Research Proposal

The role of genetic and acquired risk factors for cancer development in Galapagos residents.

Learning from an area at reduced risk for a better cancer control program worldwide

Version 4.0 prepared by Massimo Federico, Anna Iannone, Michele Maffia, Roberta Mele and Luana Conte

08 April 2015

A1 SUMMARY

Cancer is a disease worldwide accounting for 14.1 million new cancer cases and 8.2 million cancer related deaths in 2012 (GLOBOCAN). This phenomenon is likely to increase by the 2012, due to growth, ageing of the global population, inherited disease and modifiable factors including diet (diabetes, overweight, and obesity), physical activity habits and environment. It has been shown that cancer mortality among residents in the Galapagos Island is lower compared to Guayas region and much more compared to Modena in 2012, with a relative risk (RR) of cancer death (period 2004-2012) RR= 0.44 (CI95 0.33-0.58) P<0.001. The Galapagos Islands contain the most active volcanoes in the world, home of one of the longest-lived vertebrates in the world, the Galapagos tortoise. The young population accounts for more than 25,000 people predominantly Ecuadorian arrived last century and colonists from mainly Europe, America and Ecuador. The aim of this study is to investigate the habits of the Galapagos population by evaluating ***Cancer Epidemiology*** and ***Genetics***, ***Nutrition and Lifestyle***, ***Environment and Cancer risk***. The results will allow us to identify role of genetic, the protective factors, and acquired risk factors for cancer development as a basis for a better cancer control program worldwide and preventions.

A2 PROJECT

A2.1. Background

Although genetic susceptibility influences the risk of cancer, most of the variation in cancer risk across populations and among individuals is due to factors that are not inherited. Up to 30% of cancer are estimated to be preventable by behaviors such as avoiding exposure to tobacco products, maintaining a healthy weight, staying physically active throughout life, and consuming a healthy diet can substantially reduce the risk of developing or dying from cancer.
From the Anuario de Estadísticas Vitales Nacimientos y Defunciones 2012, published by Instituto Nacional de Estadística y Censos de la República del Ecuador (www.inec.gob.ec) we realized that cancer mortality among residents in the Galapagos Islands was lower than that observed in the Guayas region, in Ecuador as a whole, and, much more, in Modena (Italy) (Table 1)

Table 1

Country	*year*	*no. cases*	*STD(W)*[1]
Ecuador	2012	10,234	**70.7**
Guayas	2012	2,414	**67.1**
Galapagos	2012	10	**38.6**
Modena	2011	1,958	**94.1**

[1] *STD(W) World age-standardized rate*

The differences between Galapagos and Guayas are more evident and robust when the comparison is performed for the period 2004-2012: the Relative Risk of Cancer Death is 0.44 (CI95 0.33-0.58), P<0.001.
It is well known that the Galápagos Islands are one of the few places in the world without an indigenous population. The largest ethnic group is composed of Ecuadorian who arrived mainly in the last century from the continental part of Ecuador. In 1920s and 1930s, a small wave of European settlers arrived in the islands and a few years later, other colonists from Europe,

America and Ecuador started arriving on the islands, seeking a simpler life. Overtime, population had risen to more than 15,000 people, and in 2010 there were 25,124 people living in the Galapagos.

The finding of a reduced cancer mortality among residents of the Galapagos Islands make this population a unique source to evaluate if the specific factors that cause this low risk are also associated with a change in the frequency and distribution of specific features in the cancer types where this low risk is observed. Along with epidemiologic considerations other factors need to be taken in account. Genetics and environmental factors, alone or in combination, might have a relevance.

Cancer mortality among residents in the Galapagos Islands and the Guayas region differs substantially (38,6 vs. 67,1 cases/population year x 100.000) as described in Table 1. However, data about type and frequency of solid tumors such as gastric cancer, lung cancer, colorectal cancer, arising in these two geographical areas are scarce. In addition, no data are available on the genetic lesions underlying certain crucial genes involved in pathogenesis of the solid tumors arising in these areas. Thus, a precise molecular characterization of these tumors is still lacking: a characterization of tumors arising in Galapagos Islands compared to inland territories, by analyzing mutations of Mitogen-Activated Protein Kinases (MAPKs) pathway genes is necessary.

The **environment** is mostly related to geological factors. The Galapagos Islands are a group of the most active volcanoes in the world whose lavas were emplaced continuously for over 2 millions of years. Their age and geochemical composition are distinct from the eastern islands to western ones as well as from the northern to the southern. In support of the beneficial effect of environment there is also the observation that the Galapagos tortoise is the largest living species of tortoise and, with life spans in the wild of over 100 years, it is one of the longest-lived vertebrates.

Today, the Galapagos Islands are still one of the most active volcanoes in the world. In spite of the biological interest, their volcanologic, petrologic, and geochemical features are scarce: in particular, information on abundance of trace elements and on soils are very poor. Furthermore, these islands are characterized by a small number of native plant species (approximately 600), whose biochemical and antioxidant profiles are poorly characterized: the microclimate and the soil high salinity may greatly affect the synthesis and accumulation of biologically active molecules within allochthonous fruits and vegetables.

Another issue is **nutrition** and lifestyle among the residents. It is well known that any food or behavior related to lifestyle can promote or prevent tumor development. We know that Western diet is known to be associated with a group of frequent cancers such as breast cancer, colorectal cancer, prostate cancer and endometrial cancer, which tended to be rare in Eastern and developing countries, while stomach cancer used to be very frequent in populations with a diet poor in micronutrients and fresh foods.

Another example is the population of Okinawa, a Japanese island where the inhabitants are living longer and with a lower incidence of cancer, due to diet rich in sweet potatoes, vegetables, soybeans and low in animal protein. Okinawa people have also active lifestyles, with amount of energy expenditure that exceeds the energy intake.

Food and physical activity can *directly* promote or prevent tumor development or they can change *body composition* thus *indirectly* influencing tumor evolution.

A2.2. General aim

The aim of this study is to investigate the habits of the Galapagos population by evaluating four different task forces: (1) ***Cancer Epidemiology***, (2) ***Cancer Genetics***, (3) ***Environment and Cancer risk***, (4) ***Nutrition and Lifestyle***.

A2.3. Specific aims

(1) Cancer Epidemiology	An **epidemiologic study** is the first pivotal step of the project. We had to confirm the low cancer risk in this population. Cancer mortality data provide a reliable indication of the cancer risk in a population: although the availability of effective treatments and the presence of high-quality screenings can significantly decrease it, the overall picture of the cancer burden in a population provided by mortality data is more credible than that provided by incidence data, heavily influenced by the diagnostic pressure, with lead time and length bias, and by differences in the classification of borderline cases. As a consequence, any report of reduced (or increased) cancer mortality needs to be considered as a very strong suggestion of a truly reduced (or increased) cancer risk. Yet, the reports from the Galapagos Island need to be confirmed and further explored. To this aim, both mortality and incidence data are going to be necessary, because they provide complementary evidence. The presence of mortality data and of a cancer registry covering the population of the entire Guayas region will allow a thorough assessment of the cancer risk among Galapagos residents. Site-specific and age-specific cancer incidence and mortality rates will be also explored. It is now clear that the environmental determinants of cancer risk show a different effect on cancers of different sites: the pattern of cancer risk by site observed in a given population may provide essential clues for unraveling the environmental factors that are determining the cancer risk in that population, and, more in general, for etiologic research. The comparison between the incidence of "groups" of cancers (identified on the basis their dominant etiologic factors) in the Galapagos population and in the Guayas population will allow to tentatively identify the environmental factors that are responsible of the differences in cancer mortality and incidence between the two populations. Moreover, we aim to compare the survival and the distribution of the prognostic features of the cancer cases diagnosed among residents of the Galapagos Islands, as compared with those of the Guayas region, in order to explore the possible presence of biological differences Specific factors which are responsible of the biological heterogeneity of most solid tumors are largely unknown. The presence of a population with a very low cancer incidence and mortality provides a unique opportunity to evaluate if the specific factors that cause this low risk are also associated with a change in the frequency and distribution of specific features in the cancer types where this low risk is observed. This hypothesis will be evaluated by comparing the distribution of specific cancer characteristics routinely assessed in all cases of each cancer type (such as histology, Grade, TNM, presence of hormonal receptors In breast cancer) and various molecular marker in cases diagnosed in the Galapagos population as compared to the whole population of the Guayas regions. Furthermore, overall and stage-specific survival will be compared in the two populations for each cancer type and subtype. It will be also assessed analytical case/control study and/or cohort study of molecular biomarkers of early exposure to be compared between individuals of the study population with individuals belonging to a population with the highest cancer mortality. The study will include "site-specific" recognition of chemical contaminants in air and water matrices; the obteined data will be used for risk

	analysis (ADR). This part of the study will be carried out in collaboration with the Task Force on Environment and Cancer Risk, with regard to the analysis of chemical contaminants in the matrices, water and air, as well as in collaboration with the task force of Cancer Genetics to study the micronuclei in the cells of oral mucosa and peripheral blood.
(2) Cancer genetics	As soon as the epidemiological study will confirm the low cancer mortality in the Galapagos Island, we will continue with the other research issues. In principle, differences in cancer types are related to genetic or environmental factors, or to a combination of the two. Considering **genetics** a potential cancer risk factor, alone or in combination with environmental factors, we aim, to better understand the biological history of the solid tumors in the Galapagos Islands and Guayas areas, to analyze the genetic features of the most important growth factor signaling pathway components. Analysis of the presence of somatic mutations in driver genes, mainly the RAS GTPase family and the EGFR, will be of help to determine the features of the cancer types more frequent in the Galapagos Islands and the Guayas region and will allow their molecular sub-classification. Identification of tumor subsets by characterization of genetic lesions in these geographical areas will held interesting information about the "possible" role of environmental factors on the tumor biology. These findings will be compared to the data already observed for the same solid tumors of the western population and for some cancer (i.e. gastric cancer and NSCLC) in the Asiatic people. To this end, an exon-targeted NGS chip will be realized with 25-30 specific genes and tested starting from archival tissues of the Pathology unit. Major aim of the study is a precise molecular characterization of the most frequent cancer types in Galapagos Islands and Guayas region in order to find if a particular genetic pattern within each tumor may explain the observed difference in the incidence rate of tumors between the two regions. Four major goals will be pursued: (1) Analysis of the different distribution of solid tumors types between the Galapagos residents and the Guayas region and comparison to that observed elsewhere (2) Analysis of the frequency and types of mutations affecting Mitogen-Activated Protein Kinases (MAPKs) pathway genes in the solid tumors among Galapagos and inland territories populations cancer patients (3) Comparison of the rate and frequency of the somatic mutation data in Galapagos and Guayas populations with that of solid tumors arising in the western populations. (4) Analysis of the somatic mutations found among the tumor tissue in the corresponding germ line context. We will analyze also the genetic variation related to HSP90 gene and xenobiotic metabolizing genes (such as:. CYP2D6, GSTM1, GSTT1 and others) The rationale for these proposed analysis is explained below. HSP90- The adaptability of a population to environment partially depends on the presence of genetic variation in genes involved in the answer to external stimuli, such as the heat shock protein (HSP) genes. The expression of inducible HSP90 gene is mainly regulated at transcriptional level depending on the interaction between transcription factor and their putative cis-acting elements

	in the promoter region of HSP90 genes. A total of 18 variants were found in the HSP90 α gene thus, the response of an individual to environmental stresses or diseases may depend on his/her Hsp90 status. Xenobiotic metabolizing genes - The xenobiotic-metabolizing genes represents a group of genes that catalyze metabolism of a diverse array of endogenous substrates such as fatty acids, steroids, and vitamin D, as well as exogenous compounds including phytochemicals, environmental pollutants, and drug. Polymorphisms in these genes are involved in Adverse Reaction to environmental molecules and drugs. *They* might also increase the risk of cancer through the generation of procarcinogens whereas other variations might be protective factors for cancer. Metabolomics studies- Metabolic profiling of Galapagos patients cultured cancer cells will be also used both on lysates and culture medium in order to assess possible differences with respect to analogous non indigenous cancers such as those from European countries.
(3) Environment and cancer risk	The Galapagos tortoise is the largest living species of tortoise and, with life spans in the wild of over 100 years, it is one of the longest-lived vertebrates. Evidence-based studies indicate that two major factors contribute to an individual's longevity, genetics and lifestyle choices. Twin studies have estimated that approximately 20-30% of an individual's lifespan is related to genetics, the rest is due to individual behaviors and environmental factors. The potential for toxic effects associated with exposure to heavy metals has been well documented. Arsenic (As), lead (Pb), and mercury (Hg) are the most prevalent environmental toxic metals and, since 1997, occupy the first three positions of the U.S. Superfund list of hazardous substances (http://www.atsdr.cdc.gov/spl/), ahead of benzene and followed by cadmium (Cd) and hexavalent chromium (Cr). Many studies of heavy metal-induced carcinogenesis show that such heavy metals cause reduction in expression of DNA repair enzymes, some through epigenetic mechanisms. Public awareness toward heavy-metal contaminated water and soils and the fate of such pollutants in the environment, and in particular throughout the food chain, has steadily increased and the national and transnational legislations have imposed very tight limits to the tolerable concentrations of these chemicals in both civil and industrial wastewaters. As a consequence, there is a demand for new techniques to selectively identify and study the actions of these toxic agents. Combined with the problem of heavy metal contamination, particularly in marine settings, there are also other issues such as bioaccumulation, biomagnification and persistence in the environment. The Galapagos Islands are a group of the most active volcanoes in the world whose lavas were emplaced continuously for over 2 millions of years. Their age and geochemical composition are distinct from the eastern islands to western ones as well as from the northern to the southern. In spite of the biological interest, the Galapagos' volcanologic, petrologic, and geochemical features are still scarce. In the last two decades a number of studies looked at contribute to understanding the origin and provenance of magmas and reliable isotopic geochemical data were provided in last few years. However, no information on

	abundance of trace elements was provided. Similarly, information on Galapagos soils is very poor; they are limited to those recorded in the sixty and to some recent studies, which were mainly focused on pedological aspects (i.e., description of their profiles, properties and micromorphological characteristics). This area is also characterized by a small number of native plant species (~600), whose biochemical and antioxidant profiles are poorly characterized. Furthermore, the archipelago's microclimate and the soil high salinity may greatly affect the synthesis and accumulation of biologically active molecules within allochthonous fruits and vegetable. One approach that is gaining scientific interest in cancer therapy is the use of natural chemical compounds that occur naturally in plants.
(4) Nutrition and lifestyle	This project aims to investigate the relationship between the environment in which this population live and their Genome. Air, water, soil, foods are the essential elements to which these individuals are exposed or "e*xposome"*. Our behavior influences the choice of a healthy lifestyle or not. These choices correspond to the "b*ehavome"*, and our lifestyle is the result of our behavior in relation to the exposome in which we live. This task force will investigate and compare the *lifestyle, nutrition and physical activity,* of a group of people living in the Galapagos Island with that of a sex- and age-matched group of Ecuador inhabitants in order to find out if there is any food or behavior related to lifestyle that can promote or prevent tumor development in this population. It will be also investigated the *metabolic profile*, to assess how the life style and a specific type of diet can modify parameters which usually are correlated to a lower incidence of cancer. Metabolites are the end product of all cellular processes, and are a direct outcome of enzymatic and protein activity. Thus, metabolites are more proximal to a phenotype or disease than either genetic or proteomic information. As such, the overall goal of metabolomics is to identify the few chemical features against a large and complex background of metabolites that uniquely define the system. These few chemical features or metabolites should be directly related to the defining characteristic of the system. In particular, the study of the metabolic fingerprinting aims to determine the relative differences between the metabolomes of two or more systems, in order to understand the biological relationship among life style, food habits and cancer disease. For this reason a series of biochemical parameters will be measured in plasma and urine. Interest will be focused mainly on the evaluation of the *antioxidant and metabolic profiles* that could correlate with a lower incidence of cancer. In a parallel approach we will design specific experimental diets (formulated according to the nutritional habits of Galapagos population) to be used in animal studies. The aim of these studies is to provide the biochemical basis by which specific nutrients, nutrient combination and/or biologically active molecules are able to prevent cancer.
A2.4. Task Forces Activities	
(1) Cancer epidemiology	- Confirmation of the low cancer risk of this population; - Providing clues on the causes of this reduced risk, by evaluating site-specific and age-specific cancer incidence and mortality rates; - Comparison of the survival and the distribution of the prognostic features

	of the cancer cases diagnosed among residents of the Galapagos Islands, as compared with those of the Guayas region, in order to explore the possible presence of biological differences. - Case/Control studies
(2) Cancer Genetics	- NGS and Sanger sequencing. To determine the presence of mutations in solid tumor, an exon-targeted Next-Generation Sequencing approach will be used. - Sequencing results analysis and structure-function relationships analysis on the mutated vs. wild-type genes - FISH. For detection of some gene amplification (HER family, MET, AKT), or complex genetic alteration such as translocations or to detect chromosome instability - The metabolomic study of cultured cancer cells (supernatant and lysate), using NMR, can provide important information that could be potentially correlated with the incidence of cancer. NMR spectroscopy will be used to identify differences in cancer cell metabolites and to explore differences in pathways involved in cancer processes for Galapagos with respect to homologous non indigenous and European cancers. This part of the Cancer Genetics activity will be carried on by the same Metabolomic Unit (Salento University) involved in the Nutrition and Lifestyle activities of the present project.
(3) Environment and Cancer Risk	- Mineralogical and chemical investigations on soils and their parent materials; - Investigation of toxic effects associated with exposure to heavy metals; - Geomineralogical aspect of the atmospheric powder in order to understand the influence of the environment on human health; - Characterization of the most common plant species used in the local diet; - Determination of the effects of extracted bioactive molecules both in vitro and in vivo studies on animal models.
(4) Nutrition and Lifestyle	Assessment of lifestyle, nutrition and physical activity in the Galapagos Islands compared to inland territories, to find out food or behavior that can promote or prevent tumor development. A. Two large groups of people (*size to be defined*), age- and sex-matched, one among the inhabitants of the Galapagos and one among the inhabitants of Ecuador will be subjected to the following investigations: 1. evaluation of *energy intake* and *diet history* (eating habits, methods of cooking food) 2. assessment of *energy expenditure* 3. assessment of *body composition* B. Other two groups, formed by a smaller number of people belonging to each population (*size to be defined*), will be more deeply investigated: 1. *energy intake* evaluation 2. assessment of *energy expenditure* in *free living conditions* and lifestyle 3. assessment of *body composition:* bioimpedentiometry 4. assessment of plasma antioxidant status. 5. assessment of plasma lipid profile and diabetic markers In a parallel approach, animals will be fed with specific experimental diets,

	designed according to the nutritional habits of Galapagos population. Lipid biosynthesis and mitochondrial bioenergetics will be evaluated, since emerging studies have begun to demonstrate that targeting mitochondrial metabolism and lipogenesis is potentially a fruitful arena for the treatment of cancer.
Informatics	The exploitation of the modern Information and Communication Technologies may be very helpful in order to pursue the focal aim of this study: the comprehensive investigation of Galapagos population habits. In particular, the main technical activities will include: - design and implementation of a database able to store all data required for the analysis and monitoring of endogenous and exogenous factors that may influence the cancer risk; - desig and tuning of a multidimensional analysis system in charge to analyze several groups of people among the Galapagos population and to compare them with other populations showing similar features (food, lifestyle and weather conditions); - hot spot analysis to identify the locations of statistically significant hot and cold spots in cancer cases and cross-correlation with genetic, environmental and climate data; - Future projections analysis (based on climate change scenarios from the IPCC AR5) related to population habits indicators. Taking into account the large amount of scientific data available from the CMIP5 federated database (petabyte order), this activity will need scalable and robust big data analytics frameworks to run ensemble-based analysis; - design and implementation of web interfaces and web services needed for accessing and visualizing data; - design and tuning of software and hardware architecture; - validation and testing of the overall information system which will be able to collect, integrate, process and disseminate data on a global scale. Particular attention will be paid to issues related to data confidentiality, privacy protection and security of the whole system.
A2.5.Detailed program for each task force	
(1) Cancer epidemiology	***Confirmation of the low cancer risk of this population*** It is well known that cancer mortality data provide a reliable indication of the cancer risk in a population: although the availability of effective treatments and the presence of high-quality screenings can significantly decrease it, the overall picture of the cancer burden in a population provided by mortality data is more credible than that provided by incidence data, heavily influenced by the diagnostic pressure, with lead time and length bias, and by differences in the classification of borderline cases. As a consequence, any report of a reduced (or increased) cancer mortality needs to be considered as a very strong suggestion of a truly reduced (or increased) cancer risk. Yet, the reports from the Galapagos Island need to be confirmed and further explored. To this aim, both mortality and incidence data are going to be necessary, because they provide complementary evidence. For instance, a similar reduction in breast cancer incidence and mortality would provide definitive evidence of a true reduction in risk, not influenced by an increased diagnostic pressure (increase in incidence with a slightly reduced mortality) or

	by better therapies (same incidence with lower mortality) The presence of mortality data and of a cancer registry covering the population of the entire Guayas region will allow a thorough assessment of the cancer risk among Galapagos residents. First, the methodology of data collection and classification will be reviewed and validated. Second, age- and sex-adjusted incidence and mortality rates in the population of the Galapagos island and in the entire population of the Guayas region will be computed and compared, for all cancers and for specific cancer sites. Age- and sex- specific incidence and mortality rates will be evaluated to investigate the presence of cohort effects. ***Providing clues on the causes of this reduced risk, by evaluating site-specific and age-specific cancer incidence and mortality rates*** It is now clear that the environmental determinants of cancer risk show a different effect on cancers of different sites: to make a few examples, asbestos causes lung cancer and mesothelioma, while smoking causes lung cancer and cancers of various other sites but not mesothelioma. Western diet is known to be associated with a group of frequent cancers such as breast cancer, colorectal cancer, prostate cancer and endometrial cancer, which tended to be rare in Eastern and developing countries, while stomach cancer used to be very frequent in populations with a diet poor in micronutrients and fresh foods. As a consequence, the pattern of cancer risk by site observed in a given population may provide essential clues for unraveling the environmental factors that are determining the cancer risk in that population, and, more in general, for etiologic research. The comparison between the incidence of "groups" of cancers (identified on the basis their dominant etiologic factors) in the Galapagos population and in the Guayas population will allow to tentatively identify the environmental factors that are responsible of the differences in cancer mortality and incidence between the two populations. This information will be used to plan and interpret the results of both environmental and case-control analytical studies. Specifically, cancer specific and overall attributable risk will be computed, to set the targets for the ensuing analytical case-control studies, in order to evaluate if the difference in cancer risk can be accounted by differences in the distribution of known risk factors, or, instead, new etiological factors must be looked at. ***To compare the survival and the distribution of the prognostic features of the cancer cases diagnosed among residents of the Galapagos Islands, as compared with those of the Guayas region, in order to explore the possible presence of biological differences*** Often, the environmental determinants of cancer risk are also associated with changes in the distribution of cancer characteristics that are known to be prognostic markers and, in some instances, determinants of the efficacy of various treatments. However, this association has never been studied in depth, and the specific factors which are responsible of the biological heterogeneity of most solid tumors are largely unknown. The presence of a population with a very low cancer incidence and mortality provides a unique opportunity to evaluate if the specific factors that cause this low risk are also associated with a change in the frequency and distribution of specific features

	in the cancer types where this low risk is observed. This hypothesis will be evaluated by comparing the distribution of specific cancer characteristics routinely assessed in all cases of each cancer type such as histology, Grade, TNM, presence of hormonal receptors In breast cancer and various molecular marker in cases diagnosed in the Galapagos population as compared to the whole population of the Guayas regions. Furthermore, overall and stage-specific survival will be compared in the two populations for each cancer type and subtype.
(2) Cancer Genetics	Tissue archival samples of solid tumors stored in Pathology Unit will be used in a retrospective manner. Type and histological grade of the diagnosed solid tumors from, approximately, 2009 to 2014 will be collected and recorded. Macro dissection of areas of neoplastic cells will be performed to obtain enrichment of at least 70% of tumor cells for the subsequent experiments. In addition, normal tissue areas will be dissected and used for the analysis of mutations in the germ line genotype. **NGS and Sanger sequencing**. To determine the presence of mutations in solid tumor, an exon-targeted Next-Generation Sequencing approach will be used. This new technology has driven a paradigm shift from single gene testing to parallel multigene testing, allowing massive sequencing of mutations in 25-30 genes identified as relevant in the pathogenesis of solid tumors. The assembly of this custom targeted panel will include primarily gene mutations hot spots already identified as key mutations (like codon 12 and 13 of KRAS exone 2), but also will include those relevant genomic locations that are not included in published guidelines but will help in the molecular definition of the solid tumor (such as DDR2, a novel RTK, found mutated in squamous cell lung cancer). Tumor suppressor genes often do not have mutational hotspots like the oncogenes, but any mutation throughout their coding regions may be functionally relevant. Therefore, the gene panel assembled for this study will include full coding region coverage of tumor suppressor genes (like PTEN). The chip will then be comprised of those genes that will help the full molecular characterization of the various solid tumors for which it will be designed, such as for example the CDKN2a (the p16Ink4, an inhibitor of cell cycle progression from G1 into S phases) and CTNNB1 (a component of cadherin signaling pathway), both genes very important for the molecular characterization of sporadic Gastric Cancer. In addition, the panel will also include the p53 gene coding for an all-round tumor suppressor protein which is invariably highly mutated in all solid cancer types and contributes, through its alteration, to carcinogenesis by deregulation of some important checkpoints: DNA replication, cell proliferation and survival in response to DNA damage as well as in the control of basal cell bioenergetics metabolism. Direct Sanger sequencing will then be used to confirm the presence of the gene mutations found by NGS. **Structure-function analysis.** By mainly using a combined approach of comparative genomics, physiological genomics and bioinformatics protocols already used to study certain rare diseases (such as ARSACS, Kabuki syndrome and SLA), the relevance of the mutations found in the above indicated panel of human genes will be analyzed systematically with the aim of identifying

	evolutionary conserved motifs/domains/structures within the sequences that may be relevant for the definition of the role a DNA/RNA/protein may play in the manifestation of a diseased phenotype. Briefly, the comparative analysis will be conducted by recruiting sequence information from the genomes of non-mammalian vertebrates (birds-to-fish) and/or from model invertebrates, taking into account that in any sequences differences and similarities along the evolutionary scale are strictly related to the biological relevance of the gene product(s) in the sub-cellular/cellular/tissue/organ/system contexts. This in silico analysis will be more strongly supported by the concerted interaction with the other groups, and among these with the Informatics group, which will support the alignment-structure-function analysis actions, implement the structural-functional information in their global analysis models and/or develop new bioinformatics tools for rapid sequence analysis. **FISH.** For detection of some gene amplification (HER family, MET, AKT), or complex genetic alteration such as translocations or to detect chromosome instability, particularly important for some solid cancer (such as the gastric cancer) a different molecular approach will be used. The FISH methods allows detection of gene target sequences using fluorescent dye labeled DNA probes, one for the target gene sequence and one for the centromere sequence of chromosome bearing the target gene. The nuclear DNA of either interphase cells or of metaphase chromosomes affixed to a microscope slide is investigated with the probes and fluorescent reads will be realize through a Fluorescence microscopy. Ratio of target gene copies/centromere copies will be calculated. The advantage of this technique is high specificity and sensitivity compared to classical cytogenetics for detection of chromosomal abnormalities. **Rationale** The extracellular signal-regulated kinase Erk1 and Erk2 cascade regulate multiple cellular processes such as proliferation, differentiation, survival and transformation by phosphorylating multiple target proteins. These kinases are the last component of a signaling module composed of the small GTPase RAS and the protein kinases RAF and MEK1/2. The RAF/Mek/Erk signal transduction pathway is the best characterized MAPK pathway and is activated by growth factors, hormones and cytokines. The ultimate step of this cascade will be the phosphorylation of nuclear and cytoplasmic targets, mediated by the Erk1/2 which will control cellular processes such as proliferation, survival and growth. As this pathway have crucial roles in various physiological processes, its activity also contributes to some of the most prevalent disease and especially to cancer. Indeed, the aberrant activation of this pathway commonly occurs through gain-of-function mutation in genes encoding upstream activators such as RAS and RAF, which are frequently mutated in human cancer. These genes are part of a list of about 150 genes which are thought to be the drivers in the development of cancer processes because their mutations confer a selective growth advantage to the tumor cells. Indeed, their mutation pattern

is highly characteristic and nonrandom. The driver mutations are different from the passenger mutations that have not effect on the neoplastic process because simply they mark the time that has elapsed between successive clonal expansion.

Tumors are addicted to these genetic alterations responsible for receptor activation and continuous expression of their signaling pathways.

Indeed, the oncogenic membrane receptors (ERBB family such as HER2 or HER3 and others) linked to this pathway are often mutated and/or overexpressed in certain types of cancer and contribute to promote tumorigenesis in "in vivo" models. In addition to this pathway, other signal cascades, such as the Pi3K-AKT-mTOR pathway are activated and cross-talk with the RAF/Mek/Erk pathway. The RAS family proteins are direct activator of the PI3K pathways, since the PI3K have an amino terminal RAS-binding domain. Consequence of this talk is the up regulation of PI3K signaling. Also the PI3K/AKT/mTor pathway resulted frequently activated in human cancer (30-50%), thus regulating processes like proliferation, survival and motility. Pathological activation of this via is achieved by I) somatic mutation of the p110 alfa catalytic domain of PI3K; ii) inactivation of the antagonist PTEN lipid phosphatase; iii) amplification of AKT1 (as in gastric cancer). The network between the two signaling pathways realizes a bidirectional communication that allows the continuous cancer cell growth and survival. This interplay may also explain how cancer cells adapt to targeted therapies by using alternative compensatory pathway.

The discovery of the molecular alterations affecting components of these important pathways is one of the greatest achievements of oncology research field since their knowledge have help in the understanding of the tumor biology but also because they have addressed a new molecular definition of different subtypes of cancer among the same histological tumor. In addition, these genetic information have guided the new therapeutic approaches to cancer treatment allowing to the development of the so called targeted therapy in the last decades.

This research project will focus in analyzing the mutations of those driver genes responsible for the growth, survival and transformation of cells. In particular, the mutational analysis of genes along the RTK/ RAF/Mek/Erk and Pi3K-AKT-mTOR pathways in different solid tumors of the Galapagos Island and the Guayas area will be performed.

This study will realize an exhaustive molecular characterization of the different cancer types most frequently encountered in Galapagos Islands and Guayas region. The molecular profiles of each cancer will then be compared among the two Ecuadorian populations and with that of the western countries in order to find if a particular genetic pattern within the various tumor types may explain the observed difference in their incidence rate. Finally, the exact nature of mutations found, i.e. somatic vs. germ line will be addressed by analyzing the corresponding normal cells for that mutation. With this comprehensive molecular information, a better knowledge of the pathogenesis of solid tumors in these areas will be achieved.

Expected results

1. **RAS family genes**. RAS family genes (K-, N-, H- RAS) are the most

	frequently mutated genes in human cancer, with mutations occurring in sporadic colo-rectal cancer, melanoma cancer, lung cancer and thyroid cancer. Hot spots mutation in two out of the three family members, K- and N-RAS in codons 12, 13, 61, 117 and 146 have been demonstrated in about 50-55% of metastatic colo-rectal patients and have been associated with valuable prediction of response to the biological therapy which targets the receptor for the Epidermal Growth factor (EGFR), the antibodies cetuximab or panitumumab. Thus, in the case of mCRC patients, the finding of the molecular signature as the K or N RAS mutation, identify cancers with a genetic feature that has the powerful information to guide the optimal therapy selection. In this sense, RAS mutations are considered a primary molecular RESISTANCE to the anti-EGFR therapy. But also presence of RAS mutations in CRC identifies patients with poor prognosis. Analyzing the same kind of tumors in the Galapagos islands and Guayas area, we may realize what kind of genetic lesion are present locally and if their frequency is comparable to that worldwide. For example, mutations in the KRAS gene exon 2 are the most frequent variations, i.e. 40% of sporadic CRC cases and 80% of them are confined to codon 12 and only 20% affect codon 13. While mutations in KRAS exon 3 and 4 and NRAS exon 2 and 3 are less frequent and account for about 15% of total mCRC samples. So far, mutations in NRAS 4 have not been really quantified, being very infrequent. Thus, our studies might explore different distributions of the mentioned mutations among the tumors arising in Galapagos and Guayas areas, possibly suggesting that other factors might influence the emergence of somatic mutations in the driver genes of solid tumors. In addition, KRAS mutations are also present in about the 20% of lung cancer patients, defining subsets of patients with poor prognosis. The presence of KRAS mutations are usually linked to tobacco smoking behavior, thus incidence of these alteration is higher among cigarettes smokers patients. A higher or lower incidence of lung cancer exhibiting KRAS mutations in these geographical areas could be directly linked to the habits of the people and will give important clues on the mutational hits of the tobacco carcinogens locally. Interestingly, the third member of GTPase RAS family, the H-RAS has been showed to be non-mutated among CRC patients. It will be interesting to investigate if H-RAS mutations may be a characteristic of the CRC arising in the two Ecuadorian areas. In particular, the three gene members of the RAS family K-N-H RAS show a variable rate of mutations in both the intestinal and diffuse types of the sporadic Gastric Cancer (Cosmic Database analysis). Since Gastric cancer is common in low-resources and emerging countries and is frequent in Ecuador, the analysis of RAS mutations together with mutations in PI3K gene (the catalytic subunit p110alfa) and the CTNNB1 would have great importance. Indeed, it seems that RAS mutations express a tissue specificity in GC cases: KRAS mutations are more common together with the HRAS and the PI3K mutations in intestinal phenotype GC; while HRAS mutations seem to be restricted to the diffuse type together with the CDKN2a and CTNNB1 alteration. Whether these heterogeneous genetic patterns are also found among the GC arising in the Galapagos islands and/or the Guayas region is very interesting to investigate and may help in understanding whether this mutational pattern

may be dictated by environmental factor and exposure. In addition, the analysis of mutations in these genes may be related to the presence or the absence of H. pylori that is thought to be one of the major contributors to the development of at least the diffuse GC type in western countries. It is also of note that both the KRAS and the PI3K mutations are present in colo-rectal cancer as in the intestinal GC.

2. BRAF. The v-RAF murine sarcoma viral oncogene homolog B1 is one of the three members of RAF family of MEK and despite the fact that CRAF is the most commonly activated, BRAF is the only member found mutated in human cancer. Melanoma cases are frequently BRAF mutated (about 50% of metastatic melanomas). Indeed, scientific research on this cancer has underlined the extremely high dependency of the neoplastic cells on the activation of growth factor-dependent signaling, governed by the RAS-RAF-MEK-ERK axis. BRAF mutations appear to be more frequent in lesion from skin without chronic sun-induced damage and are mutually exclusive with NRAS mutations in melanoma cases (about 20%). Since BRAF mutations are already present in benign lesion such as melanocytic nevi, it will be interesting to analyzed malignant melanomas and other benign or pre-neoplastic conditions in the Galapagos Islands and Guayas region to eventually explore their incidence together with the rate of BRAF and NRAS mutation in these geographical areas and correlate the data to the risk factor of sun exposure. But also, analysis of BRAF mutations in other cancers such as the sporadic GC will interestingly add new genetic data to the GC puzzle that will help the molecular characterization of this type of cancer locally.

3. Receptor tyrosine Kinase (RTK). The EGFR is the best studied RTK in cancer of epithelial origin such as the lung cancer and the CRC cancer. EGFR, indeed, belongs to the ERBB family of cell-surface receptor that consists of four members, HER1 to HER4. Almost all the gene members of this family are mutated or amplified in a number of different human cancers. Interestingly, mutations in the catalytic kinase domain are almost exclusively of the lung cancer and not in other types of tumors and are therefore a hallmark of a sub population of lung cancer patients with characteristic clinical features: i.e. the EGFR mutated patients have an Adeno-Carcinoma type NSCLC, are young, female and non-smokers. These findings suggest that EGFR mutations are caused by carcinogen(s) other than those contained in tobacco smoke and that usually bear the KRAS mutation as stated above (namely, the squamous type). Asiatic people have a high rate (about 50%) of EGFR mutations among the NSCLC patients compared to that of western countries which a range between 10-16%. Thus, the determination of the frequency, the distribution and the type of EGFR mutations will be of help in subtyping the lung cancer patients of Galapagos Islands and Guayas region. These findings will be then compared with that of the western and Asiatic people and the data obtained may narrow these geographical distinct populations to one or the other ethnic group.

A particular attention will be given to two mutations, the T790M and the S492R, which are both thought to be responsible of the acquired resistance to TKI or the anti-EGFR antibody therapies in lung cancer and colorectal cancer, respectively. It will be interesting to analyze these mutations also in the

	normal cells, since at least the T790M can be demonstrated in hereditary lung cancer syndromes. Amplification of HER member genes is important in some cancer like the sporadic GC where they define a subpopulation of patients with poor outcome and more aggressive phenotype. However, these patients respond to therapy that target the HER2, i.e. the trastuzumab therapy and thus, identifying the GC patients with HER2 amplification is now entered in routine diagnostics at least in western countries and the Chinese population where incidence of GC is very high. Thus, detection of HER2 amplification in GC of these two Ecuador distinct areas will be of interesting to determine at which frequency the gene is amplified compared to western and Asiatic people. In addition, these studies may underline a possibly different correlation in HER2 expression between the intestinal and the diffuse GC types of these areas. **Metabolomic profiles** - The aim of the metabolomic research unit in this proposal is to use NMR spectroscopy for metabolomics studies and characterize the metabolic profiles of cell extracts (both lysates and supernatants) in order to evaluate the presence of variations in profiles of analogous cancers, indigenous and non indigenous, such as those from European countries. Lyophilized extracts from lysate and supernatants will be processed by NMR analysis in the metabolomic research unit of Salento University. Aliquots will be dissolved in D_2O added with 100 mM sodium phosphate buffer, pH 7.4, 0.2% NaN_3, containing 0.5 mM 3-(trimethylsilyl)-[2,2,3,3-$^2H_4$]propionate, TSP, as a chemical shift reference (δ_H = 0 ppm). Samples will be centrifuged at 14,000 *x* g for 5 min at 4°. Six hundred µL of the extract will be placed in a 5 mm outer diameter NMR tube. Samples related to analogous cancers could be differentiated taking advantage of statistical MVA, using supervised and unsupervised methods. This study could allow to assess differences in metabolic pathways for the examined cancer cell cultures. Results obtained from metabolomic and chemometric study could integrate genetic fingerprint. This fingerprint is also linked to the general lifestyle and persisting environmental factors, conditioning gene expression.
(3) Environment and cancer risk	In order to shed light on the complex relationships between environment and health, hypothetically explaining the low occurrence of oncological diseases among the inhabitants of the Galapagos Island, analyses of soils, water, atmosphere, organic molecules and biological elements and of their interactions with each other, are needed. The Galapagos Islands are a group of the most active volcanoes in the world. In spite of the biological interest, their volcanologic, petrologic, and geochemical features are still scarce. In particular, information on abundance of trace elements and on soils is very poor. The mineralogical and chemical investigations on soils and their parent materials should include: a) sampling of soils and their parent rocks in islands having different age, lithology, climate and vegetation; b) mineralogical analyses of the sampled materials; c) chemical analyses of soils and rocks mainly addressed to quantify those elements that, by their high/low concentration, may influence human health in general and cancer risk in particular; d) and chemical analyses for

	evaluating, trough a selective chemical extraction, geochemical mobility and bioavailability of a number of chemical elements. Another important environmental issue related to cancer risk is the potential for toxic effects associated with exposure to heavy metals. This occurrence has been well documented and many studies of heavy metal-induced carcinogenesis show that such exposure may cause a reduction in the expression of DNA repair enzymes, for example through epigenetic mechanisms. Combined with the problem of heavy metal contamination, particularly in marine settings, is their bioaccumulation, biomagnification and persistence in the environment. In this field it could be extremely interesting: i. the identification of polluting metals in water and soil samples from different sites of the Galapagos islands and the Guayas region, as well as from Italian areas with higher/lower incidence of cancer, ii. the quantitative determination of the identified toxic chemical compounds containing metals and of their oxidation states, iii. the determination of the chemical species defined on the base of the procedure of isolation and identification; iv. the chemical speciation and definition at the molecular level of the structure of the polluting compounds formed by toxic metals, v. the determination of the bioavailability of the toxic chemical compounds containing metals by cession tests and vi. The presence and the localization of toxic metals in cells and in organisms. Since the primary intracellular target can drive the cell response (survival, proliferation, death) the localization of metals inside the cells will be investigated by using nano-sized probe microscopies coupled with EDX and cryoTEM-tomography. A geomineralogical aspect that has to be considered for understanding the influence of the environment on human health is the characterization of the atmospheric powder. Analytical investigation should include: i. sampling of dust in different islands; ii. chemical characterization of whole sampling material; iii. mineralogical and chemical characterization of single particles; iv. determination of polycyclic aromatic hydrocarbons (PAHs), volatile organic compounds (VOCs) and other compounds with cancerous features; v. determination of the toxicity induced by the particulate materials (via ROS production and/or ER stress that in turn can cause unbalance of survival and death rates) on in vitro and in vivo models. The Galapagos Islands are characterized by a small number of native plant species (~600), whose biochemical and antioxidant profiles are poorly characterized. Furthermore, the archipelago's microclimate and the soil high salinity may greatly affect the synthesis and accumulation of biologically active molecules within allochthonous fruits and vegetable. Starting from these considerations, the proposed study aims to sample the most common plant species used in the local diet in order to characterize their biochemical diversity and test the antioxidant and anticancer activity of the phytocomplex in normal and pathological cell cultures. In the specific the activities that could be performed include: i) a initial sampling of autochthonous and allochthonous fruits and other plant species used in the traditional diet of Guayas region and Galapagos Islands; ii) a fine structural and biochemical characterization of collected plant material in order to define the exact qualitative and quantitative composition of all the

	biomolecules that are contained within; iii) specific tests (TEAC, ORAC, etc.) to study the antioxidant properties of the plant extracts, iv) mass spectrometry analysis (LCMS and MSMS GCxGC) to complete the screening of potential anticancer and anti-inflammatory substances. By using different methods, including also green and solvent-free technologies, (e.g. supercritical carbon dioxide), matrices and/or extracts from the collected plant species will be prepared and tested on cell cultures and/or preclinical assays in order to better highlight the anticancer and antiproliferative properties. In a subsequent phase, the molecular mechanism by which the extracted bioactive molecules exert the potential antioxidant and antitumoral properties could be analyzed through both in vitro cell cultures and in vivo studies on animal models. These analyses will focus primarily on the effects that these extracts exert on tumor markers expression, on the apoptotic and autophagic processes, lipogenic genes (FAS, ACC, etc.), regulatory proteins of lipogenesis, proteins that regulate vesicular trafficking. Furthermore it will be possible to characterize the chemopreventive and chemosensitizing effects of selected phytochemicals in cell culture and animal models of cancer and experimental strategies such as gene-expression profiling and mass-spectrometry-based methods will be used to aid this phase.
(4) Lifestyle and Nutrition	Cancer incidence and mortality in Galapagos: an area at reduced risk. Why? This project aims to investigate the relationship between the environment in which this population live and their Genome. Air, water, soil**,** foods are the essential elements to which these individuals are exposed or *Exposome* [(1)]. Our behavior influences the choice of a healthy lifestyle or not. These choices correspond to the *Behavome*, and our lifestyle is the result of our behavior in relation to the Exposome in which we live. It will be investigated and compared *lifestyle, nutrition and physical activity,* of a group of people living in the Galapagos Island with that of a sex- and age-matched group of Ecuador inhabitants in order to find out if there is any food or behavior related to lifestyle that can promote or prevent tumor development in this population. As shown by studies in the population of Okinawa, a Japanese island where the inhabitants are living longer and with a lower incidence of cancer, lifestyle, dietary choices and physical activity are most likely responsible for their optimal healthy status (2). In this population, the diet is rich in sweet potatoes, vegetables and soybeans while it is low in animal protein. Okinawa people have active lifestyles and their energy expenditure exceeds their energy intake leading to what has been defined as a status of *undernutrition without malnutrition*. Nutrition and physical activity are the main modifiable components of *energy balance* and determine the percentage of the two constituent of the body (or *body composition*): *fat mass* or adipose tissue and *fat free mass,* essentially skeletal muscles and internal organs (3-4). In the long term an excess of food (or *energy intake*) or a decrease in physical activity (or *activity energy expenditure*) will result in *energy imbalance* that causes a change in body composition with a pathological increase of *fat mass*. Even a sedentary lifestyle can cause an imbalance with a progressive decline

of the *fat free mass*, mainly skeletal muscle mass. Hence an absolute or relative excess of *adipose tissue* in comparison to the fat free mass may increase inflammatory cytokines (IL-6 and TNF-alfa), affect immunity, increase the level or effects of sex hormones, like estrogens and androgens, or of metabolic hormones like insulin and IGF1. As a result the new inflammatory and metabolic status may promote tumor development or growth. Thus food and physical activity can *directly* promote or prevent tumor development or they can change *body composition* thus *indirectly* influencing tumor evolution.

We are confident that the investigation of the lifestyle of the inhabitants of the Galapagos Islands will allow us to understand the reason of the low incidence of cancer in this population.

This detailed program is the following:

NUTRITIONAL STATUS EVALUATION

A. Two **large groups of people** (*size to be defined*), age- and sex-matched, one among the inhabitants of the Galapagos and one among the inhabitants of Ecuador will be subjected to the following investigations:

4. **evaluation of *energy intake* and *diet history*** (eating habits, methods of cooking food):

- food frequency questionnaires of food intake (FFQs) (5)
- dietary history and food habit questionnaires
- blended questionnaires

5. **assessment of *energy expenditure*:**

- estimation of *resting energy expenditure* through predictive equations (6)
- estimation of *activity energy expenditure* through physical activity questionnaires [(7)]

6. **assessment of *body composition:***

- anthropometric measurements (essentially weight, height, waist circumference, hip circumference, arm circumference)

B. Other two groups, formed by a smaller number of people belonging to each population (*size to be defined*), will be more deeply investigated:

6. ***energy intake* evaluation:**

- 24-hour dietary recalls (single or multiple)
- blended questionnaires

7. **assessment of *energy expenditure* in *free living conditions* and lifestyle:**

- measurement of *resting energy expenditure* through *indirect calorimetry (8)*
- assessment of *activity energy expenditure in free-living conditions* thorugh: accelerometers, pedometers, heart monitoring (9)

8. **assessment of *body composition*:**

- bioimpedentiometry

In the same groups of people we will investigate the *metabolic profile*, to assess how the life style and a specific type of diet can modify parameters which usually are correlated to a lower incidence of cancer.

For this reason a series of biochemical parameters will be measured in plasma and urine. Interest will be focused mainly on the evaluation of the *antioxidant*

	and metabolic profiles that could correlate with a lower incidence of cancer. ***BIOCHEMICAL INVESTIGATIONS*** We will proceed with the following *biochemical investigations*: **1. assessment of plasma antioxidant status** • HPLC assay of ascorbic acid/dehydroascorbic acid, carotenoids, vitamin E, flavonoids. **2. assessment of plasma lipid profile and diabetic markers** • FFA profile, total cholesterol, HDL, LDL, Apo-A1, tryglycerides, glycemia, insulin levels, insulin resistance (HOMA test) A second investigation could be the determination of the lipid metabolism and energy metabolism in animals subjected to particular dietary intake. Specifically, it will be determined: • the lipid composition of the liver and plasma; • hepatic lipogenesis; • the oxidation of fatty acids; • mitochondrial respiratory efficiency; • oxidative stress condition. It is also well known that the increment of factors such as the body mass index, dyslipidemia and diabetes mellitus is associated with an increased risk of developing several types of cancer. The purpose of the study is to correlate the individual factors or their combination with the lowest incidence of tumors in the Galapagos area. In particular, it will be explored: • weight and height, important for calculating the body mass index; • waist circumference as predictive parameter for predicting chronic inflammatory diseases; • the plasma level of total cholesterol, triglycerides, high density lipoprotein (HDL), Apolipoprotein AI and any intrinsic modifications (i.e. peroxidative processes) in the lipidic species tested. These latter will be determined by mass spectrometry Other parameters such as blood glucose, insulin and insulin resistance (HOMA test) will be taken into account. **Metabolomic analysis of biological fluids** (plasma, urine, etc.) will be also determined in order to investigate whereas the difference in incidence and cancer mortality among the Galapagos Islands and the nearby areas, are associated with specific metabolic profiles. In particular, the NMR technique, whose peculiarity consists in a high degree of reproducibility of the experiments, allows to acquire a fingerprint of the metabolic profile of any stage of a physiological and/or pathological condition, providing a more comprehensive description of pathophysiological modifications, which are reflected in NMR spectral changes. As well known, all biological systems are easily perturbed by any number of experimental or environmental factors, such as age, diet, growth, nutrients, pH, sex, and temperature. For this reason, the analysis of metabolomic data is further complicated by the inherent variability in each sample, as the number of small-molecule metabolites in a biofluid, cell lysate, tissue or organ differs wildly depending on the organism studied, ranging from several hundred to hundreds of thousands. As a result, the analysis of metabolomic data requires a robust

methodology to expose underlying trends in these highly complex and variable data sets.

• It will be used a metabolic profiler, based on a cryoprobe equipped High field Nuclear Magnetic Resonance Spectrometer 600MHz, in which analytical techniques such as HPLC, NMR and mass spectroscopy are interfaced. The metabolic profiler will be able to examine in depth and characterize all the molecules of interest for this study. Therefore, the metabolic profiles will be integrated with statistical methods able to enhance any physiological modifications caused by particular dietary intake, food habits and the individual factors that influence people lifestyle. In addition, NMR spectroscopic and pattern recognition based methods will be used to investigate the biochemical variability among different metabolic profiles of specific foods which characterize the history and the eating habits of the Galapagos population diet, in order to understand the relationship between food specificity and people life. Thus, a hallmark of metabolic fingerprinting is the use of multivariate analysis methods to identify those biologically relevant spectral features for further targeted analyses, with two of the most popular methods being principal component analysis (PCA) and partial least squares projection to latent structures (PLS). In essence, PCA and PLS aim to differentiate between classes in highly complex data sets, despite within class variability (10).

ANIMAL STUDIES

Rational

Accumulating evidence now suggests that mitochondrial bioenergetics and lipid biosynthesis are required for tumorigenesis. The metabolic change observed in the majority of tumor cells center in large part upon the different utilization of citrate compared to normal cells. It has been proposed that in cancer cells mitochondrial citrate is predominantly exported via CIC (citrate carrier) to the cytoplasm, where it is used as precursor for lipid biosynthesis. This switch from mitochondrial to cytoplasmic metabolism of citrate is thought to account for acquisition of the lipogenic phenotype that, together with the high rates of aerobic glycolysis (Warburg effect) represents a hallmark of cancer cells. It is known that inhibition of CIC hampers cancer cell proliferation through depletion of lipids biosynthesis, leads to mitochondrial dysfunction, in particular declines respiratory complex I (CI) activity, destabilizes mitochondrial membrane potential (MMP) and triggers autophagy clearance of mitochondria. Therefore, emerging studies have begun to demonstrate that targeting mitochondrial metabolism and lipogenesis is potentially a fruitful arena for the treatment of cancer.

Planned experiments

In all organisms liver plays a fundamental role in lipid metabolism, because it is involved in many different processes, such as fatty acids uptake, storage, conversion, oxidation, synthesis and secretion. Therefore, liver samples from rats fed with different experimental diets will be used in this project to evaluate biochemical pathways involved in lipid metabolism.

The following experiments will be performed:

	1.1 Lipid composition of liver. Lipids will be extracted from liver samples using chloroform and methanol. The levels of liver triglycerides, cholesterol and phospholipids will be measured using commercial kits. 1.2 Citrate transport into liver mitochondria and proteoliposomes. Citrate Carrier (CIC) plays an important role in hepatic lipogenesis, because it is responsible for the efflux of acetyl-CoA from the mitochondria to the cytosol in the form of citrate, the primer for fatty acid and cholesterol synthesis. Citrate transport into mitochondria will be measured according to standard procedures well established in our laboratory. The CIC transport reaction will be initiated by the addition of [^{14}C]citrate to malate-loaded mitochondria and terminated by adding the citrate carrier inhibitor 1,2,3-BTA. The expression of the protein will be determined by western-blotting analysis. 1.3 Hepatic enzymes involved in the synthesis of fatty acids. Rat liver cytosol will be obtained by differential centrifugation. Enzymatic activities of Acetyl-CoA Carboxylase (ACC) and Fatty Acid Synthase (FAS) will be determined spectrophotometrically, according to procedures well established in our laboratory. 1.4 Hepatic enzymes involved in the oxidation of fatty acids. Carnitine palmitoyltransferase (CPT) activity will be measured spectrophotometrically in isolated mitochondria. 1.5 Oxidative stress/oxidative damage. Lipid peroxidation will be analyzed with a commercial kit. This is a sensitive and reliable assay for the measurement of the hydroperoxides from any sample containing lipid hydroperoxides, directly utilizing the redox reactions with ferrous ions. The resulting ferric ions are detected using thiocyanate ion as the chromogen and quantified spectrophotometrically. Protein peroxidation will be analyzed with a commercial kit. This kit provides the chemical and immunological reagents necessary to perform the immunoblot detection of carbonyl groups introduced into proteins by oxidative reactions with ozone or oxides of nitrogen or by metal catalyzed oxidation. 1.6 Mitochondrial respiration efficiency. Rat liver mitochondrial respiration will be measured by a Clark oxygen electrode. The addition of different substrates will permit to evaluate mitochondrial respiration when respiratory complexes I (pyruvate, malate), II (succinate, rotenone), and IV (ascorbate, TMPD, rotenone, antimycin A) were stimulated. Instead, the addition of palmitoyl-L-carnitine and malate will permit to reveal oxygen consumption from β-oxidation pathway. For each substrate, after 2 min, state 3 respiration was induced by the addition of ADP. Respiratory control ratio (RCR) was calculated as the ratio of the rate of oxygen uptake in the presence of added ADP (state 3) to the rate observed when added ADP had been completely phosphorylated to ATP (state 4). 1.7 Mitochondrial respiration chain enzyme activities. Respiratory chain complexes activities will be measured spectrophotometrically according to procedures well established in our

	laboratory. ***Expected outputs*** The identification of possible molecular targets involved in the diet modulation of mitochondrial metabolism and lipogenesis will eventually translate into the development of novel cancer treatment strategies.
Informatics	The investigation about the epidemiological and environmental causes behind the low cancer incidence in the Galapagos Island population requires the development of a comprehensive approach to better define which aspects could represent cancer risk factors. Such a global approach is mainly based on the possibility to have at disposal all data characterizing the Galapagos population (endogenous and exogenous factors), but the solely availability is not sufficient for analysis purposes. The proper design of a data infrastructure and the adoption of multidimensional analysis techniques together with appropriate web interfaces and services, will allow to access, visualize and analyze such data in order to evaluate all cause-and-effect relationships and all potential correlations existing among Galapagos population and other similar ones (when available) in terms of lifestyle, food, weather conditions, etc. Machine learning systems will be adopted in order to train the system and improve its response to newest data. Moreover, it will be carried out a spatial analysis on local areas and people groups to map the incidence of different types of cancer and a correlation with both environmental and genetic factors. In a first step, GIS (Geographical Information System), spatial epidemiology, and hot-spot analyses will allow to identify geographic hot spots, in which the incidence of a specific tumor is particularly high. In a second step, data analytics and data mining techniques will be used to correlate these epidemiological data with environmental factors such as the temperature, the altitude, slope, the distance from waterways, parks, roads, etc. These approach will be also applied to correlate the incidence of cancer cases with the genetic pattern of the local people both healthy subjects and patients suffering by a given pathology. From an architectural point of view, this pose the need for appropriately choose and design the systems and techniques that will be taken into account, such as relational databases for the storage aspect, data warehouse for multidimensional analysis and data analytics and cloud infrastructures for storage and accessing purposes. When dealing with personal data, special care has to be paid to issues related to data confidentiality and privacy protection. In this context, anonimization techniques will be adopted to make possible the open dissemination of such data in full compliance with privacy and data security.
A2.6. Expected Results	
This research may shed new light into the relationship between particular environmental influences (diet and life style) and genetic alteration of the solid tumors in the Galapagos Island and Guayas territory. Indeed, some of the somatic mutations affecting the crucial genes under investigation may be classified as germ line mutation if already present in the normal cells, thus helping in a better understanding of their contribution to the pathogenesis of solid tumors in these areas. In	

addition, these data could eventually explain differences between Galapagos and Guayas people themselves and with that of the western populations.

A2.7. References

A2.7.1 Cancer Epidemiology

A2.7.2 Cancer Genetics

- Carlomagno F and Chiarello M Growth factor transduction pathways: paradigm of anti-neoplastic targeted therapy. J Mol Med, 92: 723 (2014).
- Ahmed A. Samatar and Poulikos I. Poulikakos**.** Targeting RAS–ERK signalling in cancer: promises and challenges Nature Rev 13:928 (2014).
- David A. Fruman and Christian Rommel. PI3K and cancer: lessons, challenges and opportunities. Nature Rev 13:141 (2014).
- Vogelstein B et al. Cancer Genome Landscapes Science 339:1546 (2013).
- Vivanco I Targeting molecular addictions in cancer. Br J Cancer 111:2033 (2014).
- Garraway LA and Lander ES (2013) Lessons from the cancer genome Cell 153:17
- Douillard J-Y et al. Panitumumab-FOLFOX treatment and RAS mutations in Colorectal cancer NEJM 369:1023 (2013).
- John R.Goffin and Kevin Zbuk. Epidermal Growth Factor Receptor: Pathway, Therapies, and Pipeline Clinical Therapeutics Volume 35: 1282 (2013).
- Wendy A. Cooper, David C. L. Lam, Sandra A. O'Toole, John D. Minna. Molecular biology of lung cancer J Thorac Dis 2013;5(S5):S479-S490 (2013).
- McLean MH and El-Omar EM. Genetics of gastric cancer Nature Rev 11:664 (2014).
- Gunturu KS et al. Gastric cancer and trastuzuma: first biologic therapy in gastric cancer Ther Adv Med Oncol 5:143 (2013).
- Lo JA, Fisher DE. The melanoma revolution: from UV carcinogenesis to a new era in therapeutics. Science. Nov 21;346(6212):945-9 (2014)
- Hagemann IS et al. Design of targeted, capture-based, next generation sequencing tests for precision cancer therapy Cancer Genetics 206:420 (2014)
- Zhang L et al. Profiling cancer gene mutations in clinical Formlin-Fixed, Paraffin-Embedded Colorectal tumor specimens usimg targeted next-generation sequencing The Oncologist 19:1 (2014).

A2.7.3 Environment and Cancer Risk

A2.7.4 Nutrition and Lifestyle

1. Wild, C. P. Complementing the Genome with an 'Exposome': The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology. Cancer Epidemiol. Biomarkers Prev. 14, 1847–1850 (2005).
2. Willcox, D. C., Willcox, B. J., Todoriki, H. & Suzuki, M. The Okinawan diet: health implications of a low-calorie, nutrient-dense, antioxidant-rich dietary pattern low in glycemic load. J. Am. Coll. Nutr. 28 Suppl, 500S–516S (2009).
3. Ellis, K. J. Human body composition: in vivo methods. Physiol. Rev. 80, 649–680 (2000).
4. Norgan, N. Laboratory and field measurements of body composition. Public Health Nutr. 8, 1108–1122 (2005).
5. Thompson, F. E. & Subar, A. F. in Nutrition in the Prevention and Treatment of Disease 5–46 (Elsevier/AP, 2013).
6. FAO. Human energy requirements, Report of a Joint FAO/WHO/UNU Expert Consultation.

Food and Nutrition Technical Report Series. Rome, 17–24 October 2001.
7. Lagerros, Y. T. & Lagiou, P. Assessment of physical activity and energy expenditure in epidemiological research of chronic diseases. Eur. J. Epidemiol. 22, 353–362 (2007).
8. Matarese, L. & Fada, R. Indirect Calorimetry: Technical Aspects. J. Am. Diet. Assoc. 97, S154–S160 (1997).
9. Bouten, C. V., Westerterp, K. R., Verduin, M. & Janssen, J. D. Assessment of energy expenditure for physical activity using a triaxial accelerometer. Med. Sci. Sports Exerc. 26, 1516–1523 (1994).
10. Worley, B., Powers, R. Multivariate analysis in metabolomics. Current Metabolomics, 1(1), 92-107 (2013).

A3	**PRINCIPAL INVESTIGATORS**	
Study Chairman	***Organisation***	***Contact e-mail***
Massimo Federico M.D Chair of Medical Oncology, Department of Diagnostic Medicine, Clinical Medicine and Public Health,	University of Modena and Reggio Emilia, Modena and Modena Cancer Registry	massimo.federico@unimore.it
Study Co-Chairman	***Organisation***	***Contact e- mail***
Anna Iannone MD, PhD, Chair of Physiopathology, Department of Diagnostic Medicine, Clinical Medicine and Public Health	University of Modena and Reggio Emilia, Modena, Italy	anna.iannone@unimore.it
Michele Maffia Phd, Chair of Physiology, department of Biological and Environmental Sciences and Technologies	University of Salento, Lecce, Italy	michele.maffia@unisalento.it
Teodoro Chisesi MD, Coordinator	Comitè de Investigacion, CISOL, SOLCA Guayaquil, Ecuador.	t.chisesi@libero.it

Informatics	
Referents	**Organisation**
Michele Colajanni (TL) CRIS	"Centro di ricerca interdipartimentale per la prevenzione dei rischi" Dep. of Engineering "Enzo Ferrari", University of Modena and Reggio Emilia, Italy
Giovanni Aloisio	University of Salento, Lecce, Italy
Mario Bochicchio	University of Salento, Lecce, Italy

Project managers	
Referents	**Organisation**
Anna Fedina	Department of Diagnostic Medicine, Clinical Medicine and Public Health, University of Modena and Reggio Emilia, Modena, Italy
Roberta Mele DReAM	"Laboratorio Diffuso di Ricerca", University of Salento, Lecce, Italy
Luana Conte DReAM	"Laboratorio Diffuso di Ricerca", University of Salento, Lecce, Italy

A4	**TASK FORCES REFERENTS**
Cancer Epidemiology	
Scientific Referents	**Organisation**
Paolo Bruzzi (TL) Director of Clinic Epidemiology	IRCCS AUO San Martino IST – Genova
Massimo Federico	Department of Diagnostic Medicine, Clinical Medicine and Public Health, University of Modena and Reggio Emilia, Modena and Modena Cancer Registry
Leyda Jaramillo	Guayaquil Cancer Registry
Rina Quinto	Guayaquil Cancer Registry
Luigi Marcheselli	Department of Diagnostic Medicine, Clinical Medicine and Public Health, University of Modena and Reggio Emilia, Modena and Modena Cancer Registry
Antonella De Donno	University of Salento, Lecce, Italy
Marcello Guido	University of Salento, Lecce, Italy
Cancer Genetics	
Scientific Referents	**Organisation**
Manlio Ferrarini (TL) Scientific Director	IRCCS AOU San Martino – IST- Genova
Maria Dono	IRCCS AOU San Martino – IST- Genova
Robin Foà Chair Director	UO Haematology, University "La Sapienza", Rome
Simona Zupo	Oncology, IRCCS AOU San Martino – IST- Genova
Pierfrancesco Tassone	Università Magna Grecia, Catanzaro
Maria Teresa Di Martino	Università Magna Grecia, Catanzaro
Maria Pia Bozzetti	University of Salento, Lecce, Italy
Tiziano Verri	University of Salento, Lecce, Italy
Vincenzo Zara	University of Salento, Lecce, Italy
Nutrition and Lifestyle	

Scientific Referents	Organisation
Massimo Pellegrini (TL)	Department of Diagnostic Medicine, Clinical Medicine and Public Health, University of Modena and Reggio Emilia, Modena , Italy
Anna Iannone	Department of Diagnostic Medicine, Clinical Medicine and Public Health, University of Modena and Reggio Emilia, Modena , Italy
Nino Carlo Battistini	Department of Diagnostic Medicine, Clinical Medicine and Public Health, University of Modena and Reggio Emilia, Modena , Italy
Sebastiano Banni	Department of Biomedical Science, University of Cagliari
Francesco Paolo Fanizzi	University of Salento, Lecce, Italy Metabolomic Unit
Ferramosca Alessandra	University of Salento, Lecce, Italy
Environment and cancer risk	
Scientific Referents	**Organisation**
Giovanni Natile (TL) Director of CIRCMSB	"Consorzio Interuniversitario di Ricerca in Chimica dei Metalli nei Sistemi Biologici", University of Bari
Marcello Lenucci	University of Salento, Lecce, Italy
Luisa Siculella	University of Salento, Lecce, Italy
Michele Maffia	University of Salento
Fabio Arnesano CIRCMSB	(Consorzio Interuniversitario di Ricerca in Chimica dei Metalli nei Sistemi Biologici), University of Bari
Francesco Cannito CIRCMSB	(Consorzio Interuniversitario di Ricerca in Chimica dei Metalli nei Sistemi Biologici), University of Bari
Saverio Fiore IMAA	"Istituto di Mineralogia per l'Analisi Ambientale", CNR, Italy
Cosimino Malitesta	University of Salento, Lecce, Italy
Giuseppe De Benedetto	University of Salento, Lecce, Italy
Alessandra Genga	University of Salento, Lecce, Italy
Luciana Dini	University of Salento, Lecce, Italy
Cecilia Bucci	University of Salento, Lecce, Italy
Anna Giudetti	University of Salento, Lecce, Italy
Loredana Capobianco	University of Salento, Lecce, Italy
Francesco Paolo Fanizzi	University of Salento, Lecce, Italy

A5	**BUDGET (in Euros)**
A5.1. Cancer epidemiology	

A5.2 .Cancer Genetics	
A5.3. Environment and Cancer Risk	
A5.4. Nutrition and Lifestyle	
A5.5. Informatics	
A5.6. Study Coordination	
TOTAL (A5.1 + A5.2 +A5.3 + A5.4 + A5.5)	**€**