WHITE PAPER AP9
BIOTECHNOLOGY

EXECUTIVE SUMMARY

The present document summaries the scientific activities of about eighty scientists and eleven CNR institutes. These activities can be traced down up to the birth of biotechnology and bioinformatics activities in Italy, which were actually born within CNR.
The growth of the biotechnology research and related industry in the recent years has been unprecedented; advancements in molecular modeling, disease characterization, drug discovery, personalized healthcare and agriculture fundamentally impact economic and social issues in the European scenario. When bioinformatics approaches and technological innovation meet the demands of biotechnology (e.g. environmental biotech, drug design and discovery, genomics, regenerative medicine), relevant breakthrough research is carried out, as recognized by the significant number of project and publications of CNR-DIITET personnel. This growing importance is not only because bioinformatics handles large volumes of data but also in the usefulness of bioinformatics tools to predict, analyze and help interpretation in clinical and preclinical findings.
One of the greatest challenges facing the CNR DIITET biotechnology community is to be able to make use of the vast and dynamic influx of omics data. The synchronized development of bioinformatics concepts and related computational tools for prediction and modelling is a prerequisite to enable the exploitation of this wealth of biological data as a source of new biotechnological applications. The rapid technological progress of DNA sequencing has greatly facilitated the fast mapping of microbial genomes. This, however, necessitates matching bioinformatics platforms for subsequent analysis.
This document presents the current activities of AP Biotechnology, divided into 6 pillars fully in line with the European priorities:

- Analysis, management and integration of biological big data
- Development of network-based approach towards precision medicine
- Biophotonics
- Bioprocess and Tissue engineering
- Mathematical and software tools for bioengineering and biology
- Bio-sensors and Bio-inspired systems

After a short description of the state of the art of each pillar, the current activities are described, together with their impact and the identification of the specific challenges of the next future. The description of the effort in terms of person/months, the available infrastructures and facilities are described in the annex.

1. STATE OF THE ART OF THE RELEVANT SCIENTIFIC AREA

Biotechnologies provide an ever changing and evolving scenario, where research has obtained outstanding scientific and technological results, by developing and integrating multidisciplinary competences. Such progresses have economic impacts in many different fields, such as welfare, healthcare, agriculture and industry, and as a key component in social challenges addressed by H2020 program. The research themes require competences both in base sciences like Chemistry, Physics, Mathematics, Engineering, Computer Science, Bioinformatics and Systems Biology and in more specialized ones like biotechnological, medical, material science and life sciences. Results from this area are strictly connected with those in Healthcare and Wellbeing, Nanotechnologies, Applied Mathematics, ICT Systems and Devices and Low Carbon Technologies, although the characterizing parts are peculiar to the present area.

Scientific and technological research in this field is fueled by a constant progress in technologies. The potential of such technologies requires tremendous efforts to be fully exploited.
The convergence of new measurement and visualization technologies, and new computational and mathematical tools, bioinformatics approached and systems approaches to disease models can be expected to allow to overpass the current medicine approach, where we wait until the patient is sick before responding, with a personalized, predictive, preventive, and participatory **(P4) medicine** that will be cost effective and increasingly focused on wellness.

In more detail, the aim of this area is to develop novel methodologies, algorithms, software tools, hardware devices, as well as processes and technologies for clinical, medical and biotechnological applications. This aim can be reached only merging competences already available in the Institutes with those belonging to national and international collaborators. Examples can be found in different biotechnology fields: new **bioinformatics** approaches have been developed for gaining more insights into fundamental biological processes and functions; realization of electronic devices mimicking essential properties of biological synapses*;* data analysis methodologies and bioelectronic and biomechanical technologies have been carried out for advanced therapies, targeted drug delivery and sensing devices; biophotonics, embracing all light-based technologies applied to the life sciences and medicine, provides a better and earlier diagnostics, novel biomarkers, multifactorial and multigenetic diseases and comorbidity discoveries and advanced therapies; dynamic **tissue engineered systems** have been designed and realized to provide more realistic and controllable environments to better simulate the real human healthy/pathological conditions in vitro, thus acting as reliable platforms for generating in vitro predictive results.
The same happens in **omics data production**: a number of high-throughput sequencing techniques (e.g. 3C, 4C, Hi-C) have recently been developed to count the internal interactions of the DNA chain in cell populations. Since the spatial structure of the chromatin has been recognized as an important factor in gene expression and epigenetics, many 3D reconstruction algorithms are being developed, e.g., to correlate the known properties of genes with their spatial positions and relationships. One of the major challenges is to integrate the huge and dynamic avalanche of biological big “omics” data, with imaging data. The development of innovative bioinformatics approaches in the spirit of reproducible research according to the FAIR (findable, accessible, interoperable and reusable) principles, taking advantage of high performance computers, becomes a prerequisite to completely fill the gap between available clinical/biological data and our ability to structure these data in information to gain novel knowledge on relevant scientific and technological problems. Among the most challenging objectives we are going to fulfill, there are the analysis of single cell sequencing, the planning of artificial genomes for biotechnological applications, sustainable biomolecular pathways, the simulation of biological processes and functions, and the subsequent realization of artificial biological systems.
Finally, the development of novel methodologies, tools and industrial biotechnological processes based on advances technologies, such as those involved in **bio-catalysis**, will give rise to a class of novel products that cannot be obtained through presently available industrial technologies. Among these, the development of microbial and enzymatic processes in the framework of the biorefinery of renewable feedstock (*e.g.* lignocellulosic biomass residues) is a cross-cutting topic connected with low-carbon technologies and nanotechnologies. Moreover, molecular transport is the basis of biochemical cell processes (e.g., transcriptomics). Single particle tracking techniques revealed that diffusion of macromolecules is highly heterogeneous, thus posing new challenges to the understanding of the underlying mechanisms.
Nevertheless, imaging and sensing are among the main areas, where **biophotonics** can provide unique solutions and play an essential role. For example, big efforts of the scientific community are related to the development of sensing platform/device capable to monitor and detect bioanalytes and chemoanalytes at very low limit of detection, down to single molecule detection. On the other hand, a great effort is devoted to the development of new light-based approach, in order to induce selective and localized effects in biological tissues or biological materials. These approaches are useful in the design of personalized and precision medicine, surgery and diagnostics.

In such ever-changing scenario, we can identify several relevant scientific and technological areas, for which we are going to describe, in the next sections, contribution and impact.

2. CONTRIBUTION TO THE RELEVANT SCIENTIFIC AREA

Analysis, management and integration of biological big data (IAC, IASI, ICAR, ISSIA, IIT, ISTI)
CNR focused on the development of new statistical and machine learning approaches needed to analyze, integrate, visualize and manage biological big and high dimensional data, recognize patterns and models of behavior, as well as search innovative solutions to discover new clinical knowledge starting from biomedical and next generation sequencing data. Our tools are conceived to support reproducible computational research and promote FAIR principles of Data Science. With respect to sequence analysis, we focused on the analysis of structural variations, absent biological sequences or nullomers, and their application in immunology, cancer development, and rare disease diagnosis.
CNR also contribute to the development of novel methods for the integration of omics and clinical data with the aim of revealing and predicting the cellular mechanisms and their disruption under pathological conditions, and for identifying genetic markers underlying the onset and progression of complex and multifactorial diseases, such as cancer. Recent activities are represented by the analysis of single cell sequencing data with the aim of better understanding diseases and stem cell differentiation.
Development of network-based approach towards precision medicine (IASI, ICAR, ISSIA, IAC, ISTI)
Novel methods for validating network communities and to study the robustness of gene–regulatory time varing networks have been developed. Additionally, we also developed advanced approaches for network-based survival regression able to identifying few potential biomarkers and discriminate patients in high-and low-risk groups. Network based approaches can also be applied to precision medicine to tailor the medical treatment to the individual characteristics of each patient based on genetic, environmental, and lifestyle factors: individuals have to be classified into subpopulations that differ in their susceptibility to a particular disease, in the biology or prognosis of those diseases they may develop, or in their response to a specific treatment. The emerging tools of network medicine offer a platform to explore systematically not only the molecular complexity of a particular disease, leading to the identification of disease modules and pathways, but also the molecular relationships among apparently distinct (patho)phenotypes. This has led to the identification of new potential therapeutic targets for the treatment of different types of cancer. Furthermore, it is increasingly being recognized that quantitative description of the mechanisms connecting and regulating the various bioprocesses that sustain growth and proliferation of eukaryotic cells requires multi-scale dynamic modeling like the “whole cell” mathematical models. CNR developed a modular, coarse-grain “whole cell” model that accounts most important cellular activities, such as metabolism, growth and cycle.
Biological processes in the cell as information transfer processes, by using Shannon’s Information Theory have been also modeled. The aim is understanding how these processes are altered in cancer, since gene work in an orchestra rather than working alone. This has necessitated the development of gene interaction networks. Estimation of these networks is based on Bayesian networks and other graphical model selection methods. We also address the problem of modelling time varying gene-interaction networks using Monte Carlo for fitting a vector linear regression with time varying coefficients, and Markov chain networks to study healthy and cancer cells and understand the difference in dynamics.
Biophotonics (IFAC)
CNR is highly involved in the development of optical resonant structures/platform for the detection and monitoring of bioanalytes for both clinical and environmental applications; moreover, CNR is involved in the implementation of advanced microscopy techniques including second harmonic generation (SHG), atomic force (AFM) and Raman microscopy for diagnostics and spectroscopy of tissues and biological fluids, microscopy and photoacoustic tomography for the early detection of tumor lesions. Preclinical and clinical methods, technologies and procedures for therapeutic and surgical use, based on the use of lasers, LEDs,

biopolymers and laser-activating contrast agents for the following applications have been also developed. CNR launched several actions on laser-induced repair of biological tissues (ophthalmology, microvascular, neuro) and haemostasis induced by high-power blue-LED (vascular dermatology).

Bioprocess and Tissue engineering (IEIIT, IRC)

Concerning advanced biofuels production through the lignocellulosic sugar platform, main open issues related to the current technical barriers to the process commercialization have been addressed. Novel integrated pretreatment and hydrolysis processes for saccharification of biomass from agro-food wastes, kinetic characterization and optimization of enzymatic hydrolysis of biomasses, and catalytic conversion of ethanol into butanol have been reported. CO_2 capture through enzymatic reactive absorption processes has been developed in terms of biocatalyst design through the characterization of free enzyme catalysis at industrially relevant conditions, the immobilization of enzymes by covalent attachment and by carrier free techniques, the modeling of absorption units operating with the developed biocatalysts.

Bioreactor technology together with in silico modeling to replicate the physiological complexities of portions of human organs. In particular, fluidic bioreactors are designed and realized with advanced technologies (3D printing) for different healthy/pathological applications, for (i) mimicking the fluidic and mechanical stimuli of articular cartilage and validating novel implants for orthopedic application, or for (ii) reproducing in vitro the fluidic stimuli of the micro-circulatory system and testing the new drugs stability within the circulatory systems.

Mathematical and software tools for bioengineering and biology (ISSIA, ISTI, IMATI, ICAR)

CNR studies cancer mutations from a statistical point of view. The cancer mutations, unlike evolutionary mutations, do not have a Poisson dynamics but a Pareto-law dynamics. We are currently working on developing parametric models for cancer mutations similar to the Kimura model for evolutionary mutations. Further work is focused on the production of personalized medical products and devices through new design approaches and technologies, in an environment closely integrated with the hospital, which guarantees the direct contact between patients, medical personnel and devices manufacturers. This has the potential to drastically improve the quality of life of people, the performance of health care services and the competitiveness of medical products manufacturers. In particular, advanced mathematical and modelling technologies will be applied to build the anatomical model of patient through the combination and analysis of bio-images; the model is sent to a “hospital factory”, which can be located inside the hospital or closely integrated with it by the adoption of real-time connection tools; the hospital factory is able to produce the personalized product in a very short time, thanks to the combination of a set of innovative technologies (such as additive manufacturing, micro-EDM, micro-extrusion) and processes to manufacture personalized products; the personalized product is implanted to the patient with the guarantee of optimal size, fitting and comfort, and with minimum time and complexity of surgery.

Besides products, the hospital factory will produce also personalized anatomical models (e.g. cardiovascular models) that can help surgeons to study patients’ pathologies treatment in advance, in order to simulate different surgery strategies.

Bio-sensors and Bio-inspired systems (IMEM)

Organic memristive devices, the element that was specially developed for mimicking synapse properties, were successfully used for the realization of artificial neural networks, systems with long-term potentiation and depression, as well as short-term potentiation and depression (similar to those in nervous system), and even for making simulations of neurodegenerative disease pathology. It has been shown that it is possible to use them in electronic circuits capable of learning according the Hebbian rule. It was also shown the possibility to provide an artificial synaptic connection between two nervous cells from the rat brain cortex. The continuation of these works has resulted in the preparation of proposals, as that submitted to the H2020 FET-Open program.

Electrochemical and field effect devices based on organic conductors and graphene layers are currently developed, together with other electronic platforms, for different applications, such as biosensing, monitoring of biological system activities, neuromorphic applications, etc. All the devices are prepared through different technological processes, such as electrospinning, Langmuir-Schaefer procedure and lithographic techniques, and are endowed with innovative features, such as Additive Manufacturing. Chemistry of functional materials and surfaces, together with specific device architectures, are currently aimed at enhancing devices performance and at conferring multifunctional aspects in terms of device operation.

In the field of smart containers for targeted delivery and induced release of drugs, several pioneering approaches for fine targeting and desease-induced release have been developed. Further development of these studies is focused on the realization of smart implants, containing diagnostic tools, containers with appropriate drugs, release triggering tools, and antenna and electronics for the communications with patients and physicians.

3. IMPACT

Analysis, management and integration of biological big data

Today it is necessary to develop new techniques of analysis on big data to cope with the increasing volume of genomic and proteomic data. Efforts to build biological data base management systems that allow efficient access to non-expert users will enlarge the number of data science studies on the huge amount of data collected in the bioscience domain, so providing a continuous increase of the knowledge about biochemical mechanisms and processes that are the base of the life and dead of cells. The development of faster machine learning algorithms together with the analysis and integration of different layers of information and big data sets will allow to deeply investigate the origin of cellular mechanisms and their disruption under pathological conditions, and, as consequence, to identify novel therapeutic targets and strategies.

Developing innovative solutions dedicated to the data analysis and predictive models allows the researchers to identify the interactions among essential components such as miRNA, mRNA and proteins that are responsible for the development of cancer diseases or chronic diseases with a high economic impact.

Moreover, the study and development of tools for information integration may enhance the analysis of biologists and physicians, to shed light on mechanisms triggering disease.

Computational tools that support reproducible computational research and FAIR (findable, accessible, interoperable, reproducible) principles of data sharing and management will lead to more transparent results.

Development of network-based approach towards precision medicine

The emerging approach of the precision medicine has enabled a new era of health care delivery and treatment. Understanding of the underlying mechanisms of high-impact diseases is allowing scientists to develop new drugs, targeted therapies, and preventive strategies. SWIM represents a solution of excellence addressing these issues that can improve the knowledge of the cellular events crucial for carcinogenesis and may unveil many potential prognostic and novel therapeutic targets. Using SWIM could provide important clues that will stimulate research activities into the causes of high-impact diseases, including but not limited to human cancer, thus supporting the planning of healthcare services such as clinical trials and disease prevention.

Moreover, the development of network-based approaches can help clinicians to provide more precise prognoses and to facilitate the subsequent clinical management of patients at risk of disease.

Predictive “whole cell” computational models are supposed to be exploited in personalized medicine, whenever a specific drug administration therapy is required to be tested on single virtual cells or populations of virtual cells, possibly adapting the individualized therapy to an environment subject to changes and uncertainties.

Biophotonics

The development of devices capable to achieve single molecule detection, or in any case very low limit of detection is a challenge which is pursued in the scientific world at international level; besides the intrinsic scientific importance to identify and monitor single molecules which can open new knowledge in the comprehension of the their physics and chemistry, the measurement of analytes at very low concentrations can play a fundamental role in medicine in the detection of the onset of pathologies at a very early stage: this can have an huge effect for the patients thanks to the administration of the correct therapy well in advance, as well as it can affect the social systems with a noticeable decrease in the healthcare expenses.

The use of light- based devices can be used in a tailor-made surgical application: this technology is useful in the realization of precise and customized surgical cuts and suturing patterns; or it can be used to design cardiovascular stents perfectly matching the patient's anatomy. All these approaches are at the basis of the precision and personalized medicine.

Bioprocess and Tissue engineering

Novel bioprocesses based on microbial and enzymatic biocatalysts play a crucial role in the impact of industrial biotechnologies on circular economy. The contribution by biochemical engineering science is mainly related with the development of novel biocatalysts and their characterization aimed at their use through rational design of bioreactor units towards the integration of unit operations (*e.g.* bio-separation processes). The integration of fermentation and recovery units for the production of bio-butanol through the ABE fermentation process is an example of how the novel industrial bioprocesses can act to improve productivity in biorefinery plants. Similarly, the investigation of cellulose hydrolysis catalyzed by cellulase cocktails aimed at the design and set up of novel bioreactor configuration for efficient recovery of fermentable sugars from lignocellulosic feedstock contributes to the development of sustainable processes for the production of second generation biofuels.

The potential impact of tissue engineering (TE) based technologies is that they can act as more predictable R&D platforms, by providing more physiologically relevant conditions than 2D culture and/or animal models. Moreover, these TE technologies could be adopted as in vitro screening platforms. Enabling the availability and effectiveness of personalized and predictive screening tools for therapies that will prevent or limit spread of diseases is one of the biggest unmet clinical challenges at the moment, both for the scientific and the industrial pharmaceutical community. Providing suitable options to the high cost, inefficient and ethically complex use of animal models in pre-clinical research and development phases, with better chances of success in clinical trials is another challenge for the scientific and industrial research community, and the society.

Mathematical and software tools for bioengineering and biology

The demand of more intensive and advanced healthcare services is clearly in contrast with the general pressure to reduce the cost of healthcare systems in the European countries, due to national financial sustainability reasons. Consequently, new innovative models are required to solve this paradox in order to continue guaranteeing the growth of social and health conditions of citizens. These models should break the usual boundaries of the existing healthcare systems, not only in terms of medical services and healthcare excellence, but also in the field of technology and industry, which is a fundamental enabler of the value provided to customers in healthcare services.

Our activity has the following scientific and technological impacts: define new statistical methods and approaches to elaborate an accurate anatomical model from bio-images in order to design personalized products; define improvements to the existing technologies for personalized products, to lower production costs, thus allowing the wide diffusion of personalized medical products; propose innovative process combinations and manufacturing approaches; demonstrate the new technological approaches in real products and medical scenarios.

Bio-sensors and Bio-inspired systems

The progress of systems for bio-mimicking information processing requires the use of synapse-like electronic elements what will allow to combine memory and processor in a single system. Such system will allow learning at the hardware level and parallel information treatment. Developed devices and networks will mimic some functions of the nervous system and brain, what will make possible experimentations that are impossible or forbidden to perform on animals and brains. Finally, such devices are expected to be interfaced with nervous system for recovering damaged parts and to control prosthesis.

As all pharmaceutical preparation have side effects, ideal container must provide their delivery to the diseased areas of the body or to the zones of risk, and the active drug must be released when the disease occurs. In this respect, containers based on nanoengineered polymeric capsules are very challenging. The whole internal part of these object can be filled with active drug, while their shell architecture (incorporation of magnetic nanoparticles and receptors for targeted delivery and special compounds for induced release) guarantees that the drug will be delivered in a proper area and the release will occur only in the case of disease.
The short-term perspectives are focused on the accumulation of a robust and diversified technological experience in the fabrication of different device structures, through the combination of several advanced and powerful technological tools, and on the implementation of novel characterization protocols.
In the medium-to-long term, organic biosensors and devices will be customized with different structural properties and features to optimize the response with respect to specific application scenarios and to widen the portfolio of potential application fields; moreover, OECTs will be miniaturized and transferred onto biocompatible substrates for in-vivo biomedical applications.

4. EMERGING RESEARCH CHALLENGES

1) Development of statistical algorithms and computational tools for connecting resources, integrating imaging, sequencing omics data for human health and ecology.
2) Novel algorithms and precision molecular tools to explore the diseases molecular complexity. Efficient, scalable algorithms for high-throughput omic-analysis, visualization methods of reconstructed structures.
3) Single-molecule detection based on optical resonating structures; realization of personalized and precise surgery, tailor made on patient's organ morphology.
4) Novel pathways for integrating thermo- / bio-chemical conversion processes. CO_2 capture and use through microbial and enzymatic processes. 3D healthy/pathological models for the study of micro-environmentally driven cellular mechanisms.
5) New models to guarantee the growth of social and health conditions of citizens.
6) Realization of implantable devices controlling drug release, and artificial analogues of nervous system damaged parts for implantation.

5. CONCLUSIONS

Biotechnology provides a constantly evolving scenario, where research has achieved significant scientific and technological progress using and integrating multidisciplinary skills. These advances have economic repercussions in many sectors, such as health, agriculture and industry, and in support of the various social challenges of the European H2020 program. The research topics addressed in this project activity require skills in both technology, medical, and life sciences. As described in the present document, the activities within the DIITET provide results with close ties with other design areas such as Healthcare and Well-being, Nanotechnology and Applied Mathematics and Low Carbon Technologies. In particular, the aims of this design activity are focused on developing new methodologies, algorithms and software tools for synthetic biology, new bioinformatics approaches in clinical, medical and biotechnological fields, industrial biotechnological processes, data analysis methodology and advanced therapy technologies, biophotonics. Scientific results are achieved by putting the existing skills in the CNR-DIITET Institutes involved with those of the national and international realities with which they work together.
The results attest a gradient ascending direction towards major achievements in terms of scientific contribution, the ability to collaborate in national and international research projects, and young scientists training.

PROJECT AREA 9: BIOTECHNOLOGY

Editorial team and Contact person (CP)	Institute	Email
ANCONA NICOLA	ISSIA	ancona@ba.issia.cnr.it
ANGELINI CLAUDIA	IAC	c.angelini@iac.cnr.it
BALDINI FRANCESCO	IFAC	f.baldini@ifac.cnr.it
D'ANGELO PASQUALE	IMEM	pasquale.dangelo@imem.cnr.it
EROKHIN VICTOR	IMEM	victor.erokhin@fis.unipr.it
GUARRACINO MARIO (**CP**)	ICAR	mario.guarracino@cnr.it
KURUOGLU ERCAN	ISTI	ercan.kuruoglu@isti.cnr.it
LANZARONE ETTORE	IMATI	ettore@mi.imati.cnr.it
PACI PAOLA	IASI	paola.paci@iasi.cnr.it
PELLEGRINI MARCO	IIT	marco.pellegrini@iit.cnr.it
RUSSO MARIA ELENA	IRC	m.russo@irc.cnr.it
SCAGLIONE SILVIA (**CP**)	IEIIT	silvia.scaglione@ieiit.cnr.it
TOCCOLI TULLIO	IMEM	toccoli@science.unitn.it