopfgartner B, Jung YL, Murn J, Ninova M, Hubmann M, Badeaux AI, Euong Ang C, Tenen D, Wesche DJ, Abazova N, Hogue M, Tasdemir N, Brumbaugh J, Rathert P, Jude J, **Ferrari F**, Blanco A, Fellner M, Wenzel D, Zinner M, Vidal SE, Bell O, Stadtfeld M, Chang HY, Almouzni G, Lowe SW, Rinn J, Wernig M, Aravin A, Shi Y, Park PJ, Penninger JM, Zuber J, Hochedlinger K. The histone chaperone CAF-1 safeguards somatic cell identity. *Nature* 2015 Dec 10;528(7581):218-24.

14. Choi J, Lee S, Mallard W, Clement K, Malagoli-Tagliazucchi G, Lim H, Choi IY, **Ferrari F**, Tsankov AM, Pop R, Lee G, Rinn JL, Meissner A, Park PJ, Hochedlinger K. A comparison of genetically matched cell lines reveals the equivalence of human iPSCs and ESCs. *Nat Biotechnol.* 2015 Nov;33(11):1173-81.

15. De Los Angeles A, **Ferrari F**, Xi R, Fujiwara Y, Benvenisty N, Deng H, Hochedlinger K, Jaenisch R, Lee S, Leitch HG, Lensch MW, Lujan E, Pei D, Rossant J, Wernig M, Park PJ, Daley GQ. Hallmarks of Pluripotency. (Review article) *Nature* 2015 Sep 23;525(7570):469-78.

16. Ho J, Jung Y, Liu T, Alver B, Lee S, Ikegami K, Sohn K, Minoda A, Tolstorukov M, Appert A, Parker S, Gu T, Kundaje A, Riddle N, Bishop E, Egelhofer T, Hu S, Alekseyenko A, Rechtsteiner A, Asker D, Belsky J, Bowman S, Chen Q, Chen R, Day D, Dong Y, Dosé A, Duan X, Epstein C, Ercan S, Feingold E, **Ferrari F**, Garrigues J, Gehlenborg N, Good P, Haseley P, He D, Herrmann M, Hoffman M, Jeffers T, Kharchenko P, Kolasinska-Zwierz P, Kotwaliwale C, Kumar N, Langley S, Larschan E, Latorre I, Libbrecht M, Lin X, Park R, Pazin M, Pham H, Plachetka A, Qin B, Schwartz Y, Shoresh N, Stempor P, Vielle A, Wang C, Whittle C, Xue H, Kingston R, Kim J, Bernstein B, Dernburg A, Pirrotta V, Kuroda M, Noble W, Tullius T, Kellis M, MacAlpine D, Strome S, Elgin S, Liu X, Lieb J, Ahringer J, Karpen G, Park PJ. Comparative analysis of metazoan chromatin organization. *Nature* 2014, Aug 28;512(7515):449-52.

17. Jung YL, Luquette LJ, Ho JW, **Ferrari F**, Tolstorukov M, Minoda A, Issner R, Epstein C, Karpen G, Kuroda MI, Park PJ. Impact of sequencing depth in ChIP-seq experiments. *Nucleic Acids Res*. 2014 May; 42(9):e74, Epub 2014 March 5.

18. Tzatsos A, Paskaleva P*, **Ferrari F***, Deshpande V, Stoykova S, Contino G, Wong K, Lan F, Trojer P, Park PJ, Bardeesy N. KDM2B promotes pancreatic cancer via Polycomb-dependent and -independent transcriptional programs, *Journal of Clinical Investigation*, 2013 Feb 1;123(2):727-39.

19. Stadtfeld M*, Apostolou E*, **Ferrari F**, Choi J, Walsh RM, Chen T, Ooi SS, Kim SY, Bestor TH, Shioda T, Park PJ, Hochedlinger K. Ascorbic acid prevents loss of Dlk1-Dio3 imprinting and facilitates generation of all-iPS cell mice from terminally differentiated B cells. *Nat Genet.* 2012 Mar 4;44(4):398-405, S1-2.

20. Agnelli L*, Forcato M*, **Ferrari F**, Tuana G, Todoerti K, Walker BA, Morgan GJ, Lombardi L, Bicciato S, Neri A. The reconstruction of transcriptional networks reveals critical genes with implications for clinical outcome of multiple myeloma. *Clin Cancer Res.* 2011 Dec 1;17(23):7402-12.

21. Gurumurthy S, Xie S.Z., Alagesan B, Kim J, Yusuf R.Z., Saez B, Tzatsos A, Oszolak F, Milos P, **Ferrari F**, Park P.J., Shirihai O, Scadden D.T., and Bardeesy N. Lkb1 regulates stem cell quiescence, energy metabolism, and cell survival in the hematopoietic system. *Nature*. 2010 Dec 2;468(7324):659-63.

22. Mosca L, Fabris S, Lionetti M, Todoerti K, Agnelli L, Morabito F, Cutrona G, Andronache A, Matis S, **Ferrari F**, Gentile M, Spriano M, Callea V, Festini G, Molica S, Lambertenghi Deliliers G, Bicciato S, Ferrarini M and Neri A. Integrative genomics analyses reveal molecularly distinct subgroups of B-cell chronic lymphocytic leukemia patients with 13q14 deletion, *Clin Cancer Res*. 2010 Dec 1;16(23):5641-53. Epub 2010 Oct 14.

23. Martello G, Rosato A, **Ferrari F**, Manfrin A, Cordenonsi M, Dupont S, Enzo E, Guzzardo V, Rondina M, Spruce T, Parenti AR, Daidone MG, Bicciato S and Piccolo S. A microRNA targeting Dicer for metastasis control, *Cell* 2010 Jun 25; 141(7):1195-207.

24. Bicciato S, Spinelli R, Zampieri M, Mangano E, **Ferrari F**, Beltrame L, Cifola I, Peano C, Solari A, Battaglia C. A computational procedure to identify significant overlap of differentially expressed and genomic imbalanced regions in cancers datasets. *Nucleic Acids Res*. 2009; 37(15):5057-70.

D. Additional Information: Research Support and/or Scholastic Performance

Lab funding: My lab research activity is currently supported by an AIRC Startup grant, a 5 years grant for young investigators (grant #16841); by IFOM institutional support for new junior PIs; and by two postdoctoral fellowships by SIPOD (Structured International Post Doc program of SEMM), a Marie Curie co-funded fellowship program.