UNIVERSITY OF NAPLES "FEDERICO II"

DEPARTMENT OF VETERINARY MEDICINE AND ANIMAL PRODUCTION

VETERINARY TEACHING HOSPITAL
INTERNAL MEDICINE UNIT

PhD thesis in Veterinary Science

Next-Generation Sequencing for the diagnosis of canine myocardial disorders: role of pathogens and study of gene expression pathways

Coordinator	**Candidate**	**Tutor**
Prof. Giuseppe Cringoli	Dr Carlo Ferravante	Prof. Antonio Di Loria

XXXIV CYCLE

To the big tree and his wondering garden.

List of contents

ACKNOWLEDGEMENTS

I would like to express my gratitude to several people who supported me during the Ph.D. experience. My special acknowledgments go to my primary supervisor, prof. Antonio Di Loria, who guided me throughout this project and encouraged me every day. I would like to thank Prof. Paolo Ciaramella, for always being present during this period with his support and precious suggestions. I wish also to extend my appreciation to Prof. Roberto Santilli, Dr. Manuela Perego, Prof. Elena Grego and Prof. Massimiliano Tursi for their contribution to this project and their professional advice. I would like to express my gratitude to Prof. Alessandro Weisz from the University of Salerno, for welcoming me into his laboratory, for having continually stimulated me, directly and indirectly, to improve my skills and my professionalism. Special thanks to prof. Alberto Riva from the l'Interdisciplinary Center for Biotechnology Research (ICBR) of the University of Florida, for his openness and patience during our hours of remote working. My special acknowledgments go to Dr. Giorgio Giurato, my mentor, and all the bioinformatics team, for all the wise advice, criticism, and for words of comfort that have helped me to grow. Thanks for having been close to me all these years. I expand my thanks to all past and present members of the Laboratory of Molecular Medicine and Genomics of the University of Salerno and the Genomix4Life team, with whom I shared stressful moments, but also lightheartedness and joy. Thanks for teaching me what working hard with loyalty means, without ever losing the desire to smile.

But the biggest thanks goes to Federica, and to my big family, for always believing in me, for being always solid support in my life, and for having always been close to me with unconditional love.

List of abbreviation

WHO	World Health Organization
ISFC	International Society and Federation of Cardiology
DCM	Dilated Cardiomyopathy
EMB	Endomyocardial Biopsy
GCM	Giant Cell Myocarditis
ECG	Electrocardiogram
PCR	Polymerase Chain Reaction
NGS	Next Generation Sequencing
mNGS	Metagenomic Next Generation Sequencing
COVID-19	CoronaVirus Disease 19
CMR	Cardiac Magnetic Resonance
CBC	Common Blood Count
RT	Reversible Termination
qRT-PCR	quantitative Reverse Transcription PCR
DGE	Differential Gene Expression
WMG	Whole Metagenome
RNA-seq	RNA sequencing
RV	Right Ventricular
RIN	RNA Integrity Number
DE	Differentially expressed
FDR	False Discovery Rate
GO	Gene Ontology
DEA	Differential Expression Analysis
KEGG	Kyoto Encyclopaedia of Genes and Genomes
GAS	Group A Streptococcus
AP	Action Potential
PCA	Principal Component Analysis

List of figures

List of tables

ABSTRACT

Cardiomyopathy, in clinical medicine, defines the primary myocardial disease triggered by an unknown etiology. The secondary myocardial disease may be described as some often reversible, inflammatory, metabolic, toxic, or infiltrative disease of the myocardium with a known etiology or causative agent. In this context, it is possible to differentiate myocardial disease in forms related to myocarditis, frequently associated with pathogens infection, from other forms secondary to endocrine abnormalities, electrolyte imbalances, trauma and neoplasias. Myocarditis can evolve in dilated cardiomyopathy, with consequent devastating effects on dogs' health. Despite the problematic clinical conditions that characterize canine myocarditis, few data are present in the veterinary literature. Nonetheless, in recent years, more attention has been focused on this heart muscle disease, probably due to the advances in diagnostic technology and the more frequent use of endomyocardial biopsies that require advanced cardiology division heart failure programs, the technical capabilities and the expertise to appropriately analyze the specimens. Analyzing tissue specimens with more sophisticated methods, rather than routine staining used in light microscopy, may improve the diagnostic yield and clinical and research utility of endomyocardial biopsies. In this context, Next-Generation Sequencing (NGS) may be considered a supreme advantage. This high-throughput technology can be used to achieve a comprehensive and unbiased sequencing of the nucleic acids present in a sample. The sequencing of the whole nucleic acids content of a specimen enables the exploration of both host and exogenous DNA/RNA sequences. The large amount of data produced by the NGS approach, requires comprehensive and user-friendly pipelines for data analysis, that speed up the bioinformatics steps.

In this thesis, the transcriptomic responses of the dogs to myocardial disorders were investigated along with the role of pathogens in myocardial disease. The sequencing of total RNA molecules from endomyocardial biopsies of 13 dogs affected with myocardial disorders, was performed using NextSeq 500 Illumina platform. Each endomyocardial biopsy was histologically evaluated, to reveal the presence of inflammation signs. Histological diagnosis of fibrosis was identified in 10/13 animals enrolled in this study. Three dogs with fibrosis showed also the presence of inflammatory infiltrates.

In order to discover the presence of pathogens in cardiac tissues, a metagenomic approach was performed on RNA-seq data. Considering the lack of exhaustive and easy-to-use bioinformatics pipelines publicly available for the identification of pathogens, a new software (HOME-BIO (sHOtgun MEtagenomic analysis of BIOlogical entities)) was developed to provide a comprehensive taxonomy profiling of the endomyocardial biopsies. The

metagenomic analysis results showed no presence of cardiotropic pathogen commonly related to myocarditis and myocardial damage, associated with a high number of RNA fragments. The expression profiles of enrolled samples were compared with those of the control healthy group, downloaded from public dataset (NCBI BioProject PRJNA78827), in order to identify cardiac tissue genes and gene pathways differentially regulated between the 2 populations. The analysis highlighted genes enriched in the dilated cardiomyopathy disease pathway and involved in cellular energy metabolism, along with altered cardiac structural proteins in affected dogs. Moreover, important results were achieved from the analysis of genes that characterized samples with clear signs of inflammation. In these samples, differentially expressed genes were involved in cardiac altered activity, as well as in structural cell reorganization and metabolism. By evaluating the different pathways and genes involved during myocardial disease, the results revealed molecular and genetic events that could play a key role in the progression of myocardial disorders and in particular may provide important insights into the pathogenesis of dilated cardiomyopathy.

.

Myocardial disease indicates acquired damage or metabolic derangement of the myocardium. Frequently, the term myocardial disease is often used interchangeably with the term cardiomyopathy in clinical medicine, although the latter should be reserved for the primary myocardial disease with an unknown etiology (idiopathic), with a strong suspicion or a genetic cause. Secondary myocardial disease may be defined as some often reversible, inflammatory, metabolic, toxic, or infiltrative disease of the myocardium with a known etiology or causative agent. The suspicion of myocardial disease includes a history of recent onset of exercise intolerance or weakness, from mild to most severe, diminished cardiac performance and frequently cardiac arrhythmias. The clinical signs may have been chronic and progressive, or are of acute onset, depending on the etiology. In this context, it is possible to differentiate myocardial disease in forms related to myocarditis frequently associated with pathogens infection from other forms secondary to endocrine abnormalities, electrolyte imbalances, trauma and neoplasias. As defined in 1995 by the World Health Organization (WHO)/International Society and Federation of Cardiology (ISFC), myocarditis is an inflammatory disease of the heart muscle, diagnosed considering histological, immunological, and immunohistochemical criteria. Thanks to the diagnostic advances achieved, the term has been refined, and presently applies to acute or chronic inflammatory responses of the heart to environmental or endogenous triggers [Heymans et al., 2016]. For this reason, if the pathological presentation of fulminant myocarditis is closely related to an infiltrative inflammatory picture, during a chronic evolution different pathological polymorphisms can be reached in the pathological tissue. Myocarditis can evolve in cardiomyopathy and be viewed as a chronological sequence of three pathogenetically distinct phases: in the first phase an initial insult to the myocardium occurs, (viral infection in the great majority of cases) usually may go unnoticed; after the resolution of the initial injury, the second phase develops as a result of autoimmunity triggered by that injury; in the third phase, a typical picture of dilated cardiomyopathy (DCM) develops and may progress despite the cessation of the first two processes. Clinically overt congestive heart failure often develops in this phase as a result of extensive additional myocardial injury. Although myocarditis is, by definition, an inflammatory disorder and DCM is in most cases idiopathic, accumulating data has revealed an important inflammatory component in the pathogenesis of cardiomyopathy, considering them both closely related. The immunity response leads to a process of chronic autoantigen-driven inflammation, which can progress to dilative cardiomyopathy and end-stage heart

failure [Mahrholdt et al., 2006, Cunha-Neto et al., 2006]. During the transition from acute inflammation to chronic fibrosis, inflammation can persist and symptoms can be absent.

Inflammatory infiltration of the myocardium may develop based on three different etiopathogenic mechanisms: idiopathic, non-infectious (e.g. autoimmune) and infectious. If idiopathic myocarditis may occur when the causative agent remains unidentifiable, among non-infectious causes, Giant cell myocarditis (GCM) and cardiac sarcoidosis are well-documented specific forms only for human myocarditis. In non-infectious forms, myocarditis is frequently related to inflammatory responses of the heart to environmental or endogenous triggers. Toxic myocarditis may be triggered by several prescribed drugs such as dobutamine, phenytoin, antibiotics (e.g., ampicillin, azithromycin, cephalosporins and tetracyclines), psychiatric medications (tricyclic antidepressants, benzodiazepines and clozapine), illicit drugs (e.g. methamphetamine or cocaine), heavy metals (copper, lead and arsenicals) and antineoplastic agents (e.g., anthracyclines, cyclophosphamide, 5-fluorouracil and tyrosine kinase inhibitors). In the last years, the increasing use of Immune-Checkpoint Inhibitors to treat a variety of cancers was reported to lead to lethal myocarditis as a potential adverse effect of therapy [Mahmood et al., 2018].

The existence of an inflammatory process or a myocardial damage in the heart muscle, related to a progressive worsening of myocardial function, has been known for at least two centuries, but only recently there has been a greater interest in the etiopathogenetic processes and related clinical-therapeutic implications [Heymans et al., 2016]. This emerging attention is probably due to multiple factors such as the introduction of endomyocardial biopsy (EMB) for in vivo diagnosis, the study of new standard criteria for histological diagnosis and the identification of etiological agents. The introduction of EMB as gold-standard for diagnosis of myocardial damage and in particular of myocarditis using molecular techniques, has certainly contributed to better defining the physio-pathological bases of the inflammatory process, furthermore, it helped to clarify the histological and immunohistochemical aspects of the illness [Tschöpe et al., 2021]. Despite the clear definition and the recent evolution of molecular and immunohistochemical investigation, natural history, classification, diagnosis and treatment of myocarditis remain controversial. This confusion lay, partially, in the high clinical polymorphism of the disease, the broad etiopathogenetic spectrum, and the lack of precise guidelines regarding the assessment, prognostic stratification and his treatment.

In dogs the diagnostic approach of myocarditis appears more difficult and oriented when it is possible to perform EMB, to confirm or exclude an infectious etiology; in this sspecies studies relating to non-infectious or idiopathic causes resulted until now absent or fragmentary.

In dogs, myocarditis is related to different pathogen infections [Molesan et al., 2017], but in clinical practice, the underlying cause is not even detected, probably due to the few confirmatory clinical tests for canine myocarditis. When infectious myocarditis is considered, except for South America, where *Trypanosoma cruzi* is the main etiologic agent (Chagas disease), viral and post-viral myocarditis remains the major cause of acute myocarditis in humans such as in dogs [Heymans et al., 2016, Molesan et al., 2019]. Cardiac involvement during viral infection is usually supported both by host characteristics (genetic, immunological factors, age) and by the virulence of the infecting agent. In humans, infectious etiology can be supported by different viruses: Enterovirus (EV), Adenovirus (ADV), Coxsackievirus, Cytomegalovirus (CMV), Epstein-Barr virus (EBV), HIV-1 [Imanaka-Yoshida et al., 2020].

Myocarditis incidence in the canine population is unknown, although a histopathologic diagnosis of myocarditis was reported in 1.5% of dogs in a retrospective necropsy study [Lakhdhir et al., 2020]. Literature pertaining to the presentation, diagnosis, and treatment of myocarditis in canines is limited, consisting mostly of single-case reports and small case series focused on a single causative etiology. At the physical examination, the most common presenting findings recorded in dogs with myocarditis, include lethargy and hyporexia, fever, heart murmur, and pulse abnormalities. Complete blood count and biochemistry abnormalities usually are consistent with non-specific inflammation (neutrophilia, left shift, monocytosis, and anemia) or decreased organ perfusion (elevated ALT and azotemia). Ventricular ectopy is common on ECG, indeed dogs exhibiting ventricular premature complexes although ventricular tachycardia can be also detected. The most noted echocardiographic abnormality is represented by a decreased left ventricular systolic function (DCM-like), evident in more than 50% of cases reported in the literature [Lakhdhir et al., 2020]. Although in the last years some veterinary clinic centers start to include EMB during diagnostic cardiological procedures, the antemortem diagnosis of canine myocarditis is generally presumptive and based on an individual clinician's interpretation of case presentation and clinicopathologic information. Considering the high mortality that distinguishes infectious myocarditis, achieving early diagnosis becomes extremely important. Most of the diagnoses of myocarditis in dogs are made post mortem, although when the clinician has clinical suspicion of myocarditis, an endomyocardial biopsy should be recommended [Santilli et al., 2017]. Ideally, this procedure

should be performed at centers with advanced cardiology division heart failure programs that have the technical capabilities and the expertise to appropriately analyze the specimens. Analyzing tissue specimens with more sophisticated methods, rather than routine staining used in light microscopy, may improve the diagnostic yield and clinical and research utility of EMB. The development of highly sensitive molecular biological methods such as in-situ hybridization and polymerase chain reaction (PCR) enabled to detect nucleic acids of selected pathogens in canine endomyocardial biopsies for in-vivo investigation. EMB samples of dogs with unexplained myocardial and rhythm disorders showed the presence of viral nucleic acid (NA) via PCR, similar to observations reported in humans. This first study in vivo confirms the presence of viral and/or bacteria genomes and infiltration of inflammatory cells in suspected dogs with myocarditis, while no pathogens and no inflammatory signs were detected in tissue obtained from control dogs [Santilli et al., 2019]. However, other authors, despite a comprehensive diagnostic screening, reported that the cause of myocardial lesions remained frequently undetermined, demonstrating the limitations of available molecular tests. Agent detection by PCR, indeed, did not imply causation, Molesan and colleagues showed how the presence of M. haemocanis, canine adenovirus (CAV), canine respiratory coronavirus (CRCoV), canine parainfluenza virus (CPiV), canine distemper virus CDV, and canine herpesvirus (CHV) were not associated with myocarditis as detection occurred with equivalent frequency in myocarditis cases and controls. The authors reported how many of the PCR-detected pathogens (CDV, CHV, and CAV) are associated with subclinical, latent, or otherwise inapparent infection, particularly in adult animals [Molesan et al., 2019]. As immunization with modified live vaccine (CDV, CPV-2, CAV-2, and CPiV) is recommended for dogs beginning as early as 6 weeks of age, PCR may be detecting vaccine virus and not field strains. In this context, PCR-based methods could represent a limit because an oriented selection of pathogens could exclude other possible agents of myocardial infection. In recent years, Next-Generation Sequencing (NGS) technology has become a very attractive instrument, changing the landscape of diseases and giving the opportunity to obtain a great amount of information in an accelerated time and at relative low costs. This high-throughput methodology, supporting a broad range of usages, enabled the fast sequencing (or “reading”) of the bases in DNA or RNA samples. This technology and bioinformatics analyses generated data, offered a promising strategy for the rapid identification of nucleic acids related to pathogens in clinical settings. Compared with other standard methods, NGS may be considered a supreme advantage to the discovery of novel infectious agents and characterize different known pathogens simultaneously. Traditional

methods for microorganism's isolation, typically lack the sensitivity required for the detection of pathogens that could be present at low abundance [Liu et al., 2011]. The high-throughput technology can be used to achieve a comprehensive and unbiased sequencing of the nucleic acids present in a sample (i.e. tissues and/or biological fluids). Metagenomics NGS (mNGS), the comprehensive analysis of microbial and host genetic material, has emerged as a powerful approach to analyze and survey microbiota composition in the field of infectious diseases [Franzosa et al., 2015]. By directly sequencing millions of DNA/RNA molecules in a sample and matching the sequences to those available in databases, pathogens of an infectious disease can be inferred. Furthermore, the sequencing of whole nucleic acids content of a specimen enables the exploration of both host and exogenous DNA/RNA sequences. This large amount of data produced, requires comprehensive and user-friendly pipelines for data analysis, that speed up the bioinformatics steps. Data about this aspect in veterinary medicine are not investigated yet.

Currently, most of the knowledge about the mechanisms underlying myocarditis is provided by experiments on rodent susceptible models as mentioned above. In these animals, the inflammation is caused by experimental infection with cardiovirulent pathogens (virus, bacteria or parasites), or induced with cardiac antigens and strong adjuvant. Unlike the rodent model, canine species, sharing with humans some pathogens related to myocarditis at risk of zoonosis, represents the ideal animal model of study. The dog, due to its close and ancestral relationship with humans, is also a natural infection model that may include the environmental risk factors common to the two species.

The aim of this thesis is to study the role of pathogens and gene expression pathways: i) considering for the first time NGS technology on EMB samples obtained in vivo in dogs with myocardial disorders; ii) developing a taylored pipeline for bioinformatics analysis of data obtained by metagenomic NGS approach; iii) valuating the different pathways and genes involved during myocardial disease.

The experimental part of this thesis will be preceded by a general part that performs an in-depth discussion of the studies performed on pathogens involved in miocardial disorders and the new diagnostic method emploing NGS.

Cunha-Neto E, Bilate AM, Hyland KV, Fonseca SG, Kalil J, Engman DM. Induction of cardiac autoimmunity in Chagas heart disease: a case for molecular mimicry. Autoimmunity. 2006 Feb;39(1):41-54. doi: 10.1080/08916930500485002. PMID: 16455581.

Franzosa EA, Hsu T, Sirota-Madi A, Shafquat A, Abu-Ali G, Morgan XC, Huttenhower C. Sequencing and beyond: integrating molecular 'omics' for microbial community profiling. Nat Rev Microbiol. 2015 Jun;13(6):360-72. doi: 10.1038/nrmicro3451. Epub 2015 Apr 27. PMID: 25915636; PMCID: PMC4800835.Cooper LT Jr. Myocarditis. N Engl J Med. 2009 Apr 9;360(15):1526-38. doi: 10.1056/NEJMra0800028. PMID: 19357408; PMCID: PMC5814110.

Heymans S, Eriksson U, Lehtonen J, Cooper LT Jr. The Quest for New Approaches in Myocarditis and Inflammatory Cardiomyopathy. J Am Coll Cardiol. 2016 Nov 29;68(21):2348-2364. doi: 10.1016/j.jacc.2016.09.937. PMID: 27884253.

Imanaka-Yoshida K. Inflammation in myocardial disease: From myocarditis to dilated cardiomyopathy. Pathol Int. 2020 Jan;70(1):1-11. doi: 10.1111/pin.12868. Epub 2019 Nov 5. PMID: 31691489.

Lakhdhir S, Viall A, Alloway E, Keene B, Baumgartner K, Ward J. Clinical presentation, cardiovascular findings, etiology, and outcome of myocarditis in dogs: 64 cases with presumptive antemortem diagnosis (26 confirmed postmortem) and 137 cases with postmortem diagnosis only (2004-2017). J Vet Cardiol. 2020 Aug;30:44-56. doi: 10.1016/j.jvc.2020.05.003. Epub 2020 May 29. PMID: 32668360; PMCID: PMC7256493.

Liu S, Vijayendran D, Bonning BC. Next generation sequencing technologies for insect virus discovery. Viruses. 2011 Oct;3(10):1849-69. doi: 10.3390/v3101849. Epub 2011 Oct 10. PMID: 22069519; PMCID: PMC3205385.

Lombard CW Differentiating Primary Myocardial from Secondary Myocardial Diseases in Dogs. World Small Animal Veterinary Association World Congress Proceedings. Vetsuisse Faculty, University of Bern, Bern, Switzerland, 2011.

Mahmood SS, Fradley MG, Cohen JV, Nohria A, Reynolds KL, Heinzerling LM, Sullivan RJ, Damrongwatanasuk R, Chen CL, Gupta D, Kirchberger MC, Awadalla M, Hassan MZO, Moslehi JJ, Shah SP, Ganatra S, Thavendiranathan P, Lawrence DP, Groarke JD, Neilan TG. Myocarditis in Patients Treated With Immune Checkpoint Inhibitors. J Am Coll Cardiol. 2018 Apr 24;71(16):1755-1764. doi: 10.1016/j.jacc.2018.02.037.

Mahrholdt H, Wagner A, Deluigi CC, Kispert E, Hager S, Meinhardt G, Vogelsberg H, Fritz P, Dippon J, Bock CT, Klingel K, Kandolf R, Sechtem U. Presentation, patterns of myocardial damage, and clinical course of viral myocarditis. Circulation. 2006 Oct 10;114(15):1581-90. doi: 10.1161/CIRCULATIONAHA.105.606509. Epub 2006 Oct 2. PMID: 17015795.

Molesan A, Goodman L, Ford J, Lovering SJ, Kelly K. The Causes of Canine Myocarditis and Myocardial Fibrosis Are Elusive by Targeted Molecular Testing: Retrospective Analysis and Literature Review. Vet Pathol. 2019 Sep;56(5):761-777. doi: 10.1177/0300985819839241.

Santilli RA, Battaia S, Perego M, Tursi M, Grego E, Marzufero C, Gianella P. Bartonella-associated inflammatory cardiomyopathy in a dog. J Vet Cardiol 2017;19:74-81.

Santilli RA, Grego E, Battaia S, Gianella P, Tursi M, Di Girolamo N, Biasato I, Perego M. Prevalence of selected cardiotropic pathogens in the myocardium of adult dogs with

unexplained myocardial and rhythm disorders or with congenital heart disease. J Am Vet Med Assoc. 2019 Nov 15;255(10):1150-1160. doi: 10.2460/javma.255.10.1150. PMID: 31687895.

Tschöpe C, Ammirati E, Bozkurt B, Caforio ALP, Cooper LT, Felix SB, Hare JM, Heidecker B, Heymans S, Hübner N, Kelle S, Klingel K, Maatz H, Parwani AS, Spillmann F, Starling RC, Tsutsui H, Seferovic P, Van Linthout S. Myocarditis and inflammatory cardiomyopathy: current evidence and future directions. Nat Rev Cardiol. 2021 Mar;18(3):169-193. doi: 10.1038/s41569-020-00435-x.

Chapter 1

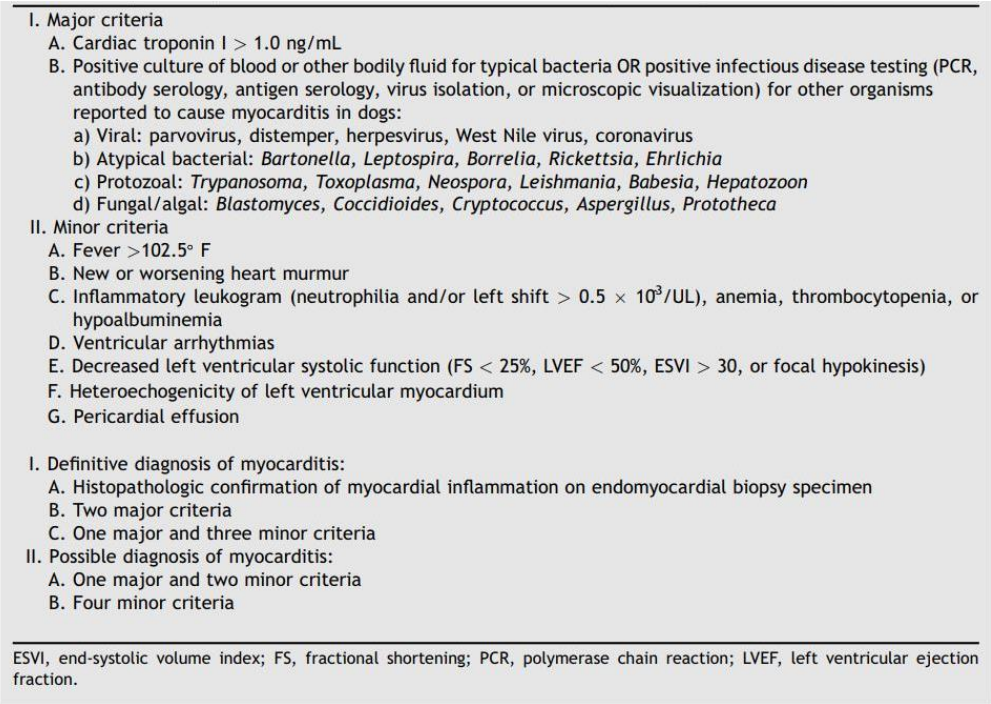
MYOCARDITIS IN DOG

1.1 MYOCARDITIS IN DOG

Canine myocarditis is a non-specific condition related to the inflammation of the myocardium. Although it is associated with devastating effects on dogs' health, myocarditis incidence remains often unknown. The literature related to myocarditis presentation, treatment and diagnosis in dogs, consisted of a few case reports and small case series. One of the largest studies is provided by Lakhdhir and collaborators (2020), which reported clinical features, clinicopathologic and cardiovascular findings, treatment, and outcome in dogs with a presumptive antemortem diagnosis of myocarditis and histopathologic findings in dogs with postmortem confirmation of myocarditis. Infectious and non-infectious agents have been described in association with the development of canine myocarditis. Noninfective myocarditis is associated with toxins or pharmacological drugs, heatstroke, trauma, hemodynamic shock, immune-mediated diseases, and idiopathic forms of myocarditis [Janus et al., 2014]. Lakhdhir and collaborators reported non-infectious myocarditis in 17% of dogs examined, including rodenticide toxicity, immune-mediated hemolytic anemia, and neoplasia. In this study, non-specific clinical abnormalities were recorded, lethargy and hyporexia were the most common symptoms reported less frequently fever, heart murmur, and pulse abnormalities. Other clinical findings reported in this work regarded common blood count (CBC) and biochemistry abnormalities, that were consistent with non-specific inflammation (neutrophilia, left shift, monocytosis, and anemia) or decreased organ perfusion (elevated ALT and azotemia). Arrhythmias were commonly distributed among analyzed specimens, with more than half of dogs, exhibiting ventricular premature complexes and one-third of dogs experiencing ventricular tachycardia. Decreased left ventricular systolic function was the most common echocardiographic abnormality, described in more than 50% of cases. Generally, high levels of cardiac troponin I (cTnI) (>0.2 ng/ml), have been described in dogs affected with the myocardial disease and myocardial neoplasia [Chun et al., 2010, Wesselowski et al., 2019], as well as dogs with various congenital and acquired cardiac diseases [Wess et al., 2010]. Concordantly, Lakhdhir and collaborators showed how cardiac troponin I concentrations >1.0 ng/ml may support a diagnosis of myocarditis in combination with other proposed diagnostic criteria. For humans, a set of defined criteria for diagnosis of myocarditis by non-invasive testing was proposed and has been shown to support clinicians when myocarditis is suspected [Caforio et al., 2013]. Similarly, for endocarditis, the modified Duke criteria, are an important diagnostic tool and were suited for dogs and cats [Macdonald 2010, Palerme et al., 2016]. These modified set of criteria included some major and minor criteria, as described in figure 1 and in

order to diagnose definitive or possible myocarditis, combined major or minor criteria are required.

Figure 1



I. Major criteria
- A. Cardiac troponin I > 1.0 ng/mL
- B. Positive culture of blood or other bodily fluid for typical bacteria OR positive infectious disease testing (PCR, antibody serology, antigen serology, virus isolation, or microscopic visualization) for other organisms reported to cause myocarditis in dogs:
  - a) Viral: parvovirus, distemper, herpesvirus, West Nile virus, coronavirus
  - b) Atypical bacterial: *Bartonella, Leptospira, Borrelia, Rickettsia, Ehrlichia*
  - c) Protozoal: *Trypanosoma, Toxoplasma, Neospora, Leishmania, Babesia, Hepatozoon*
  - d) Fungal/algal: *Blastomyces, Coccidioides, Cryptococcus, Aspergillus, Prototheca*

II. Minor criteria
- A. Fever >102.5° F
- B. New or worsening heart murmur
- C. Inflammatory leukogram (neutrophilia and/or left shift > 0.5 × 10^3/UL), anemia, thrombocytopenia, or hypoalbuminemia
- D. Ventricular arrhythmias
- E. Decreased left ventricular systolic function (FS < 25%, LVEF < 50%, ESVI > 30, or focal hypokinesis)
- F. Heteroechogenicity of left ventricular myocardium
- G. Pericardial effusion

I. Definitive diagnosis of myocarditis:
- A. Histopathologic confirmation of myocardial inflammation on endomyocardial biopsy specimen
- B. Two major criteria
- C. One major and three minor criteria

II. Possible diagnosis of myocarditis:
- A. One major and two minor criteria
- B. Four minor criteria

ESVI, end-systolic volume index; FS, fractional shortening; PCR, polymerase chain reaction; LVEF, left ventricular ejection fraction.

Figure 1: summary of major and minor criteria for the diagnosis of myocarditis in dogs proposed by Lakhdhir et al., 2020

Although the definitive diagnosis of myocarditis is obtained with histopathology which demonstrates inflammation of the myocardium, for antemortem diagnosis, endomyocardial biopsy (EMB) is required. EMBs are commonly performed in the clinical practice for humans [Biesbroek et al., 2015], but not for the dogs, in which are not often executed, probably due to the particular devices used and high-level expertise required, also combined with the risk associated with these procedures. Nonetheless, in the last years, some veterinary clinic centers included EMB during diagnostic cardiological procedures, the antemortem diagnosis of myocardial disorders is generally presumptive and based on an individual clinician's interpretation of case presentation and clinicopathological information [Santilli et al., 2017, 2019].

1.2 ETIOLOGY OF INFECTIOUS MYOCARDITIS IN DOG

Concerning infectious myocarditis Lakhdhir and collaborators (2020) show that the most common cause of myocarditis was bacterial sepsis or extension of infective endocarditis (12/64 dogs in the antemortem population), although other etiologic agents, including canine parvovirus, Toxoplasma, Neospora, Leptospira, Trypanosoma, Bartonella, and Dirofilaria, were also detected. Myocarditis occurs most commonly secondary to systemic bacterial infection and septicemia, however in some geographic areas endemic to vector-borne infections myocarditis can be related to trypanosomiasis or leishmaniasis infection. In particular, *Leishmania spp.* is a flagellate protozoan parasite responsible for the zoonotic disease of leishmaniasis. The disease is distributed worldwide and is considered endemic in more than 70 countries, principally in Latin America, Asia, Africa, and Mediterranean regions. This protozoon infects both humans and dogs and is transmitted by the bite of the insect Phlebotominae, which is the intermediate host; while canines are the natural hosts and reservoirs of the disease. Histological examination of dogs affected with leishmaniasis, shows mononuclear inflammatory infiltrate, mainly lymphohistioplasmacytic, and myocardial inflammatory infiltrates in both the epicardium and the endocardium with focal, focally extensive and multifocal lesions. In addition, focal myocardial fibrosis and mild necrotizing subepicardial vasculitis have been reported [Costagliola et al., 2016]. *Toxoplasma gondii* and *Neospora caninum* are other protozoa parasites that were found in association with myocarditis, mainly in immunocompromised dogs [Odin et al., 1993, Meseck, et al., 2005]. In dogs, also viral-related myocarditis has been described, but, while in humans, viruses are known as the most common etiology, (about 69% of myocarditis cases), canine myocarditis related to viral etiology is undetermined [Fairley et al., 1996]. Examples of viral etiologies are canine parvovirus, as well as West Nile virus, even if they are not commonly distributed among canine myocarditis [Cannon et al., 2006]. Since the late 1970s, the canine parvovirus (CPV) has been identified in association with myocardial damage in dogs, resulting in the rapid progression of congestive heart failure, arrhythmias, and sudden death. Moreover, CPV is also described in other significant worldwide canine diseases in two further different forms: the enteric and the neurologic form. CPV is originated from a mutation of the feline panleukopenia virus (FPV), is classified into several serotypes, CPV-1 (minute virus), CPV-2, CPV 2a, and CPV 2b.

The association between CPV and myocarditis is also described in puppies that are infected in utero or within the first weeks of life. In their study Sime and coll. (2015) found that the viral infection may cause myocardial damage, without gastrointestinal signs, and with

respiratory sympthoms that occured in the acute phase. Animals that passed this phase could develop chronic cardiac changes and juvenile DCM. Finally, they proposed that when young puppies and juvenile dogs presented signs of acute respiratory distress and evidence of heart failure, myocardial parvovirus shoud be considered as a biomarker for a more precise diagnosis [Sime et al., 2015]. However, in 2017, Ford and collegues, described a significant relationship between canine parvovirus 2 and myocarditis/myocardial fibrosis in young dogs [Ford et al., 2017].

Bacteria, such as Borrelia, Bartonella, Leptospira, Citrobacter, Staphylococcus, and Streptococcus, are rarely described as etiology of myocarditis in the veterinary literature, nonetheless, bacteria are the most common confirmed etiology identified in the Lakhdhir study [Lakhdhir et al., 2020]. These findings suggested that bacterial septicemia may represent a risk of myocardial involvement, which resulted underestimated, and bacteria could be considered as possible trigger agents of myocarditis in some geographic areas.

Treatment of myocarditis can be challenging because of difficulty obtaining a definitive antemortem diagnosis. The treatment plan was not standardized and generally reflected clinician preference in the context of suspected etiology. The prognosis for dogs diagnosed with myocarditis is poor to grave, while for dogs who survived at least 2 weeks after diagnosis is more encouraging, with the median survival time between 82 days, and over 2 years.

1.3 STRATEGIES FOR DIAGNOSIS OF INFECTIOUS MYOCARDITIS

Current diagnostic tests during infectious myocarditis require different approaches including immunohistochemistry, PCR, in Situ Hybridization. However, despite a comprehensive diagnostic screening, the cause of myocardial lesions remained frequently undetermined, demonstrating the limitations of available molecular tests. Furthermore, pathogen detection by PCR did not imply causation; Molesan and colleagues (2019), in their work, described a retrospective study obtained with the analysis of nucleic acid extracted from archived (from 2007 to 2015) tissues of myocarditis cases and control dogs without myocardial lesions. Their study shows how the presence of Mycoplasma haemocanis, canine adenovirus (CAV), canine respiratory coronavirus (CRCoV), canine parainfluenza virus (CPiV), canine distemper virus (CDV), and canine herpesvirus (CHV) were not associated with myocardial inflammation, as pathogens' detection with equivalent frequency, characterized both myocarditis cases and controls. The authors reported PCR-detected pathogens (CDV, CHV, and CAV) are probably

associated with subclinical or latent infection, particularly in adult animals. As immunization with modified live vaccine (CDV, CPV-2, CAV-2, and CPiV) is recommended for dogs beginning as early as 6 weeks of age, PCR may be detecting vaccine virus and not field strains. The co-detection of multiple agents infrequently associated with disease (due to widespread vaccination) suggested the detection of vaccine virus [Molesan et al., 2019]. PCR detection at least represents evidence of exposure to an infectious agent; thus, ultimately, it is unclear if PCR-detected agents are cardio-pathogenic, incidental, or latent as clinical recognition of myocarditis usually occurs as viral titers are waning. Considering the limits of PCR, new diagnostic approaches are needed to identify infectious agents of myocarditis able to reveal emerging, unrecognized, or novel pathogens. In this context, NGS may results in a useful approach with several advantages, including high sensitivity detection levels in combination with the comprehensive achievement of information.

Biesbroek PS, Beek AM, Germans T, Niessen HW, van Rossum AC. Diagnosis of myocarditis: Current state and future perspectives. Int J Cardiol. 2015 Jul 15;191:211-9. doi: 10.1016/j.ijcard.2015.05.008. Epub 2015 May 6. PMID: 25974197.

Caforio AL, Pankuweit S, Arbustini E, Basso C, Gimeno-Blanes J, Felix SB, Fu M, Heliö T, Heymans S, Jahns R, Klingel K, Linhart A, Maisch B, McKenna W, Mogensen J, Pinto YM, Ristic A, Schultheiss HP, Seggewiss H, Tavazzi L, Thiene G, Yilmaz A, Charron P, Elliott PM; European Society of Cardiology Working Group on Myocardial and Pericardial Diseases. Current state of knowledge on aetiology, diagnosis, management, and therapy of myocarditis: a position statement of the European Society of Cardiology Working Group on Myocardial and Pericardial Diseases. Eur Heart J. 2013 Sep;34(33):2636-48, 2648a-2648d. doi: 10.1093/eurheartj/eht210. Epub 2013 Jul 3. PMID: 23824828.

Cannon AB, Luff JA, Brault AC, MacLachlan NJ, Case JB, Green EN, Sykes JE. Acute encephalitis, polyarthritis, and myocarditis associated with West Nile virus infection in a dog. J Vet Intern Med. 2006 Sep-Oct;20(5):1219-23. doi: 10.1892/0891-6640(2006)20[1219:aepama]2.0.co;2. PMID: 17063720.

Chun R, Kellihan HB, Henik RA, Stepien RL. Comparison of plasma cardiac troponin I concentrations among dogs with cardiac hemangiosarcoma, noncardiac hemangiosarcoma, other neoplasms, and pericardial effusion of nonhemangiosarcoma origin. J Am Vet Med Assoc. 2010 Oct 1;237(7):806-11. doi: 10.2460/javma.237.7.806. PMID: 20919846.

Costagliola A, Piegari G, Otrocka-Domagala I, Ciccarelli D, Iovane V, Oliva G, Russo V, Rinaldi L, Papparella S, Paciello O. Immunopathological Features of Canine Myocarditis Associated with Leishmania infantum Infection. Biomed Res Int. 2016;2016:8016186. doi: 10.1155/2016/8016186. Epub 2016 Jun 19. PMID: 27413751; PMCID: PMC4930798.

Fairley CK, Ryan M, Wall PG, Weinberg J. The organisms reported to cause infective myocarditis and pericarditis in England and Wales. J Infect. 1996 May;32(3):223-5. doi: 10.1016/s0163-4453(96)80023-5. PMID: 8793712.

Ford J, McEndaffer L, Renshaw R, Molesan A, Kelly K. Parvovirus Infection Is Associated With Myocarditis and Myocardial Fibrosis in Young Dogs. Vet Pathol. 2017 Nov;54(6):964-971. doi: 10.1177/0300985817725387. Epub 2017 Aug 16. PMID: 28812526.

Janus I, Noszczyk-Nowak A, Nowak M, Cepiel A, Ciaputa R, Pasławska U, Dzięgiel P, Jabłońska K. Myocarditis in dogs: etiology, clinical and histopathological features (11 cases:

2007-2013). Ir Vet J. 2014 Dec 24;67(1):28. doi: 10.1186/s13620-014-0028-8. PMID: 25642323; PMCID: PMC4311452.

Lakhdhir S, Viall A, Alloway E, Keene B, Baumgartner K, Ward J. Clinical presentation, cardiovascular findings, etiology, and outcome of myocarditis in dogs: 64 cases with presumptive antemortem diagnosis (26 confirmed postmortem) and 137 cases with postmortem diagnosis only (2004-2017). J Vet Cardiol. 2020 Aug;30:44-56. doi: 10.1016/j.jvc.2020.05.003. Epub 2020 May 29. PMID: 32668360; PMCID: PMC7256493.

Macdonald K. Infective endocarditis in dogs: diagnosis and therapy. Vet Clin North Am Small Anim Pract. 2010 Jul;40(4):665-84. doi: 10.1016/j.cvsm.2010.03.010. PMID: 20610018.

Meseck EK, Njaa BL, Haley NJ, Park EH, Barr SC. Use of a multiplex polymerase chain reaction to rapidly differentiate Neospora caninum from Toxoplasma gondii in an adult dog with necrotizing myocarditis and myocardial infarct. J Vet Diagn Invest. 2005 Nov;17(6):565-8. doi: 10.1177/104063870501700607. PMID: 16475515.

Molesan A, Goodman L, Ford J, Lovering SJ, Kelly K. The Causes of Canine Myocarditis and Myocardial Fibrosis Are Elusive by Targeted Molecular Testing: Retrospective Analysis and Literature Review. Vet Pathol. 2019 Sep;56(5):761-777. doi: 10.1177/0300985819839241.

Odin M, Dubey JP. Sudden death associated with Neospora caninum myocarditis in a dog. J Am Vet Med Assoc. 1993 Sep 15;203(6):831-3. PMID: 8226237.

Palerme JS, Jones AE, Ward JL, Balakrishnan N, Linder KE, Breitschwerdt EB, Keene BW. Infective endocarditis in 13 cats. J Vet Cardiol. 2016 Sep;18(3):213-225. doi: 10.1016/j.jvc.2016.04.003. Epub 2016 Jun 6. PMID: 27283084.

Santilli RA, Battaia S, Perego M, Tursi M, Grego E, Marzufero C, Gianella P. Bartonella-associated inflammatory cardiomyopathy in a dog. J Vet Cardiol 2017;19:74-81.

Santilli RA, Grego E, Battaia S, Gianella P, Tursi M, Di Girolamo N, Biasato I, Perego M. Prevalence of selected cardiotropic pathogens in the myocardium of adult dogs with unexplained myocardial and rhythm disorders or with congenital heart disease. J Am Vet Med Assoc. 2019 Nov 15;255(10):1150-1160. doi: 10.2460/javma.255.10.1150. PMID: 31687895.

Sime TA, Powell LL, Schildt JC, Olson EJ. Parvoviral myocarditis in a 5-week-old Dachshund. J Vet Emerg Crit Care (San Antonio). 2015 Nov-Dec;25(6):765-9. doi: 10.1111/vec.12347. Epub 2015 Jul 28. PMID: 26220397.

Wess G, Simak J, Mahling M, Hartmann K. Cardiac troponin I in Doberman Pinschers with cardiomyopathy. J Vet Intern Med. 2010 Jul-Aug;24(4):843-9. doi: 10.1111/j.1939-1676.2010.0516.x. Epub 2010 Apr 16. PMID: 20412436.

Wesselowski S, Cusack K, Gordon SG, Jeffery N, Saunders AB. Artificial cardiac pacemaker placement in dogs with a cohort of myocarditis suspects and association of ultrasensitive cardiac troponin I with survival. J Vet Cardiol. 2019 Apr;22:84-95. doi: 10.1016/j.jvc.2018.09.003. Epub 2018 Dec 18. PMID: 30573438.

Chapter 2

MYOCARDITIS IN HUMAN

2.1 MYOCARDITIS IN HUMAN

Myocarditis concerns multiple inflammatory conditions related to the heart muscle with a broad spectrum of clinical and histopathological signs [Cooper 2009]. In the past, the word myocarditis referred to infarction of myocardium and chronic ischemic heart disease, but recently the term has been refined. Nowadays the definition of myocarditis involves acute or chronic inflammatory responses of the heart muscle to environmental or endogenous triggers. In this context, myocarditis may be associated with infectious and non-infectious causes. Among non-infectious causes, drug-induced hypersensitivity, cardiac manifestations of systemic autoimmunities, such as sarcoidosis or systemic lupus erythematosus, and giant cell myocarditis (GCM) are well-documented specific forms of human myocarditis, that occur rarely. As described by the International Classification of Diseases, 9th Revision diagnoses, in humans, the incidence of myocarditis was 22 per 100,000 people (about 1.5 million cases in 2013) [Global Burden of Disease Study 2013 Collaborators]. The prevalence of myocarditis in 2014 varied from 0.5% to 4.0%, depending mainly on the age and the geographic area of patients [Cooper et al., 2014]. The global number of deaths caused by myocarditis and cardiomyopathy in 2015 was assessed to be around 354,000, with a death rate of 4.8 per 100,000 people. Furthermore, myocarditis was responsible for sudden cardiovascular death in approximately 2% of infants, 5% of childhood, and 5% to 12% of young athletes' sudden death [Maron et al., 2015].

2.2 ETIOLOGY OF HUMAN MYOCARDITIS

The most common causes of myocarditis in humans include viruses and less frequently bacteria, fungi and parasites [Schultheiss et al., 2011]. In the past, from the 1950s to the 1990s, frequently detected infective agents in North America and Western Europe were enteroviruses, particularly the Coxsackie virus. In the last 2 decades, with the improvement of detection techniques, such as polymerase chain reaction and *in situ* hybridization, several cardiotropic viruses were identified in heart biopsy samples, including Adenovirus [Pauschinger et al., 1999, Bowles et al., 2003], Cytomegalovirus, Echovirus, Parvovirus B-19, and Epstein-Barr virus [Chimenti et al., 2004], Hepatitis C virus [Matsumori et al., 2000, Omura et al., 2005], Influenza A virus (strain H1N1) [Baruteau et al., 2010]. In addition, patients with chronic symptoms and "inflammatory cardiomyopathy", are characterized by the predominant presence of Human herpesvirus 6 [Kühl et al., 2005] and Parvovirus B19 [Breinholt et al.,

2010] genomes. Nonetheless, non-viral infection and autoimmune syndromes are reported as important causes of myocarditis. Bacteria, such as *Corynebacterium diphtheria*, *Beta-haemolytic streptococci*, *Meningococci*, *Rickettsie*, *Salmonella typhior paratyphi*, *Borrelia burgdorferi* (Lyme disease), *Mycoplasma pneumonia* and *Chlamydia psittaci* may lead to myocarditis. *Corynebacterium diphtheria* infection, for example, is a major public health problem in many underdeveloped countries and probably could be considered as the most common cause of myocarditis worldwide, while it is less frequent in western regions [Golpour et al., 2021]. Also parasites are related to myocarditis; in South America, *Trypanosoma cruzi* is reported as the main etiologic agent (Chagas disease) [Marin-Neto et al., 2007]. Infections, injury and autoantigen-specific mechanisms activate the host immune response, which may lead to organ dysfunction, pathological remodeling and heart failure. Among non-infectious causes, hypersensitivity myocarditis may occur from drug- or vaccine-related heart-specific autoimmunity. Several drugs such as tricyclic antidepressants, benzodiazepines, furosemide, methyldopa, azithromycin, hydrochlorothiazide, ampicillin, tetracycline, aminophylline, phenytoin, and tumor necrosis factor antagonists are associated with myocarditis [Leone et al., 2019]. In most cases of hypersensitivity myocarditis, inflammation evolves early during drug treatment, nonetheless up to 15% of myocarditis induced by clozapine, develops later, even 2 years after the first dose of therapy [Cook et al., 2015]. Vaccine-related myocarditis occurred in the case of smallpox vaccination in up to 6 per 10,000 vaccines [Engler et al., 2015]. More recently, studies focused attention on the role of COVID-19 vaccination and myocarditis. Perez and colleagues reported cases of myocarditis as a rare adverse effect associated with COVID-19 mRNA vaccines both in young and adult patients. As expected, cardiovascular disease or history of myocarditis should be considered as a risk factor for COVID-19 vaccine-related myocarditis. Furthermore, they found that inflammation is more often associated with the second dose of the COVID-19 mRNA vaccine. Nonetheless, literature is still lacking and more data are needed [Perez et al., 2021]. Cardiac inflammation in the context of radiation or systemic autoimmune diseases (i.e. systemic sclerosis, lupus erythematosus, celiac disease or antinuclear antibody-related vasculitis) is also reported as non-infectious causes of myocarditis [Zawadowski et al., 2012]. GCM, previously mentioned, is a rare form of myocardial disease associated with rapid progression and frequently fatal outcomes in young and middle-aged adults. GCM is associated with T lymphocyte-mediated inflammation of the heart muscle and characterized about 20% of cases with systemic autoimmune diseases [Cooper et al., 1997]. Currently, most of the knowledge about the mechanism underlying human myocarditis derives from studies based on susceptible rodent models. In these experiments, the inflammation

evolves after cardiomyocyte infection with specific virus strain, bacteria or parasites [Cihákova et al., 2004]. Other strategies to develop inflammatory cellular infiltrates, involve the use of cardiac antigens delivered with strong adjuvant or carried within dendritic cells in noninfectious models. Models not based on infection, are not exhaustive and only provide information on the final phases of myocarditis disease [Watanabe et al., 2015].

2.3 PATHOGENESIS OF HUMAN MYOCARDITIS

The changes that lead from acute viral infection to a DCM, through active inflammation, may be considered as a multiphase model summarized in figure 2. The selection of proper therapy, as well as the underlying mechanisms of pathogenesis, are dependent on the stage of the disease and the infective agents. For viral myocarditis, the interaction between virus and cardiomyocytes triggers the type 1 interferons production in the first hours with myocytes death based on apoptosis and autophagy processes. In the case of coxsackie B virus strains infection, for example, specific cell surface receptors, such as the coxsackie-adenovirus receptor and CD55 receptor are required. The host inflammatory signaling pathways play an important role in viral replication and contribute to coxsackie B virulence [Abston et al., 2012]. On the other hand, Chagas myocarditis, related to *Trypanosoma cruzi* infection, is characterized by tissue damage caused by not only host inflammatory cell invasion, but also parasite bioactive metabolites, released during host-pathogen interaction [Bonney et al., 2015]. After infection, a second phase occurs over hours to days and is characterized by the activation of the host innate immune response that includes modified regulatory T-cell function, nitric oxide, natural killer (NK) cells. In this context, the host releasing of mediators depends on infective agents: type I interferon for viral pathogens and granule components from eosinophils and polymorphonuclear cells for bacterial or parasites infections. A third phase develops with the migration of antigen-bearing cells to regional lymph nodes in a proper cytokine environment. Furthermore, cellular infiltrates in cardiac tissue with autoantigen or pathogen-specific T cells, antibodies or macrophages are specific for this phase [Noutsias et al., 2011, Jenke et al., 2014]. Finally, with the clearance of the infective agent, usually cardiac muscle recovers its normal functions. In humans and experimental models, a process of chronic autoantigen-driven inflammation may occur consequently to the loss of T-cell tolerance to cardiac self-antigens. This chronic scenario may progress to DCM and end-stage heart failure [Caforio et al., 2013]. The role of cardiac autoimmunity in disease evolution is dependent on multiple factors: the

genetic sensibility of the host, the modifications in the pathogens' genomes, the molecular similarity between microbial proteins and host cardiac structures as well as environmental elements. In addition, heat shock proteins or immunogenic fragments released after injury, non-infectious microbial elements or drug treatments may be also considered as trigger agents for autoimmune inflammation development. The switch to chronic fibrosis, from acute inflammation, is a critical step that may occur with persistent inflammation in absence of visible symptoms [Blyszczuk et al., 2009].

Figure 2

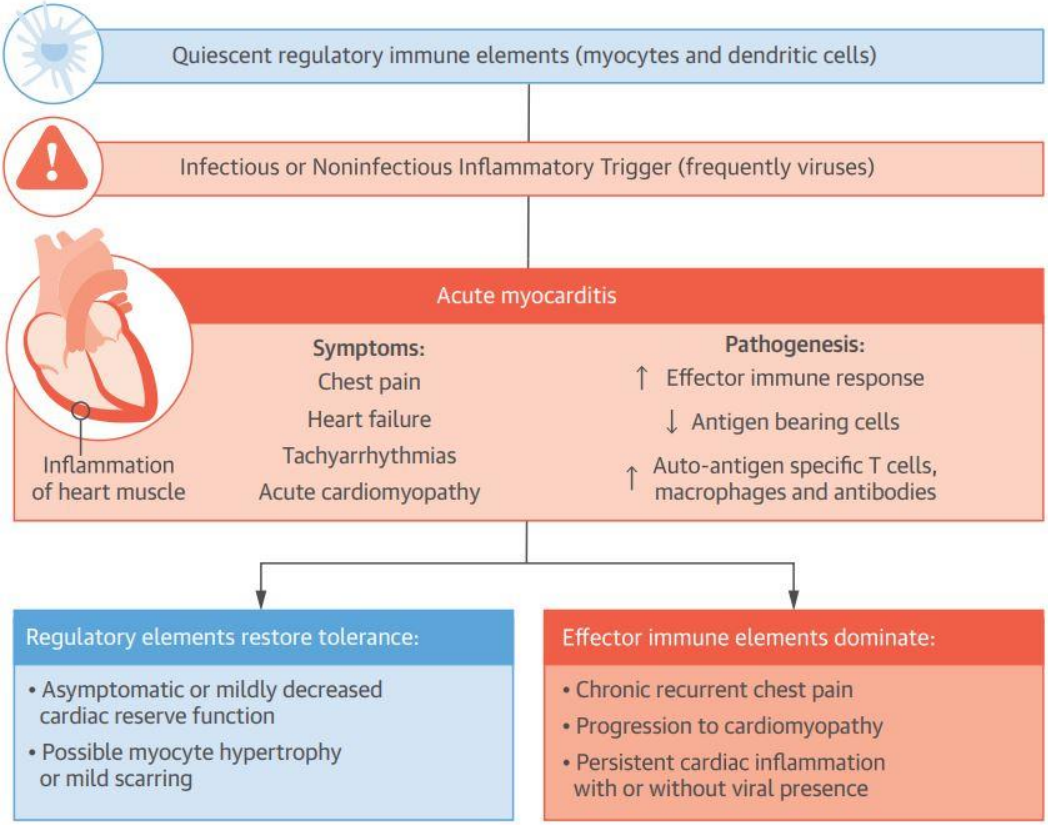


Figure 2: Conceptualization of multiphase model of myocarditis. Acute myocarditis may derive from several injuries. that often develop immune responses to both pathogens and myocardial tissue. Clinical recovery or progression to chronic cardiomyopathy or recurrent chest pain depends on the balance of regulatory and effector elements. Chronic injury may be associated with ongoing viral infection [Heymans et al., 2016].

2.4 TISSUE INVOLVEMENT DURING CARDIAC INFLAMMATION

The myocardial inflammation and remodeling involve several cellular compartments:

- The endothelial compartment, which represents a barrier that modulates the access of circulating bone marrow-derived cells to the heart.
- The bone-marrow compartment which includes T-cells and B-cells, as an effector of adaptive immunity, and myeloid-derived population represented by macrophages, granulocytes, eosinophils, dendritic cells and mast cells. In addition, several precursors are involved, with the capacity to differentiate into myofibroblasts and/or other inflammatory cells.
- The interstitial compartment represented by myofibroblasts, fibroblasts and other stromal cells that are adsorbed in an interstitial matrix and are involved in the process of shaping of inflammatory phenotype.
- The cardiomyocyte compartment, which provides a specific adaptive response to inflammatory triggers represented by altered calcium signaling and hypertrophy [Frieler et al., 2015].

The pathological anatomic phenotype of the inflamed heart is defined by the modifications in these cellular compartments following a trigger event (ischemia, infection, toxins or immunological balance). The regulation of the changes in cellular compartments, including activation, deactivation, or transdifferentiation of specific cell subtype, is provided by cell-cell interactions and autocrine and paracrine mediators (cytokines, hormones, chemokines, degradation or danger products). These molecules are involved in the regulation of gene expression through relative specific receptors and their intracellular downstream pathways [Gratchev et al., 2006]. In this scenario, gene silencing is an additional mechanism involved in the post-transcriptional regulation of myocarditis. In conclusion, the extracellular matrix, that fills the extracellular space, represents another element able to modulate cardiac inflammation, remodeling and tissue fibrosis through structural support and tissue organization [Rienks et al., 2016].

2.5 INNATE IMMUNE RESPONSES IN THE INFLAMMATORY PROCESS

The innate immune system is represented by several cells including granulocytes, eosinophils, histiocytes, mast cells, dendritic cells and macrophages. These cells are able to recognize, through their intracellular and surface receptors, conserved microbial structures, that are not present in host molecular patterns [Epelman et al., 2015]. The microbial molecules in myeloid cells induce the formation of a multiprotein oligomer known as inflammasome, composed of PYCARD, caspase 1 and NOD-like receptors (Nucleotide-binding Oligomerization Domain). The inflammasome triggers the caspase 1 cascade, with the consequent production of proinflammatory cytokines. On the other hand, Toll-like receptors (TLRs) is a family of pattern recognition receptor molecules that are present on the cell surface, in human, eleven different TLRs have been characterized. They play an important role in the early innate immune response in case of pathogen infections and are particularly present on macrophages and dendritic cells surface. Among them, TLR2, TLR3, TLR4, TLR7, and TLR9 and their downstream adaptors (MyD88 and TRIF) have a specific function in inflammatory heart disease and myocarditis [Pagni et al., 2010, Riad et al., 2012, Jenke et al., 2013]. Interestingly, in experimental mice with no expression of TLR3, has been described a high rate of mortality consequently to enteroviruses infection. In addition, a common gene polymorphism of TLR3 characterizes patients with enteroviral myocarditis or cardiomyopathy [Gorbea et al., 2010, Abston et al., 2013]. Polymorphic variants of the TLR3 gene reduce TLR3-mediated signaling after recognition of enterovirus-specific pathogen-related molecular patterns and increase viral replication. Furthermore, in non-ischemic cardiomyopathy, the expression of PAR2 (Protease Activated Receptor 2), a TLR3 interactor, is strongly associated with high inflammation rates and negatively correlates with IFN-β expression in cardiac tissue [Weithauser et al., 2013]. In myocarditis and DCM, the releasing of cytokines, such as tumor necrosis factor (TNF)-α, interleukin (IL)-1β, IL-18, or high-mobility group box 1 protein, is induced by the activation TRLs combined with the inflammasome. In addition, TLR2 and TLR4 regulate TNF releasing and IL-1β secretion and may be directly activated by self-proteins, like cardiac myosin [Zhang et al., 2009, Su et al., 2016]. In coxsackievirus B3 infection, for example, it has been described upregulation of TRL4 in macrophages and mast cells during the early innate immune response. Interestingly, patients with myocarditis, are characterized by upregulation of TRL4 mRNA (messenger ribonucleic acid) expression, which is also directly associated with high levels of viral RNAs. In experimental TRL4-deficient mice, cardiac inflammation is reduced with consequently decreased levels of IL-1β, which is directly involved in cardiac fibrosis,

pathological remodeling and heart failure progression [Blyszczuk et al., 2009]. Conversely, heart tissue inflammation is supported by high-mobility group box 1 protein as described in mouse models of experimental autoimmune myocarditis [Su et al., 2016].

2.6 LEUKOCYTES SUBSETS INVOLVED IN THE INFLAMMATION PROCESS

In human and experimental myocarditis, most infiltrating cells are represented by cells of macrophages and monocyte lineages. Macrophages are mononuclear phagocytic cells with regulatory and microbicidal roles in heart tissue inflammation, in particular, $Ly6C^{hi}$ inflammatory macrophages are recruited in cardiac injury upon myocarditis. Blocking chemokines involved in $Ly6C^{hi}$ recruitment (CCR2 ligands CCL2/MCP1 or CCL3/MIP1a) may improve outcomes in autoimmune myocarditis [Barin et al., 2012, Leuschner et al., 2015]. Furthermore, the differentiation of monocytes in proinflammatory or classical M1 macrophages is strongly affected by $CD4^+$ T (type 1 T helper – Th1) cells, the antimicrobial activity and antigen presentation of macrophages, indeed, are promoted by Th1 IFN-γ in cardiac inflammation. Alternatively-activated macrophages M2 type ($Ly6C^{low}$ M2 macrophages), are the most abundant white cells during the myocardial recovery process and promote cardiac fibrotic healing with reduction of inflammatory response. In this context, IL-4 and IL-13, two types of Th2-related cytokines, support the activation of anti-inflammatory M2 macrophages, that characterize the transition from acute myocarditis to chronic pathological remodeling [Martinez et al., 2008].

Despite they represent the minority of cardiac infiltrating cells, both in the mouse model and in human biopsies, T cells are the relevant actors in the development of myocarditis [Afanasyeva et al., 2004]. Among them, in order to overcome early acute viral infections, the responses of γδ T cells, NK, $CD4^+$ and $CD8^+$ are fundamental. T cells responses, in particular $CD4^+$, even after removing pathogens, may switch initial myocarditis to an autoimmune reaction, in part caused by intolerance to alpha-myosin heavy chain (α-MyHC) in thyme. The expansion of autoreactive $CD4^+$ T cells and heart inflammation is promoted by activated, self-antigen–expressing dendritic cells [Lv et al., 2011]. In rodent models, with overexpression of α-MyHC-specific T cell receptors (TCR) in $CD4^+$ T cells, in fact, autoimmune myocarditis develops. Based on the cytokines they produce, $CD4^+$ effector T cells may be categorized into 3 groups:

- Th1 $CD4^+$ cells characterized by the production of IFN-γ, they predominate in autoimmune myocarditis in mice and human patients;
- Th2 $CD4^+$ T cells that generate IL-4, IL-5, and IL-13;
- Th17 $CD4^+$ T cells, mainly produce IL-17.

Among those subsets, Th17 cells and related cytokines, play a critical role in chronic myocarditis and cardiac myosin-Th17 response is associated with heart failure in humans. Th1 cells, instead, act as regulators of cardiac inflammation and as mediators in early autoimmune T effector cell expansion [Myers et al., 2016]. Furthermore, both human and mice model myocarditis and DCM, are characterized by the presence of heart-specific, T-helper, cell-dependent immunoglobulin G autoantibodies, in particular, they may predict the outcome of DCM patients.

2.7 EPIGENETIC REGULATION OF INFLAMMATORY PROCESS

In the context of post-translational regulation of inflammation of myocarditis, the expression of polygenetic susceptibility is influenced by epigenetic factors, such as micro-ribonucleic acids (miRNAs). They are small non-coding molecules of RNA (~21 nucleotide length), that have emerged as epigenetic regulators of gene expression. By imperfectly binding to specific target regions, a single miRNA may control the expression of transcripts for several proteins, clustered into a biological mechanism or pathway, such as inflammation or fibrosis [Small et al., 2011]. As miRNA action may be blocked by the injection of complementary filament (anti-miRs), they could be considered a new therapeutic strategy. Recent studies, in fact, support the relationships between myocarditis and the uncontrolled expression of miRNA. Novel miRNA therapeutic targets are suggested, such as miRNA-155, miRNA-146b, and miRNA-21, which are up-regulated in cardiac tissues during the acute phase of human myocarditis and murine coxsackie B3 myocarditis [Corsten et al., 2012]. Inhibition of these molecules, with delivered anti-miRs, for example, has been reported to reduce heart inflammation and cardiac tissue damage, in mice models with autoimmune myocarditis [Yan et al., 2016, Zhang et al., 2016]. Moreover, the virulence of cardiotropic viruses may be controlled by miRNAs, such as miRNA-221 or miRNA-222. In case of infection with the enterovirus CVB3 via coxsackie and adenovirus receptor in cardiomyocytes of mice models, systemic inhibition of miRNA-221 and miRNA-222 results in an increasing of viral load with consequent worsening of cardiac

inflammation and heart damage. These 2 miRNAs target several proteins, like interferon regulatory factor 2, ETS1/2, TOX, Bcl2-like 11, Bcl2 modifying factor, and CXCL12, that modulate viral replication and tissues inflammation [Corsten et al., 2015]. miRNAs are also investigated as novel instruments for diagnosis and prognosis: a recent study proposes 8 miRNAs (miR- 135b, -155, -190, -422a, -489, -590, -601, and -1290) as biomarkers of acute inflammation, as they are strongly related to viral myocarditis characterized by cardiac dysfunction and viral persistence in heart tissues [Kuehl et al., 2015].

2.8 SYMPTOMS AND DIAGNOSTIC RELATED TO MYOCARDITIS

Most of the common symptoms of acute myocarditis are dyspnea or chest pain. The clinical syndrome is usually preceded by a viral syndrome that occurs several days or weeks before, such as upper respiratory tract or gastrointestinal diseases. When there is a suspect of myocarditis, endomyocardial biopsies should be recommended for patients that present the following characteristics: demand of inotropic or mechanical circulatory support, development of Mobitz type2 second-degree or heart block, symptomatic tachycardia, lack of responsiveness to guideline-based medical management within 1 to 2 weeks. Furthermore, also for suspected cases of eosinophilic myocarditis and myocarditis associated with systemic inflammatory disorders, an endomyocardial biopsy procedure could be also considered. Other noninvasive diagnostic criteria may be used to reveal probable acute myocarditis if patients' conditions do not allow heart muscle biopsy. Moreover, in specific forms of acute myocarditis, or severe undiagnosticated DCM, it is possible to detect cardiac troponins in blood as support for diagnosis; unfortunately, biomarkers of myocardial damage are usually normal in subacute or chronic myocarditis presentations. Nonetheless, the measurements of cardiac troponin level are recommended in cases of clinically suspected acute myocarditis.

Early myocarditis, even if most of the electrocardiographic and echocardiographic changes are nonspecific, may be strongly suspected by use of speckle-tracking-derived strain. With the advent of new and more sensitive techniques of cardiac magnetic resonance (CMR), diagnostic and prognostic values have been improved even during myocarditis. In particular, T2-weighted and T1-weighted CMR sequences of heart are able to detect edema, inflammation or fibrosis respectively [Gilotra et al., 2016].

2.9 ENDOMYOCARDIAL BIOPSY (EMB)

EMBs play a critical role in the diagnosis of myocarditis, nevertheless, due to the costs, the sensitivity and the specialized expertise, their use as gold standard is limited [Sagar et al., 2012]. The sensitivity of EMBs varies in cases of GCM (approximately 80% to 93%) and for common lymphocytic and sarcoid myocarditis (around 20% to 30%); for DCM instead, immunohistochemistry to detect specific inflammatory cell types and viral genome analysis, have improved the sensitivity to 40%. The immunohistochemical criteria recommended, regard the presence of an inflammatory infiltrate of ≥14 leukocytes/mm2, including up to 4 monocytes/mm2 with the presence of $CD3^+$ T lymphocytes ≥7 cells/mm2 [Caforio et al., 2013]. In cases of subacute or chronic myocarditis (inflammatory cardiomyopathy) characterized by a lower rate of normalization of left ventricular function (about 20%), these criteria result particularly informative and useful. The yield in left ventricular disorders results increased by the use of biopsies, with a stroke risk of about 1:300 to 1:500 [Cooper 2013]. Cardiac biopsies may also be performed in order to measure viral genome levels. In this context, serology assays are not useful as a substitute for biopsy-based diagnosis of viral infection, because serological tests have a low correlation rate with viral genomes observed in heart biopsies. Furthermore, they may be measured both in normal heart and in ischemic or valvular heart disease [Mahfoud et al., 2011, Moimas et al., 2012]. Thus, tests for replicative RNA intermediates, in order to better define active infection, may be required. Nowadays, the presence of cardiac virus related to myocardial disorders has been described only for parvovirus B19 and human herpesvirus (>250 copies/mg DNA and >500 copies/mg DNA respectively), but the only detection of viral nucleic acids (DNA or RNA) is not sufficient to measure active replication of viruses. Only a few studies are focused on testing viral RNA intermediates, to determine active viral replication, and they were mainly oriented to the detection of parvovirus B19. In these cases, the active parvovirus B19 replication was described in patients with severe inflammation state, acute myocarditis and chronic inflammatory cardiomyopathy, but not in DCM samples without inflammation [Bock et al., 2010]. Although antiviral therapies have a great impact in patients affected with viral myocarditis characterized by active virus replication, there is a lack, in clinical practice, of standardized methods focused on the detection of virus-related molecules. Nonetheless, several clinical studies are oriented on the investigation and the use of molecular tests in order to detect active viral replication to increase the clinical impact of EMBs. Other characteristics of EMBs may influence diagnosis and prognosis, for example, transcriptome

analysis was used to discern GCM from other forms of myocarditis such as cardiac sarcoidosis [Kandolin et al., 2015].

Abston ED, Coronado MJ, Bucek A, Bedja D, Shin J, Kim JB, Kim E, Gabrielson KL, Georgakopoulos D, Mitzner W, Fairweather D. Th2 regulation of viral myocarditis in mice: different roles for TLR3 versus TRIF in progression to chronic disease. Clin Dev Immunol. 2012;2012:129486. doi: 10.1155/2012/129486

Abston ED, Coronado MJ, Bucek A, Onyimba JA, Brandt JE, Frisancho JA, Kim E, Bedja D, Sung YK, Radtke AJ, Gabrielson KL, Mitzner W, Fairweather D. TLR3 deficiency induces chronic inflammatory cardiomyopathy in resistant mice following coxsackievirus B3 infection: role for IL-4. Am J Physiol Regul Integr Comp Physiol. 2013 Feb 15;304(4):R267-77. doi: 10.1152/ajpregu.00516.2011. Epub 2012 Dec 19. PMID: 23255589; PMCID: PMC3567355.

Afanasyeva M, Georgakopoulos D, Belardi DF, Ramsundar AC, Barin JG, Kass DA, Rose NR. Quantitative analysis of myocardial inflammation by flow cytometry in murine autoimmune myocarditis: correlation with cardiac function. Am J Pathol. 2004 Mar;164(3):807-15. doi: 10.1016/S0002-9440(10)63169-0. PMID: 14982835; PMCID: PMC1613271.

Barin JG, Baldeviano GC, Talor MV, Wu L, Ong S, Quader F, Chen P, Zheng D, Caturegli P, Rose NR, Cihákova D. Macrophages participate in IL-17-mediated inflammation. Eur J Immunol. 2012 Mar;42(3):726-36. doi: 10.1002/eji.201141737. Epub 2012 Jan 23. PMID: 22161142; PMCID: PMC4292791.

Baruteau AE, Boimond N, Ramful D. Myocarditis associated with 2009 influenza A (H1N1) virus in children. Cardiol Young. 2010 Jun;20(3):351-2. doi: 10.1017/S104795111000020X. Epub 2010 Apr 15. PMID: 20392309.

Blyszczuk P, Kania G, Dieterle T, Marty RR, Valaperti A, Berthonneche C, Pedrazzini T, Berger CT, Dirnhofer S, Matter CM, Penninger JM, Lüscher TF, Eriksson U. Myeloid differentiation factor-88/interleukin-1 signaling controls cardiac fibrosis and heart failure progression in inflammatory dilated cardiomyopathy. Circ Res. 2009 Oct 23;105(9):912-20. doi: 10.1161/CIRCRESAHA.109.199802. Epub 2009 Sep 17. PMID: 19762681.

Bock CT, Klingel K, Kandolf R. Human parvovirus B19-associated myocarditis. N Engl J Med. 2010 Apr 1;362(13):1248-9. doi: 10.1056/NEJMc0911362. PMID: 20357294.

Bonney KM, Engman DM. Autoimmune pathogenesis of Chagas heart disease: looking back, looking ahead. Am J Pathol. 2015 Jun;185(6):1537-47. doi: 10.1016/j.ajpath.2014.12.023. Epub 2015 Apr 7. PMID: 25857229; PMCID: PMC4450315.

Bowles NE, Ni J, Kearney DL, Pauschinger M, Schultheiss HP, McCarthy R, Hare J, Bricker JT, Bowles KR, Towbin JA. Detection of viruses in myocardial tissues by polymerase chain reaction. evidence of adenovirus as a common cause of myocarditis in children and adults. J Am Coll Cardiol. 2003 Aug 6;42(3):466-72. doi: 10.1016/s0735-1097(03)00648-x. PMID: 12906974.

Breinholt JP, Moulik M, Dreyer WJ, Denfield SW, Kim JJ, Jefferies JL, Rossano JW, Gates CM, Clunie SK, Bowles KR, Kearney DL, Bowles NE, Towbin JA. Viral epidemiologic shift in inflammatory heart disease: the increasing involvement of parvovirus B19 in the myocardium of pediatric cardiac transplant patients. J Heart Lung Transplant. 2010 Jul;29(7):739-46. doi: 10.1016/j.healun.2010.03.003.

Caforio AL, Marcolongo R, Jahns R, Fu M, Felix SB, Iliceto S. Immune-mediated and autoimmune myocarditis: clinical presentation, diagnosis and management. Heart Fail Rev. 2013 Nov;18(6):715-32. doi: 10.1007/s10741-012-9364-5. PMID: 23114995.

Chimenti C, Russo A, Pieroni M, Calabrese F, Verardo R, Thiene G, Russo MA, Maseri A, Frustaci A. Intramyocyte detection of Epstein-Barr virus genome by laser capture microdissection in patients with inflammatory cardiomyopathy. Circulation. 2004 Dec 7;110(23):3534-9. doi: 10.1161/01.CIR.0000148823.08092.0E. Epub 2004 Nov 22. PMID: 15557377.

Cihákova D, Sharma RB, Fairweather D, Afanasyeva M, Rose NR. Animal models for autoimmune myocarditis and autoimmune thyroiditis. Methods Mol Med. 2004;102:175-93. doi: 10.1385/1-59259-805-6:175. PMID: 15286386.

Cook SC, Ferguson BA, Cotes RO, Heinrich TW, Schwartz AC. Clozapine-Induced Myocarditis: Prevention and Considerations in Rechallenge. Psychosomatics. 2015 Nov-Dec;56(6):685-90. doi: 10.1016/j.psym.2015.07.002.

Cooper LT Jr, Berry GJ, Shabetai R. Idiopathic giant-cell myocarditis--natural history and treatment. Multicenter Giant Cell Myocarditis Study Group Investigators. N Engl J Med. 1997 Jun 26;336(26):1860-6. doi: 10.1056/NEJM199706263362603. PMID: 9197214.

Cooper LT Jr, Keren A, Sliwa K, Matsumori A, Mensah GA. The global burden of myocarditis: part 1: a systematic literature review for the Global Burden of Diseases, Injuries, and Risk Factors 2010 study. Glob Heart. 2014 Mar;9(1):121-9. doi: 10.1016/j.gheart.2014.01.007. PMID: 25432122.

Cooper LT Jr. Myocarditis. N Engl J Med. 2009 Apr 9;360(15):1526-38. doi: 10.1056/NEJMra0800028. PMID: 19357408; PMCID: PMC5814110.

Cooper LT Jr. Role of left ventricular biopsy in the management of heart disease. Circulation. 2013 Oct 1;128(14):1492-4. doi: 10.1161/CIRCULATIONAHA.113.005395. Epub 2013 Sep 4. PMID: 24004502.

Corsten MF, Heggermont W, Papageorgiou AP, Deckx S, Tijsma A, Verhesen W, van Leeuwen R, Carai P, Thibaut HJ, Custers K, Summer G, Hazebroek M, Verheyen F, Neyts J, Schroen B, Heymans S. The microRNA-221/-222 cluster balances the antiviral and inflammatory response in viral myocarditis. Eur Heart J. 2015 Nov 7;36(42):2909-19. doi: 10.1093/eurheartj/ehv321. Epub 2015 Jul 22. PMID: 26206211.

Corsten MF, Papageorgiou A, Verhesen W, Carai P, Lindow M, Obad S, Summer G, Coort SL, Hazebroek M, van Leeuwen R, Gijbels MJ, Wijnands E, Biessen EA, De Winther MP, Stassen FR, Carmeliet P, Kauppinen S, Schroen B, Heymans S. MicroRNA profiling identifies microRNA-155 as an adverse mediator of cardiac injury and dysfunction during acute viral myocarditis. Circ Res. 2012 Aug 3;111(4):415-25. doi: 10.1161/CIRCRESAHA.112.267443. Epub 2012 Jun 19. PMID: 22715471.

Engler RJ, Nelson MR, Collins LC Jr, Spooner C, Hemann BA, Gibbs BT, Atwood JE, Howard RS, Chang AS, Cruser DL, Gates DG, Vernalis MN, Lengkeek MS, McClenathan BM, Jaffe AS, Cooper LT, Black S, Carlson C, Wilson C, Davis RL. A prospective study of the incidence of myocarditis/pericarditis and new onset cardiac symptoms following smallpox and influenza vaccination. PLoS One. 2015 Mar 20;10(3):e0118283. doi: 10.1371/journal.pone.0118283. PMID: 25793705; PMCID: PMC4368609.

Epelman S, Liu PP, Mann DL. Role of innate and adaptive immune mechanisms in cardiac injury and repair. Nat Rev Immunol. 2015 Feb;15(2):117-29. doi: 10.1038/nri3800. PMID: 25614321; PMCID: PMC4669103.

Frieler RA, Mortensen RM. Immune cell and other noncardiomyocyte regulation of cardiac hypertrophy and remodeling. Circulation. 2015 Mar 17;131(11):1019-30. doi: 10.1161/CIRCULATIONAHA.114.008788. PMID: 25779542; PMCID: PMC4367123.

Gilotra NA, Minkove N, Bennett MK, Tedford RJ, Steenbergen C, Judge DP, Halushka MK, Russell SD. Lack of Relationship Between Serum Cardiac Troponin I Level and Giant Cell

Myocarditis Diagnosis and Outcomes. J Card Fail. 2016 Jul;22(7):583-5. doi: 10.1016/j.cardfail.2015.12.022. Epub 2016 Jan 6. PMID: 26768222.

Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2015 Aug 22;386(9995):743-800. doi: 10.1016/S0140-6736(15)60692-4.

Golpour A, Patriki D, Hanson PJ, McManus B, Heidecker B. Epidemiological Impact of Myocarditis. J Clin Med. 2021 Feb 5;10(4):603. doi: 10.3390/jcm10040603. PMID: 33562759; PMCID: PMC7915005.

Gorbea C, Makar KA, Pauschinger M, Pratt G, Bersola JL, Varela J, David RM, Banks L, Huang CH, Li H, Schultheiss HP, Towbin JA, Vallejo JG, Bowles NE. A role for Toll-like receptor 3 variants in host susceptibility to enteroviral myocarditis and dilated cardiomyopathy. J Biol Chem. 2010 Jul 23;285(30):23208-23. doi: 10.1074/jbc.M109.047464. Epub 2010 May 14. PMID: 20472559; PMCID: PMC2906314.

Gratchev A, Kzhyshkowska J, Köthe K, Muller-Molinet I, Kannookadan S, Utikal J, Goerdt S. Mphi1 and Mphi2 can be re-polarized by Th2 or Th1 cytokines, respectively, and respond to exogenous danger signals. Immunobiology. 2006;211(6-8):473-86. doi: 10.1016/j.imbio.2006.05.017. Epub 2006 Jul 21. PMID: 16920487.

Heymans S, Eriksson U, Lehtonen J, Cooper LT Jr. The Quest for New Approaches in Myocarditis and Inflammatory Cardiomyopathy. J Am Coll Cardiol. 2016 Nov 29;68(21):2348-2364. doi: 10.1016/j.jacc.2016.09.937. PMID: 27884253.

Jenke A, Holzhauser L, Löbel M, Savvatis K, Wilk S, Weithäuser A, Pinkert S, Tschöpe C, Klingel K, Poller W, Scheibenbogen C, Schultheiss HP, Skurk C. Adiponectin promotes coxsackievirus B3 myocarditis by suppression of acute anti-viral immune responses. Basic Res Cardiol. 2014 May;109(3):408. doi: 10.1007/s00395-014-0408-y. Epub 2014 Apr 2. PMID: 24691762.

Jenke A, Wilk S, Poller W, Eriksson U, Valaperti A, Rauch BH, Stroux A, Liu P, Schultheiss HP, Scheibenbogen C, Skurk C. Adiponectin protects against Toll-like receptor 4-mediated cardiac inflammation and injury. Cardiovasc Res. 2013 Aug 1;99(3):422-31. doi: 10.1093/cvr/cvt118. Epub 2013 May 13. PMID: 23674516.

Kandolin R, Lehtonen J, Kupari M. Cardiac sarcoidosis and giant cell myocarditis as causes of atrioventricular block in young and middle-aged adults. Circ Arrhythm Electrophysiol. 2011 Jun;4(3):303-9. doi: 10.1161/CIRCEP.110.959254. Epub 2011 Mar 22. PMID: 21427276.

Kuehl U, Lassner D, Gast M, Stroux A, Rohde M, Siegismund C, Wang X, Escher F, Gross M, Skurk C, Tschoepe C, Loebel M, Scheibenbogen C, Schultheiss HP, Poller W. Differential Cardiac MicroRNA Expression Predicts the Clinical Course in Human Enterovirus Cardiomyopathy. Circ Heart Fail. 2015 May;8(3):605-18. doi: 10.1161/CIRCHEARTFAILURE.114.001475. Epub 2015 Mar 11. PMID: 25761932.

Kühl U, Pauschinger M, Noutsias M, Seeberg B, Bock T, Lassner D, Poller W, Kandolf R, Schultheiss HP. High prevalence of viral genomes and multiple viral infections in the myocardium of adults with "idiopathic" left ventricular dysfunction. Circulation. 2005 Feb 22;111(7):887-93. doi: 10.1161/01.CIR.0000155616.07901.35

Leone O, Pieroni M, Rapezzi C, Olivotto I. The spectrum of myocarditis: from pathology to the clinics. Virchows Arch. 2019 Sep;475(3):279-301. doi: 10.1007/s00428-019-02615-8.

Leuschner F, Courties G, Dutta P, Mortensen LJ, Gorbatov R, Sena B, Novobrantseva TI, Borodovsky A, Fitzgerald K, Koteliansky V, Iwamoto Y, Bohlender M, Meyer S, Lasitschka F, Meder B, Katus HA, Lin C, Libby P, Swirski FK, Anderson DG, Weissleder R, Nahrendorf M. Silencing of CCR2 in myocarditis. Eur Heart J. 2015 Jun 14;36(23):1478-88. doi: 10.1093/eurheartj/ehu225. Epub 2014 Jun 20. PMID: 24950695; PMCID: PMC4465633.

Lv H, Havari E, Pinto S, Gottumukkala RV, Cornivelli L, Raddassi K, Matsui T, Rosenzweig A, Bronson RT, Smith R, Fletcher AL, Turley SJ, Wucherpfennig K, Kyewski B, Lipes MA. Impaired thymic tolerance to α-myosin directs autoimmunity to the heart in mice and humans. J Clin Invest. 2011 Apr;121(4):1561-73. doi: 10.1172/JCI44583. Epub 2011 Mar 23. PMID: 21436590; PMCID: PMC3069776.

Mahfoud F, Gärtner B, Kindermann M, Ukena C, Gadomski K, Klingel K, Kandolf R, Böhm M, Kindermann I. Virus serology in patients with suspected myocarditis: utility or futility? Eur Heart J. 2011 Apr;32(7):897-903. doi: 10.1093/eurheartj/ehq493. Epub 2011 Jan 8. PMID: 21217143.

Marin-Neto JA, Cunha-Neto E, Maciel BC, Simões MV. Pathogenesis of chronic Chagas heart disease. Circulation. 2007 Mar 6;115(9):1109-23. doi: 10.1161/CIRCULATIONAHA.106.624296. PMID: 17339569.

Maron BJ, Udelson JE, Bonow RO, Nishimura RA, Ackerman MJ, Estes NA 3rd, Cooper LT Jr, Link MS, Maron MS; American Heart Association Electrocardiography and Arrhythmias Committee of Council on Clinical Cardiology, Council on Cardiovascular Disease in Young, Council on Cardiovascular and Stroke Nursing, Council on Functional Genomics and Translational Biology, and American College of Cardiology. Eligibility and Disqualification Recommendations for Competitive Athletes With Cardiovascular Abnormalities: Task Force 3: Hypertrophic Cardiomyopathy, Arrhythmogenic Right Ventricular Cardiomyopathy and Other Cardiomyopathies, and Myocarditis: A Scientific Statement From the American Heart Association and American College of Cardiology. Circulation. 2015 Dec 1;132(22):e273-80. doi: 10.1161/CIR.0000000000000239.

Martinez FO, Sica A, Mantovani A, Locati M. Macrophage activation and polarization. Front Biosci. 2008 Jan 1;13:453-61. doi: 10.2741/2692. PMID: 17981560.

Matsumori A, Yutani C, Ikeda Y, Kawai S, Sasayama S. Hepatitis C virus from the hearts of patients with myocarditis and cardiomyopathy. Lab Invest. 2000 Jul;80(7):1137-42. doi: 10.1038/labinvest.3780120. PMID: 10908160.

Moimas S, Zacchigna S, Merlo M, Buiatti A, Anzini M, Dreas L, Salvi A, Di Lenarda A, Giacca M, Sinagra G. Idiopathic dilated cardiomyopathy and persistent viral infection: lack of association in a controlled study using a quantitative assay. Heart Lung Circ. 2012 Dec;21(12):787-93. doi: 10.1016/j.hlc.2012.07.013. Epub 2012 Aug 14. PMID: 22901461.

Myers JM, Cooper LT, Kem DC, Stavrakis S, Kosanke SD, Shevach EM, Fairweather D, Stoner JA, Cox CJ, Cunningham MW. Cardiac myosin-Th17 responses promote heart failure in human myocarditis. JCI Insight. 2016 Jun 16;1(9):e85851. doi: 10.1172/jci.insight.85851. PMID: 27366791; PMCID: PMC4924810.

Noutsias M, Rohde M, Göldner K, Block A, Blunert K, Hemaidan L, Hummel M, Blohm JH, Lassner D, Kühl U, Schultheiss HP, Volk HD, Kotsch K. Expression of functional T-cell markers and T-cell receptor Vbeta repertoire in endomyocardial biopsies from patients presenting with acute myocarditis and dilated cardiomyopathy. Eur J Heart Fail. 2011 Jun;13(6):611-8. doi: 10.1093/eurjhf/hfr014. Epub 2011 Mar 19. PMID: 21422001.

Omura T, Yoshiyama M, Hayashi T, Nishiguchi S, Kaito M, Horiike S, Fukuda K, Inamoto S, Kitaura Y, Nakamura Y, Teragaki M, Tokuhisa T, Iwao H, Takeuchi K, Yoshikawa J. Core

protein of hepatitis C virus induces cardiomyopathy. Circ Res. 2005 Feb 4;96(2):148-50. doi: 10.1161/01.RES.0000154263.70223.13. Epub 2004 Dec 23. PMID: 15618537.

Pagni PP, Traub S, Demaria O, Chasson L, Alexopoulou L. Contribution of TLR7 and TLR9 signaling to the susceptibility of MyD88-deficient mice to myocarditis. Autoimmunity. 2010 Jun;43(4):275-87. doi: 10.3109/08916930903509056. PMID: 20187710.

Pauschinger M, Bowles NE, Fuentes-Garcia FJ, Pham V, Kühl U, Schwimmbeck PL, Schultheiss HP, Towbin JA. Detection of adenoviral genome in the myocardium of adult patients with idiopathic left ventricular dysfunction. Circulation. 1999 Mar 16;99(10):1348-54. doi: 10.1161/01.cir.99.10.1348. PMID: 10077520.

Perez Y, Levy ER, Joshi AY, Virk A, Rodriguez-Porcel M, Johnson M, Roellinger D, Vanichkachorn G, Huskins WC, Swift MD. Myocarditis Following COVID-19 mRNA Vaccine: A Case Series and Incidence Rate Determination. Clin Infect Dis. 2021 Nov 3:ciab926. doi: 10.1093/cid/ciab926. Epub ahead of print. PMID: 34734240.

Riad A, Meyer zu Schwabedissen H, Weitmann K, Herda LR, Dörr M, Empen K, Kieback A, Hummel A, Reinthaler M, Grube M, Klingel K, Nauck M, Kandolf R, Hoffmann W, Kroemer HK, Felix SB. Variants of Toll-like receptor 4 predict cardiac recovery in patients with dilated cardiomyopathy. J Biol Chem. 2012 Aug 3;287(32):27236-43. doi: 10.1074/jbc.M112.369728. Epub 2012 May 29. PMID: 22645142; PMCID: PMC3411065.

Rienks M, Papageorgiou AP. Novel regulators of cardiac inflammation: Matricellular proteins expand their repertoire. J Mol Cell Cardiol. 2016 Feb;91:172-8. doi: 10.1016/j.yjmcc.2016.01.008. Epub 2016 Jan 12. PMID: 26791544.

Sagar S, Liu PP, Cooper LT Jr. Myocarditis. Lancet. 2012 Feb 25;379(9817):738-47. doi: 10.1016/S0140-6736(11)60648-X. Epub 2011 Dec 18. PMID: 22185868; PMCID: PMC5814111.

Schultheiss HP, Kühl U, Cooper LT. The management of myocarditis. Eur Heart J. 2011 Nov;32(21):2616-25. doi: 10.1093/eurheartj/ehr165. Epub 2011 Jun 23. PMID: 21705357.

Small EM, Olson EN. Pervasive roles of microRNAs in cardiovascular biology. Nature. 2011 Jan 20;469(7330):336-42. doi: 10.1038/nature09783. PMID: 21248840; PMCID: PMC3073349.

Su Z, Zhang P, Yu Y, Lu H, Liu Y, Ni P, Su X, Wang D, Liu Y, Wang J, Shen H, Xu W, Xu H. HMGB1 Facilitated Macrophage Reprogramming towards a Proinflammatory M1-like Phenotype in Experimental Autoimmune Myocarditis Development. Sci Rep. 2016 Feb 22;6:21884. doi: 10.1038/srep21884. PMID: 26899795; PMCID: PMC4761996.

Watanabe K, Arumugam S, Sreedhar R, Thandavarayan RA, Nakamura T, Nakamura M, Harima M, Yoneyama H, Suzuki K. Small interfering RNA therapy against carbohydrate sulfotransferase 15 inhibits cardiac remodeling in rats with dilated cardiomyopathy. Cell Signal. 2015 Jul;27(7):1517-24. doi: 10.1016/j.cellsig.2015.03.004.

Weithauser A, Bobbert P, Antoniak S, Böhm A, Rauch BH, Klingel K, Savvatis K, Kroemer HK, Tschope C, Stroux A, Zeichhardt H, Poller W, Mackman N, Schultheiss HP, Rauch U. Protease-activated receptor-2 regulates the innate immune response to viral infection in a coxsackievirus B3-induced myocarditis. J Am Coll Cardiol. 2013 Nov 5;62(19):1737-45. doi: 10.1016/j.jacc.2013.05.076. Epub 2013 Jul 17. PMID: 23871888; PMCID: PMC4077621.

Yan L, Hu F, Yan X, Wei Y, Ma W, Wang Y, Lu S, Wang Z. Inhibition of microRNA-155 ameliorates experimental autoimmune myocarditis by modulating Th17/Treg immune response. J Mol Med (Berl). 2016 Sep;94(9):1063-79. doi: 10.1007/s00109-016-1414-3. Epub 2016 Apr 6. PMID: 27052830.

Zawadowski GM, Klarich KW, Moder KG, Edwards WD, Cooper LT Jr. A contemporary case series of lupus myocarditis. Lupus. 2012 Nov;21(13):1378-84. doi: 10.1177/0961203312456752. Epub 2012 Aug 14. PMID: 22892209.

Zhang P, Cox CJ, Alvarez KM, Cunningham MW. Cutting edge: cardiac myosin activates innate immune responses through TLRs. J Immunol. 2009 Jul 1;183(1):27-31. doi: 10.4049/jimmunol.0800861. Epub 2009 Jun 17. PMID: 19535635; PMCID: PMC2720835.

Zhang Y, Zhang M, Li X, Tang Z, Wang X, Zhong M, Suo Q, Zhang Y, Lv K. Silencing MicroRNA-155 Attenuates Cardiac Injury and Dysfunction in Viral Myocarditis via Promotion of M2 Phenotype Polarization of Macrophages. Sci Rep. 2016 Mar 2;6:22613. doi: 10.1038/srep22613. PMID: 26931072; PMCID: PMC4773853.

Chapter 3

NEXT-GENERATION SEQUENCING AND ANALYSIS

3.1 NEXT-GENERATION SEQUENCING

Next-generation Sequencing (NGS) technologies, also known as massively parallel sequencing or high-throughput sequencing, has completely changed the approaches in several fields of research, offering the opportunity to perform a global investigation of nucleic acids by simultaneous sequencing of millions of DNA or RNA molecules in a short time and relatively low cost [Zhong et al., 2021].

Historically, the first method developed by Frederick Sanger and colleagues in 1977 to determine the nucleotide sequence of DNA by using modified nucleotides (dideoxytriphosphate, ddNTPs), is considered a milestone in the procedure of nucleic acid sequencing. This first-generation technique provided the basis for the development of automated Sanger sequencers, used for the realization of the Human Genome Project, started in 1990, with the aim of determining the three billion base pairs of the human genome [Lander et al., 2001].

The first-generation platforms, have been used for decades in research and clinical genetics [Sanger et al., 1975] and have rapidly evolved, leading to the invention of second and third-generation sequencing technologies, characterized by a decrease of time and experiment costs combined with a huge amount of data produced. Compared with Sanger sequencing, which required many years and billions of dollars to determine human genome sequence, NGS technology currently allows the complete human genome sequencing within a few days, for less than $1,000.

NGS methods have been improved by several commercial companies (i.e. Roche (Roche, Basel, Switzerland), Thermo Fisher (Thermo Fisher, Waltham, MA, USA) and Illumina (San Diego, CA, USA)), which developed their own unique chemistry and laboratory approaches to prepare and sequencing DNA/RNA. Generally, the workflow of different sequencing technology includes these main steps:

1. nucleic acid extraction;
2. library preparation;
3. clonal amplification;
4. sequencing of fragments.

However, as this thesis focuses on the production and analysis of Illumina sequencing data, it will describe the methodologies for Illumina data generation and analyses.

Currently, Illumina platforms are involved in several NGS studies [Wilhelm et al., 2009, Franzosa et al., 2015, Zhou et al., 2015] becoming most popular both in clinical and research settings. Among Illumina platforms, NextSeq 500 (released in 2014), represents a marked improvement compared with previous MiSeq (2011) and HiSeq (2012), characterized by the four-channel sequencing by synthesis (SBS) system, in which each base of cDNA is revealed by individual images. NextSeq500 platforms have a two-channel SBS system which allows the detection of all four base calls with only two images, resulting in reduced imaging capture time and marked decreased costs and time sequencing. NovaSeq (2017), last released Illumina platforms, is the most powerful sequencer to date, with increased performance compared to the others, it produces up to six terabytes of data over 20 billion reads in approximately two days [Zhong et al., 2021].

3.2 ILLUMINA SEQUENCING

Illumina method is based on bridge PCR for the clonal amplification and sequencing by synthesis with reversible termination (RT) technology.

In the library preparation step, the nucleic acids are randomly reduced in fragments of 200-500bp length. Specific adapter sequences are then ligated onto blunt-ended of the small segments, alternatively, “tagmentation” combines the fragmentation and ligation reactions into single step that greatly increases the efficiency of the library preparation process. Adapter-ligated fragments are then PCR amplified and gel purified to create library molecules [Van Dijk et al., 2014].

Bridge amplification and sequencing-by-synthesis method

The Illumina platforms involve the use of solid support called flow cell, which represents a core element of the sequencing. On the flow cell, the solid phase “bridge” amplification is performed. Briefly, single-stranded oligonucleotides, complementary to adapter sequences of library molecules, are hybridized on the flow cell surface. The fragments are denatured and the individual filaments are attached to the surface by one of the 2 ends (figure 3). Also the other terminal is complementary to the oligonucleotides on the flow cell, consequently, the filament can fold to form a bridge. The addition of DNA polymerase and the 4 specific nucleotides

allows the synthesis of a new strand starting from the template. Under denaturing conditions, the two filaments separate and can be anchored to the flow cell at one terminal. The repetition of this amplification phase creates a colony of millions of copies of the filaments on the flow cell to form clusters. Illumina sequencing relies on sequencing by synthesis technology. During the bridge amplification, library molecules act as a template and each independent fragment creates its own cluster on the flow cell. In this context, the main actors are:

- the four specific nucleotides (ddATP, ddGTP, ddCTP, ddTTP) in a modified form, which are labeled with a fluorescent marker of different colors for each base. They are also characterized by the presence of a 3' reversible terminator (RT) which allow the binding of one nucleotide per cycle;
- the particular DNA Polymerase, to include the marked nucleotide in the growing chain.

The use of different fluorescent markers is fundamental in order to discriminate the position of the different bases; each cluster is sequenced by reading the color produced after the addition of each single labeled nucleotide. Briefly, using a laser, excitation and emission of light by the fluorophores are induced. A detector memorizes the light emitted in the individual points of the slide: the specific color of the light identifies which nucleotide has bonded to each point in the clusters. Before proceeding with a second synthesis cycle, both the fluorescent and terminator group are removed from the nucleotide. The incorporated base is then digitally read and converted into a nucleotide sequence. A quality score is assigned to each detected base, and it is determined by the type and intensity of light produced by the cluster. The order of bases from a single cluster is read by the machine and output as a "read" (figure 3) [Metzker 2010].

Figure 3

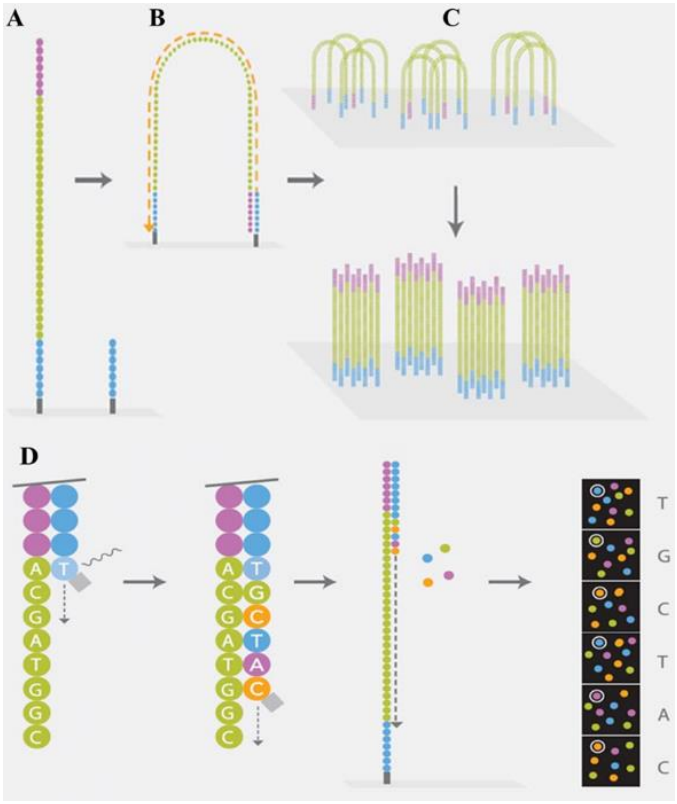


Figure 3 (ref: https://www.cegat.com/services/next-generation-sequencing/). Representation of the bridge amplification phase of the samples and sequencing for the Illumina method. A: the resulting library is hybridized to the flow cell. B: "bridge" folding and amplification. C: cluster generation. D: reading sequences.

3.3 RNA-seq

RNA is a biomolecule that plays an essential role in several biological pathways. RNA-based assays provide a powerful tool for various applications among different fields of human and veterinary health, including disease diagnosis, prognosis and therapeutic selection. Traditional methods for RNA assay represent a limited approach for the study of genome functions because only a single gene at a time (or a small set of them) could be considered. With the evolution of next-generation sequencing (NGS) technologies, the use of RNA sequencing (RNA-seq), to explore the landscape of genomic features, revealed new opportunities. The global approach developed in RNA-seq methods provides an in-depth investigation of the transcriptome (the entire pool of messenger RNA (mRNA) molecules of a sample) and also represents a powerful approach for studying the diversity of RNAs. Compared with other historical methods for the analysis of RNA, such as quantitative reverse transcription PCR (qRT-PCR) for example,

RNA-seq could be considered an open platform technology, that allows quantification of transcripts without pre-defining the RNA targets of interest, resulting in the improved capability to identify and quantify both known and novel transcripts. Transcriptome analysis results in a formidable instrument that is used for a better understanding of the underlying pathways controlling cell fate, development, and disease progression in a host [Byron et al., 2016].

Currently, RNA-seq is most often used in the analysis of differential gene expression (DGE) and represents a useful method for the quantification of transcriptome-wide gene expression. It provides comprehensive, high-quality data that detect quantitative expression levels over the transcriptome. The main stages of a DGE assay begin in the laboratory with the extraction of RNA molecules and the synthesis of cDNA. As previously described, other steps involve the library preparation with adaptor ligation, PCR amplification and then sequencing. Generally, an RNA-seq experiment is performed in order to obtain 20-30 million reads per sample, the data need to be computationally processed, assessing the fraction of reads associated with individual genes or transcripts [Stark et al., 2019].

Recently RNA-seq is also used for pathogen detection in infectious diseases. Considering a large number of clinically important RNA viruses (HIV, West Nile, dengue, hepatitis A, hepatitis D, hepatitis E, coxsackie and influenza viruses, the severe acute respiratory syndrome (SARS) viruses and Middle East respiratory syndrome (MERS) coronaviruses), RNA-seq becomes an interesting approach that could improve knowledge and management of pathogen-related diseases. For example, unbiased non-targeted metagenomic RNA-seq has recently been used to evaluate influenza virus RNA in respiratory fluids along with other viral species [Fischer et al., 2015] and to explore Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV2) coinfection in 2019-2020 pandemic [Abouelkhair 2020]. RNA-seq was also involved in monitoring the Ebola virus during the 2014 outbreak in West Africa and in the case of human papillomavirus (HPV) infection, for example, RNA-seq improved markedly diagnostics [Byron et al., 2016]. Lie et al., in their study, aimed that the detection of viral DNA is not sufficient to diagnose HPV-related tumors, because HPV DNA is observed in ~14% of healthy control women. RNA-based diagnostics have been used for mRNA detection of early viral oncoprotein E6/E7, as a surrogate for active infection, resulting in a better predictive value for cervical cancer [Lie et al., 2008].

One of the advantages of RNA-seq based approaches in the context of infectious diseases, relies on the possibility to evaluate both pathogens and host RNA, exploring consequently the transcriptomic host response to infection. Recently, it was demonstrated for genotype 3 HCV infection, that upregulation of specific host immune genes (interferon beta 1 (IFNB1) interferon lambda 2 (IFNL2) and interferon lambda 3 (IFNL3)) could increase the severity of liver fibrosis. Compared with non-genotype 3 HCV infection, these differentially expressed genes may be considered a risk factor for hepatocellular carcinoma [Mitchell et al., 2015].

3.4 NGS and METAGENOMICS

High throughput sequencing technology, as described above, markedly improved the knowledge and the clinical approach to infectious disease which remain one of the most problematic causes of morbidity and mortality worldwide [Zhong et al., 2021]. Proper and specific detection of pathogens is critical for patient management; delayed or wrong diagnosis, in fact, could be responsible for several adverse events: unnecessary or incorrect use of antibiotics, increased healthcare costs, and worsened patient outcomes [Weng et al., 2017]. As a diagnostic tool, targeted NGS panels of known pathogen sequences, are used to screen clinical isolates and detect microorganism genomes. The panels can be specific for one microorganism or target multiple types of pathogens including bacteria, viruses, and eukaryotic organisms. High specificity and sensitivity are the main advantages of these panels, unfortunately, they are limited to a small number of microorganisms and are not able to detect novel microorganisms related to the investigated disease. In the infectious disease research context, metagenomic NGS (mNGS) can solve these issues. This unbiased approach allows for the detection of different types of pathogens in one sample (and even the host response to them) and can be particularly useful when targeted or less comprehensive tests are not diagnostic [Graf et al., 2016, Doan et al., 2017].

Metagenomics is a new interdisciplinary research field combining molecular genetics, microbial ecology, and data processing [Sudarikov et al., 2017]. Nowadays, the advances in high-throughput sequencing technology allow analyzing the total genomic content of clinical or environmental samples, detecting a large number of organisms and viruses at once, and allowing to correlate the taxonomic composition to the metabolic processes where they are involved [Mendes et al., 2017, Quince et al., 2017]. The main sequencing methods applied for the analysis of metagenome are, the so-called, ‘marker gene’ and ‘shotgun’ sequencing

approaches. These are now widely used compared to the standard methods requiring isolation and cultivation of each single microbial species of interest, which can only provide a limited view, even considering that several microorganisms are not easily cultured in vitro [Lagier et al., 2018]. Marker gene approach (e.g., amplicon sequencing of 16S rRNA gene of bacteria), focuses on a polymorphic region of the microbial genome, that is amplified and sequenced in order to infer at once the composition of the (nearly) entire microbial community present in a given sample [Turnbaugh et al., 2007]. This strategy foresees sequencing of PCR-amplified sequences corresponding to the 16S bacterial rRNA gene which shows some distinctive characteristics that make it suitable for taxonomic profiling and analysis in a single step of a large spectrum of bacteria. In contrast, metagenomic shotgun sequencing (e.g. Whole Metagenome, WMG sequencing) targets all microbial nucleic acids isolated from a sample, followed by deep sequencing of the DNAs drawn randomly from the mixture. The main advantage of this procedure is represented by the large amount of data and information obtained. Contrary to 16S marker gene sequencing, this approach avoids PCR amplification, a major source of potential bias, and provides in-depth knowledge of the microbiome present in the sample, consisting not only of bacterial but also of fungal, protozoal, and viral entities [Martoz et al., 2018, Conrads et al., 2019]. The presence of several microorganisms and their interaction with each other and the host, play an important role in many physio-pathological processes such as reproduction, immune system activity, cancer, and metabolic disorders [Benson et al., 2010, Atreya 2020, Nejman et al., 2020].

3.5 BIOINFORMATIC ANALYSIS

RNA-seq technology for the analysis of differential gene expression (DGE) produces a huge amount of data, that represents a challenge for bioinformatic downstream analysis and requires specific skills and computational resources. Standard analysis workflow starts with the examination of the quality of data files (FASTQ files), which are composed by base-called sequencing reads. Other steps involve the mapping of sequence read to a known transcriptome or annotated genome. This process named “alignment” is achieved using well-documented mapping tools such as STAR, HISAT or TopHat, which assign sequence reads to a specific genome/transcriptome location. After the alignment phase, the analysis proceeds in order to determine the abundance of transcripts. Quantification tools in common use include RSEM, CuffLinks, MMSeq and HTSeq (or the R equivalent, featureCounts). All of them provide

quantification of read abundances for individual genes by counting sequence reads that overlap known genes, using a transcriptome annotation. This process results in a combined expression matrix with a row for each detected expression feature (gene or transcript) and a column for each sample analyzed, with values of read counts. Generally, the counts are filtered and normalized in order to consider differences in read depth, expression patterns and technical biases. However, the next goal in the analysis of DGE is to detect which transcript features are likely to have changed their level of expression. EdgeR, DESeq2 and limma+voom accomplish this information using generalized linear models, that are able to evaluate complex experimental set-ups [Stark et al., 2019].

3.6 BIOINFORMATIC ANALYSIS OF METAGENOMIC DATA

As other NGS technology, data produced by mNGS (WMG or non-targeted metagenomic RNA-seq) consist of large sets of sequenced reads and represents a challenge for bioinformatics analysis and biological interpretation, particularly in clinical samples. WMG data analysis, for example, requires the implementation of multiple ad hoc tools designed, whose installation and configuration, including reference library and computational environment variables, is complex and troublesome in particular for investigators with limited computational expertise. Furthermore, most pipelines available for non-expert users are designed to measure the diversity of microbiota composition from environmental samples (e.g. soil, water) rather than their detection and characterization in biological matrices. Moreover, these pipelines do not provide a global view of sample metagenomic content, since generally they have been designed for specific entities [Rampelli et al., 2016, Milani et al., 2018, Piper et al., 2019, Ji et al., 2020]. In this context, Kodoja [Baizan-Edge et al., 2019] is a workflow specifically developed for plant datasets, identifying only viral sequences from mixed RNAseq data. MetaFlow [Sobih et al., 2016], instead, perform taxonomy profiling using genomic sequences (DNA). It is implemented in C++ and needs a specific input file format (LEMON's LGF): a graph-based representation of the alignments of metagenomic reads in a collection of reference genomes derived from a bacterial domain. Other pipelines such as nf-rnaSeqMetagen [Mpangase et al., 2020] are designed for RNAseq data, and therefore of limited use for metagenomics. Furthermore, this pipeline is implemented in Nextflow [Di Tommaso et al., 2017] and all the applications needed to execute the workflow are containerized in Singularity [Kurtzer et al., 2017]. Nonetheless, it performs quality check step, host filtering, taxonomic profiling of

unmapped reads, and, in order to make taxonomic classification faster, it also runs an assembly de novo step with Trinity [Grabherr et al., 2011]. Metaphlan2 [Segata et al., 2012] is another tool for profiling the composition of microbial communities but, in its present version, is unable to query protozoal databases. In addition, other freely available solutions, such as ASaiM [Batut et al., 2018] or Galaxy [Afgan et al., 2018], are able to perform numerous steps in a metagenomic investigation, incorporating several exhaustive integrated tools. ASaiM, for example, is an Open-Source Galaxy-based framework dedicated to microbiota data analyses, distributed also via Docker and Conda. MGnify [Mitchell et al., 2020], instead, provides a platform for the assembly, analysis and archiving of microbiome data, it requires an internet connection and registration. Therefore, an easy-to-use, comprehensive and specific analysis workflow for shotgun sequencing from samples with variable content of host genetic material, is still unavailable and much required.

Abouelkhair MA. Non-SARS-CoV-2 genome sequences identified in clinical samples from COVID-19 infected patients: Evidence for co-infections. PeerJ. 2020 Nov 2;8:e10246. doi: 10.7717/peerj.10246. PMID: 33194423; PMCID: PMC7643552.

Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Čech M, Chilton J, Clements D, Coraor N, Eberhard C, Grüning B, Guerler A, Hillman-Jackson J, Von Kuster G, Rasche E, Soranzo N, Turaga N, Taylor J, Nekrutenko A, Goecks J. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Res. 2016 Jul 8;44(W1):W3-W10. doi: 10.1093/nar/gkw343. Epub 2016 May 2. PMID: 27137889; PMCID: PMC4987906.

Atreya CE, Turnbaugh PJ. Probing the tumor micro(b)environment. Science. 2020 May 29;368(6494):938-939. doi: 10.1126/science.abc1464. PMID: 32467376.

Baizan-Edge A, Cock P, MacFarlane S, McGavin W, Torrance L, Jones S. Kodoja: A workflow for virus detection in plants using k-mer analysis of RNA-sequencing data. J Gen Virol. 2019 Mar;100(3):533-542. doi: 10.1099/jgv.0.001210. Epub 2019 Jan 24. PMID: 30676315.

Batut B, Gravouil K, Defois C, Hiltemann S, Brugère JF, Peyretaillade E, Peyret P. ASaiM: a Galaxy-based framework to analyze microbiota data. Gigascience. 2018 Jun 1;7(6):giy057. doi: 10.1093/gigascience/giy057. PMID: 29790941; PMCID: PMC6007547.

Benson AK, Kelly SA, Legge R, Ma F, Low SJ, Kim J, Zhang M, Oh PL, Nehrenberg D, Hua K, Kachman SD, Moriyama EN, Walter J, Peterson DA, Pomp D. Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. Proc Natl Acad Sci U S A. 2010 Nov 2;107(44):18933-8. doi: 10.1073/pnas.1007028107. Epub 2010 Oct 11. PMID: 20937875; PMCID: PMC2973891

Byron SA, Van Keuren-Jensen KR, Engelthaler DM, Carpten JD, Craig DW. Translating RNA sequencing into clinical diagnostics: opportunities and challenges. Nat Rev Genet. 2016 May;17(5):257-71. doi: 10.1038/nrg.2016.10. Epub 2016 Mar 21. PMID: 26996076; PMCID: PMC7097555.

Conrads G, Abdelbary MMH. Challenges of next-generation sequencing targeting anaerobes. Anaerobe. 2019 Aug;58:47-52. doi: 10.1016/j.anaerobe.2019.02.006. Epub 2019 Feb 12. PMID: 30769104.

Di Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, Notredame C. Nextflow enables reproducible computational workflows. Nat Biotechnol. 2017 Apr 11;35(4):316-319. doi: 10.1038/nbt.3820. PMID: 28398311.

Doan T, Acharya NR, Pinsky BA, Sahoo MK, Chow ED, Banaei N, Budvytiene I, Cevallos V, Zhong L, Zhou Z, Lietman TM, DeRisi JL. Metagenomic DNA Sequencing for the Diagnosis of Intraocular Infections. Ophthalmology. 2017 Aug;124(8):1247-1248. doi: 10.1016/j.ophtha.2017.03.045. Epub 2017 May 16. PMID: 28526549; PMCID: PMC5533302.

Fischer N, Indenbirken D, Meyer T, et al. Evaluation of Unbiased Next-Generation Sequencing of RNA (RNA-seq) as a Diagnostic Method in Influenza Virus-Positive Respiratory Samples. J Clin Microbiol. 2015;53(7):2238-2250. doi:10.1128/JCM.02495-14

Franzosa EA, Hsu T, Sirota-Madi A, Shafquat A, Abu-Ali G, Morgan XC, et al. Sequencing and beyond: integrating molecular 'omics' for microbial community profiling. Nat Rev Microbiol 2015;13:360-72.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011 May 15;29(7):644-52. doi: 10.1038/nbt.1883. PMID: 21572440; PMCID: PMC3571712.

Graf EH, Simmon KE, Tardif KD, Hymas W, Flygare S, Eilbeck K, Yandell M, Schlaberg R. Unbiased Detection of Respiratory Viruses by Use of RNA Sequencing-Based Metagenomics: a Systematic Comparison to a Commercial PCR Panel. J Clin Microbiol. 2016 Apr;54(4):1000-7. doi: 10.1128/JCM.03060-15. Epub 2016 Jan 27. PMID: 26818672; PMCID: PMC4809917.

Ji Y, Huotari T, Roslin T, Schmidt NM, Wang J, Yu DW, Ovaskainen O. SPIKEPIPE: A metagenomic pipeline for the accurate quantification of eukaryotic species occurrences and intraspecific abundance change using DNA barcodes or mitogenomes. Mol Ecol Resour. 2020 Jan;20(1):256-267. doi: 10.1111/1755-0998.13057. Epub 2019 Aug 7. PMID: 31293086.

Kurtzer GM, Sochat V, Bauer MW. Singularity: Scientific containers for mobility of compute. PLoS One. 2017 May 11;12(5):e0177459. doi: 10.1371/journal.pone.0177459. PMID: 28494014; PMCID: PMC5426675.

Lagier JC, Dubourg G, Million M, Cadoret F, Bilen M, Fenollar F, Levasseur A, Rolain JM, Fournier PE, Raoult D. Culturing the human microbiota and culturomics. Nat Rev Microbiol. 2018 May 1;16:540-550. doi: 10.1038/s41579-018-0041-0. PMID: 29937540.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C,Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C,Stange-Thomann N, Stojanovic N, Subramanian A et al. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature. 2001 Feb 15;409(6822):860-921.

Lie AK, Kristensen G. Human papillomavirus E6/E7 mRNA testing as a predictive marker for cervical carcinoma. Expert Rev Mol Diagn. 2008 Jul;8(4):405-15. doi: 10.1586/14737159.8.4.405. PMID: 18598223.

Marotz CA, Sanders JG, Zuniga C, Zaramela LS, Knight R, Zengler K. Improving saliva shotgun metagenomics by chemical host DNA depletion. Microbiome. 2018 Feb 27;6(1):42. doi: 10.1186/s40168-018-0426-3. PMID: 29482639; PMCID: PMC5827986.

Mendes LW, Braga LPP, Navarrete AA, Souza DG, Silva GGZ, Tsai SM. Using Metagenomics to Connect Microbial Community Biodiversity and Functions. Curr Issues Mol Biol. 2017;24:103-118. doi: 10.21775/cimb.024.103. Epub 2017 Jul 6. PMID: 28686570.

Metzker ML. Sequencing technologies - the next generation. Nat Rev Genet. 2010 Jan;11(1):31-46. doi: 10.1038/nrg2626. Epub 2009 Dec 8. PMID: 19997069.

Milani C, Casey E, Lugli GA, Moore R, Kaczorowska J, Feehily C, Mangifesta M, Mancabelli L, Duranti S, Turroni F, Bottacini F, Mahony J, Cotter PD, McAuliffe FM, van Sinderen D, Ventura M. Tracing mother-infant transmission of bacteriophages by means of a novel analytical tool for shotgun metagenomic datasets: METAnnotatorX. Microbiome. 2018 Aug 20;6(1):145. doi: 10.1186/s40168-018-0527-z. PMID: 30126456; PMCID: PMC6102903.

Mitchell AL, Almeida A, Beracochea M, Boland M, Burgin J, Cochrane G, Crusoe MR, Kale V, Potter SC, Richardson LJ, Sakharova E, Scheremetjew M, Korobeynikov A, Shlemov A, Kunyavskaya O, Lapidus A, Finn RD. MGnify: the microbiome analysis resource in 2020. Nucleic Acids Res. 2020 Jan 8;48(D1):D570-D578. doi: 10.1093/nar/gkz1035. PMID: 31696235; PMCID: PMC7145632.

Mitchell AM, Stone AE, Cheng L, Ballinger K, Edwards MG, Stoddard M, Li H, Golden-Mason L, Shaw GM, Khetani S, Rosen HR. Transmitted/founder hepatitis C viruses induce cell-type- and genotype-specific differences in innate signaling within the liver. mBio. 2015 Feb 24;6(2):e02510. doi: 10.1128/mBio.02510-14. PMID: 25714713; PMCID: PMC4357998.

Mpangase PT, Frost J, Ramsay M, Hazelhurst S. nf-rnaSeqMetagen: a nextflow metagenomics pipeline for identifying and characterizing microbial sequences from RNA-seq data. Med Microecol. 2020;4:100011.

Nejman D, Livyatan I, Fuks G, Gavert N, Zwang Y, Geller LT, Rotter-Maskowitz A, Weiser R, Mallel G, Gigi E, Meltser A, Douglas GM, Kamer I, Gopalakrishnan V, Dadosh T, Levin-Zaidman S, Avnet S, Atlan T, Cooper ZA, Arora R, Cogdill AP, Khan MAW, Ologun G, Bussi Y, Weinberger A, Lotan-Pompan M, Golani O, Perry G, Rokah M, Bahar-Shany K, Rozeman EA, Blank CU, Ronai A, Shaoul R, Amit A, Dorfman T, Kremer R, Cohen ZR, Harnof S, Siegal T, Yehuda-Shnaidman E, Gal-Yam EN, Shapira H, Baldini N, Langille MGI, Ben-Nun A, Kaufman B, Nissan A, Golan T, Dadiani M, Levanon K, Bar J, Yust-Katz S, Barshack I, Peeper DS, Raz DJ, Segal E, Wargo JA, Sandbank J, Shental N, Straussman R. The human tumor microbiome is composed of tumor type-specific intracellular bacteria. Science. 2020 May 29;368(6494):973-980. doi: 10.1126/science.aay9189. PMID: 32467386; PMCID: PMC7757858.

Piper AM, Batovska J, Cogan NOI, Weiss J, Cunningham JP, Rodoni BC, Blacket MJ. Prospects and challenges of implementing DNA metabarcoding for high-throughput insect surveillance. Gigascience. 2019 Aug 1;8(8):giz092. doi: 10.1093/gigascience/giz092. PMID: 31363753; PMCID: PMC6667344.

Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis. Nat Biotechnol. 2017 Sep 12;35(9):833-844. doi: 10.1038/nbt.3935. Erratum in: Nat Biotechnol. 2017 Dec 8;35(12):1211. PMID: 28898207.

Rampelli S, Soverini M, Turroni S, Quercia S, Biagi E, Brigidi P, Candela M. ViromeScan: a new tool for metagenomic viral community profiling. BMC Genomics. 2016 Mar 1;17:165. doi: 10.1186/s12864-016-2446-3. PMID: 26932765; PMCID: PMC4774116.

Sanger F and Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol 1975; 94:441-8.

Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. Nat Methods. 2012 Jun 10;9(8):811-4. doi: 10.1038/nmeth.2066. PMID: 22688413; PMCID: PMC3443552.

Sobih A, Tomescu AI, Mäkinen V. Metaflow: Metagenomic profiling based on whole-genome coverage analysis with min-cost flows. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics). 2016;9649:111–21.

Stark R, Grzelak M, Hadfield J. RNA sequencing: the teenage years. Nat Rev Genet. 2019 Nov;20(11):631-656. doi: 10.1038/s41576-019-0150-2. Epub 2019 Jul 24. PMID: 31341269.

Sudarikov K, Tyakht A, Alexeev D. Methods for The Metagenomic Data Visualization and Analysis. Curr Issues Mol Biol. 2017;24:37-58. doi: 10.21775/cimb.024.037. Epub 2017 Jul 6. PMID: 28686567.

Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. Nature. 2007 Oct 18;449(7164):804-10. doi: 10.1038/nature06244. PMID: 17943116; PMCID: PMC3709439.

Van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. Ten years of nextgeneration sequencing technology. Trends Genet 2014;30:418-26

Weng QY, Raff AB, Cohen JM, Gunasekera N, Okhovat JP, Vedak P, Joyce C, Kroshinsky D, Mostaghimi A. Costs and Consequences Associated With Misdiagnosed Lower Extremity Cellulitis. JAMA Dermatol. 2017 Feb 1;153(2):141-146. doi: 10.1001/jamadermatol.2016.3816. PMID: 27806170.

Wilhelm BT and Landry JR. RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing. Methods 2009; 48:249-57

Zhong Y, Xu F, Wu J, Schubert J, Li MM. Application of Next Generation Sequencing in Laboratory Medicine. Ann Lab Med. 2021 Jan;41(1):25-43. doi: 10.3343/alm.2021.41.1.25. Epub 2020 Aug 25. PMID: 32829577; PMCID: PMC7443516.

Zhou J, He Z, Yang Y, Deng Y, Tringe SG, Alvarez-Cohen L. Highthroughput metagenomic technologies for complex microbial community analysis: open and closed formats. mBio 2015;6:e02288-14 .

Chapter 4

EXPERIMENTAL PART

4.1 MATERIALS AND METHODS

4.1.1 SAMPLE COHORT

This study includes 13 dogs with myocardial disorders documented with the presence of a severe arrhythmia and three control samples downloaded from the short read archive under NCBI BioProject PRJNA78827. This control dataset contained RNAseq data from heart tissue of healthy dog euthanized following owner consent and ethical approval. Ethical approval was granted by Uppsala Djurförsöksetiska Nämnd (Permit Number: C 2/12) [Hoeppner et al., 2014].

Concerning 13 dogs enrolled in the study, supraventricular tachycardia (1/13), ventricular tachycardia (4/13) premature ventricular contractions (8/13) were recorder at 12-lead surface ECG and holter monitoring. All the animals enrolled in this study required right ventricular (RV) catheterization, performed at Clinica Veterinaria Malpensa (Dipartimento di Cardiologia, Clinica Veterinaria Malpensa, Samarate, Italy). The study population included 4 females and 9 males. The median weight was 27.3 kg (range: 10-49 kg). Breeds of dogs included mixed breed (n=4), Corso (n=2), Amstaff (n=1), Border Collie (n=1), Setter Gordon (n=1), Vizsla (n=1), Weimaraner (n=1), Labrador retriever (n=1) and Whippet (n=1) as described in table 1. Of all the animals enrolled, 11 of them showed a dilatated cardiomopathy (DMC) and 1 showed an arrhythmogenic right ventricular cardiomyopathy (ARVC) (table 1). The animals were undergone echocardiographic evaluation that included M-Mode, 2-D, Colour Flow Mapping, pulse wave and continuous wave Doppler using specific transducers (GE Logic-9 Expert Machine, GE medical system, Phoenix, AZ, USA). Routinary hemato-biochemical parameters were within the normal reference ranges for dogs (data not shown).

The study protocol was approved by the Ethical Animal Care and Use Committee of University of Naples Federico II (Protocol No. 67990/2015). A valid informed consent was obtained from all owners prior to inclusion in the study.

Table 1

Sample ID	*Age*	*Weight*	*Breed*	*Sex*	*Diagnosis*
Sample_001	10 y 8 m	24 kg	Amstaff	female	DCM
Sample_002	4 y	21 kg	Border Collie	male	DCM
Sample_003	6 m	32 kg	Mixed-breed	male	DCM
Sample_004	4 m	18.8 kg	Setter Gordon	male	Supraventricular Tachycardia
Sample_005	5 y	27.3 kg	Mixed-breed	male	DCM
Sample_006	3 y	33 kg	Labrador	male	ARVC
Sample_007	4 y 8 m	NA	Weimaraner	male	DCM
Sample_008	6 m	19.8 kg	Vizsla	male	DCM
Sample_009	4 y	49 kg	Corso	male	VPC
Sample_010	7 y 9 m	44 kg	Corso	female	DCM
Sample_011	9 m	10 kg	Mixed-breed	male	DCM
Sample_012	3 y	22.6 kg	Mixed-breed	female	DCM
Sample_013	5 y	NA	Whippet	female	DCM

Table 1: table summarizing the characteristic of EMB samples used in the analysis (DCM= dilated cardiomyopathy, ARCV= arrythmogenic right ventricular tachycardia, VPC= ventricular premature contraction).

4.1.2 EMBs COLLECTION

EMBs were collected under general anesthesia using a modified Seldinger technique throughout the jugular vein access, using a long introducer (Cook, 7Fr, Flexor, 45 cm) placed across the tricuspid valve annulus. Along with this device, a bioptome (Sholten, Su103-50, 7F pre-curved) was introduced to get the biopsies. For each animal, at least four biopsy was collected. All biopsy samples were submitted for histologic examination, PCR assay (for pathogen detection) and NGS experiment, that was performed at Laboratorio di Medicina Molecolare e Genomica (University of Salerno). Part of EMB samples was collected in RNAse

free tubes with RNAlater solution (Thermo Fisher Scientific) in order to preserve the integrity of RNA during collection and shipping. The biopsies were stored at 2-8 °C overnight and then frozen at -20 °C (o -80 °C). These samples were then shipped in dry ice at Laboratorio di Medicina Molecolare e Genomica (University of Salerno), to perform the NGS experiment. Moreover, EMB samples, from each dog, underwent evaluation by qualitative PCR assay (ad hoc optimized) for detection of cardiotropic pathogens' nucleic acids, as summarized in table 2. EMBs were shipped to the laboratory of Microbiology at Dipartimento di Scienze Veterinarie, University of Torino (Italy). In detail, using specific primers, different pathogens were tested, including canine coronavirus, canine herpesvirus 1, canine distemper virus, canine adenovirus 1, canine adenovirus 2, canine parvovirus 2, West Nile virus, and Bartonella, according to Santilli and collaborators (2019).

Table 2

Purpose	***Target gene***	***Primer sequence (forward)***	***Primer sequence (reverse)***
Control extraction			
Deoxyribonucleic acid	GAPDH	GTTCCAGTATGATTCCACCC	TCCCTCCACGATGCCAAA
Ribonucleic acid	Na+/K+ ATPase α	GCTGACTTGGTCATCTGC	AGGTAGGTTTGAGGGGATAC
Pathogen detection			
Canine adenovirus 1	E3	CGCGCTGAACATTACTACCTTGTC	CCTAGAGCACTTCGTGTCCGCTT
Canine adenovirus 2	F	TGTCAACAAGGTTTTGTCTTTT	TTTTCAAGGGAGGTGCGT
Bartonella spp	16S-23S	CTTCGTTTCTCTTTCTTCA	GGATAAACCGGAAAACCTTC
Canine coronavirus	M	TCCAGATATGTAATGTTCGG	TCTGTTGAGTAATCACCAGCT
Canine herpesvirus 1	K	TGCCGCTTTTATATAGATG	AAGCGTTGTAAAAGTTCGT
Canine parvovirus 2	VP2	CATTGGGCTTACCACCATTTCCAACC	TCAGCTGGTCTCAT
Canine distemper virus	N	GATAAAGCATGTCATTATAGTCCTAA	CTTGAGCTTTCGACCCTTC
West Nile virus	NS5	GCMATHAGGTWCATGTGG	GTRTCCCAKCCDGCNGTRTC

Table 2: Primers used for PCR assay, to detect pathogens in EMB samples collected from dogs enrolled in this study, during RV catheterization.

4.1.3 HISTOLOGIC EXAMINATION OF EMB SAMPLES

Light microscopy—Formalin-fixed EMB samples were processed by routine methods, embedded in paraffin wax, and sectioned at a thickness of 5 μm. Twenty-four serial sections were prepared for each EMB sample. The samples were valued at the laboratory of Animal Pathology at Dipartimento di Scienze Veterinarie, University of Torino (Italy); 12 were evaluated histologically to provide a good representation of the entire sample, and 12 were used for immunohistochemical analysis. Sections were numbered sequentially, and the 12 sections for histologic examination were mounted on glass slides and stained with H&E stain (slides 1, 6, 11, 16, 21, and 24) for morphological evaluation or with Masson trichrome stain (slides 2, 7, 12, 17, 22, and 23) to detect collagen deposition. Each EMB sample was assessed by a single individual for evidence of myocyte hypertrophy, sarcoplasm vacuolization, fibrosis, and lymphocytic inflammation.

4.1.4 SAMPLE PREPARATION AND SEQUENCING

Endomyocardial biopsies were processed to perform total RNA sequencing. Total RNA was extracted from a single biopsy using Norgen's RNA/DNA Purification kit (Norgen Biotek Corporation), following the manufacturer's instructions. This kit allows to sequentially isolate and purify both total RNA and DNA from a single sample. The purification step takes advantage of spin column chromatography that uses Norgen's proprietary Silicon Carbide resin separation matrix. In particular, a two-column system is used: one for DNA and one for RNA; moreover, the RNA column permits the purification of total RNA. Before proceeding with the synthesis of the libraries, the extracted RNA samples were subjected to standard quality controls. The integrity of each sample (RIN) was assessed by automated electrophoresis on the Tapestation system (Agilent Technologies). To exclude proteins and organic solvent contamination, the absorbance ratios 260/280 and 260/230 at the Nanodrop-2000C instrument were measured. Finally, the total RNA was quantized by Qubit fluorimeter (Thermofisher) with RNA HS Assay Kit. Indexed libraries were prepared from 7 ng of purified RNA with SMARTer Stranded Total RNA-seq kit v2-pico Input Mammalian (Takarabio) according to the manufacturer's instructions. This kit enables to produce indexed cDNA libraries suitable for NGS on any Illumina platform and is particularly indicated for poorly concentrated RNA samples. Indeed, the kit takes advantage of SMART (Switching Mechanism At 5' end of RNA Template) cDNA synthesis technology for the generation of Illumina-compatible libraries via

PCR amplification. Briefly, total RNA was converted to cDNA, then adapters for Illumina sequencing (with specific barcodes) were ligated by PCR. PCR products were purified with magnetic beads and then ribosomal cDNA was depleted. This last step is particularly important for the success of the final analysis since it improves mapping statistics. In fact, it is known that Ribosomal RNA (rRNA) represents more than 90% of all RNA molecules present in total RNA extracted; as consequence removing these abundant transcripts from the samples prior to generating libraries provides an important vantage. cDNA fragments were amplified via PCR using universal primers. Finally, the PCR products were purified to increase the final yield of the libraries. After synthesis, libraries were quantified using the Agilent 2100 Bioanalyzer (Agilent Technologies) and pooled with a final concentration of 2 nM. The pooled samples were sequenced using an Illumina NextSeq500 System (Illumina Inc.) in a 2x75 paired-end read format at a final loading concentration of 1,7 pmol.

4.1.5 RNA-Seq DATA ANALYSIS

Raw Illumina intensity data were demultiplexed and converted to fastq files (.fastq) using the Illumina software bcl2fastq (Illumina Inc.). Fastq files from 3 control samples were downloaded from the short read archive under NCBI BioProject PRJNA78827 [Hoeppner et al., 2014]. Dataset consisting in RNAseq experiment to generate a comprehensive catalog of transcripts for the dog genome, from different tissue types. Only raw sequence files from hearth samples, with accession SRX111067, were used in the analysis.

Fastq files were quality checked using FastQC software [http://www.bioinformatics.babraham.ac.uk/projects/fastqc/], then adapter sequences were removed and low-quality reads were filtered out using Trimmomatic software (version 0.39) [Bolger et al., 2014] with parameters set as follow: LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15 MINLEN:20. The resulting high-quality reads were then mapped to the dog reference genome (Dog10K_Boxer_Tasha, RefSeq accession: GCF_000002285.5) with STAR software (version 2.7.4a) [Dobin et al., 2012] with default parameters. The number of reads mapping to each transcript within the reference genome was computed with FeatureCounts v2.0.0 [Liao et al., 2014]. The counts were then imported in in R package DESeq2 v1.26.0 (R version 3.6.3) [Love et al., 2014] and DGE analysis was performed comparing sequenced samples from dog EMBs and control samples from the downloaded project. Differentially expressed (DE) mRNAs were identified after the DESeq2 internal normalization method. Differential expression was reported as fold change |1.5| along with

associated adjusted *p* values (FDR ≤ 0.05), computed according to Benjamini–Hochberg (Benjamini et al., 2001). Gene Ontology (GO) of DE genes was performed with ShinyGO [Ge et al., 2020] and only functions and pathways showing a p-value ≤ 0.01 were considered.

4.1.6 HOME-BIO PIPELINE IMPLEMENTATION

HOME-BIO was implemented using both tools mainly used in metagenomic analyses and custom python scripts to produce tables and charts for immediate and easier interpretation of the results. The central core of this pipeline is its modularity, being composed of three main blocks: 'Quality Control', 'Metagenomic Shotgun' and 'Assembly de novo' modules (figure 4). They can be run all together or separately, according to the user's needs. The "Quality Control" module allows to perform sequence read quality checks and includes FASTQC and MultiQC [Ewels et al., 2016], to perform the quality check and the summary of quality control, respectively, while the adapter trimming and removal of low-quality reads is performed by Cutadapt [Martin 2011]. If required, HOME-BIO performs a filtering step to remove host and contaminant sequence fragments by mapping each of them on the corresponding genomes. This alignment is performed with bowtie [Langmead et al., 2009], for input reads with length less than 50 bp and with bowtie2 [Langmead et al., 2012] for those longer than 50 bp. The 'Metagenomic Shotgun' module performs taxonomic profiling by classifying unmapped reads with Kraken2 [Wood et al., 2019]. By default, the confidence score threshold of 0.5 is used to define the quality of taxonomic classification. To extract as much information as possible, bacterial, viral, and protozoal NCBI databases built by Kraken2 are provided. Nonetheless, if users intend to explore only one of the domains described before (or only viral entities), it is possible to set custom options to query only one (or more) databases and obtain the taxonomic information desired. Compared to other taxonomic profiling pipelines freely available, such as MicroPro [Zhu et al., 2019], METAwrap [Uritskiy et al., 2018], HOME-BIO performs an additional protein-validation step for non-eukaryotic entities. This implementation makes taxonomic classification outcomes more robust by using protein-level classification with Kaiju [Menzel et al., 2016], thereby increasing the reliability and sensitivity of the analysis. Kaiju carries out a comparison to a reference database containing microbial and viral protein sequences. Unmapped metagenomic input reads are thus translated into amino acid sequences and then searched in the database using a modified backward search on a memory-efficient implementation of the Burrows-Wheeler transform, which finds maximum exact matches. As default, it was set the evaluation in run mode greedy, with an e-value cutoff of 0.001. All the

entities classified with Kaiju are then processed with Krona [Ondov et al., 2011], to obtain a comprehensive graphic visualization. As outputs, the 'Metagenomic shotgun' module generates a table containing the Kranken2 taxonomy profile and related Kaiju protein-validation information. A given taxon is considered protein-validated when both tools classify and assign reads to it. In addition, HOME-BIO generates output pie-charts with the top 15 represented species, with an estimation of the relative abundance of each of them. In the 'Assembly de novo' module analysis, HOME-BIO uses SPAdes [Nurk et al., 2013] in metagenomic mode (option -meta). It takes in input unmapped reads resulting from the 'Quality Control' module, generating sequence contigs that are classified with Kraken2, and protein-validated with Kaiju as described before. As mentioned above, HOME-BIO takes advantage of several reference databases. This allows users to investigate in depth the content of their biological samples. If it ran in end-to-end mode, the pipeline provides comprehensive profiling of specimens by querying the above-mentioned databases. This aspect makes HOME-BIO a powerful tool in the hands of users dealing with metagenomic data. To make its installation easier, HOME-BIO is inserted in a Docker [Merkel 2014] image, with installed Ubuntu 18.04.4 LTS and Anaconda 3 (V. 02/2020) (https://anaconda.com). This allows that all the required tools are always compatible with each other, automatically updated, and handled by Anaconda itself. All the parameters are set using a configuration file that is given as input file when the pipeline is launched.

The pipeline was tested on a public dataset [Mitra et al., 2015] composed by 12 samples, divided in patients affected with symptomatic atherosclerotic plaques (n=7) and matched asymptomatic controls (n=5). Furthermore, pipeline performances were tested in a benchmark analysis using mock communities from Singer and coll. [Singer et al., 2016] dataset. HOME-BIO results were compared with those obtained on the same dataset with Metaphlan2.

Figure 4

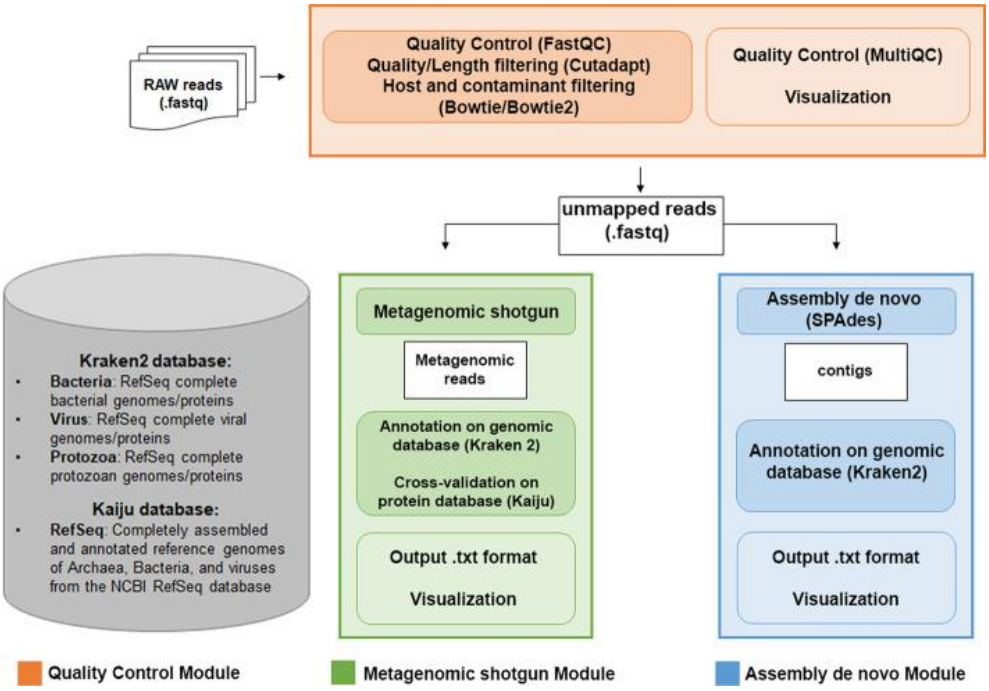


Figure 4 (reference HOME-BIO DOI: 10.1186/s12859-021-04004-y): HOME-BIO workflow. The pipeline accepts NGS data as input and then proceeds to perform the analysis according to the different steps comprised in the three modules. The Quality Control module (in orange) performs QCs of the reads, low-complexity reads removal, and host and contaminant sequencing reads filtering (indicated in the dark orange rectangle), generating as output quality reports and graphical charts (indicated in the light orange rectangle). The sequence reads maintained after this step are processed using the Metagenomic shotgun module (in green) and/or the Assembly de novo module (in blue). In the first case, the reads are processed by Kraken2 and Kaiju (indicated in the dark green rectangle), producing as output text tables and charts. The Assembly de novo module processes the sequences by assembling with SPAdes and annotating with Kraken2 (dark blue rectangle), generating as output text tables and charts. For both modules, a custom database (indicates in the grey box) is used.

4.1.7 HOME-BIO ANALYSIS

For taxonomy profiling, raw fastq files were imported in the HOME-BIO pipeline performing all three modules analysis. Config file was set on RNA molecule in paired-end mode, in order to perform quality control before and after adapter removal. Adapter sequence given in input corresponded to Illumina universal adapter: AGATCGGAAGAG. Dog reference genome (RefSeq accession: GCF_000002285.5) was used for the host filtering process, by fixing the "path host genome" option. Taxonomy classification of samples was obtained by querying the kraken2 database containing bacterial (RefSeq complete bacterial genomes/proteins), viral (RefSeq complete viral genomes/proteins) and protozoal (RefSeq complete protozoan

genomes/proteins) information. Furthermore, in order to exclude possible contaminations not filtered out in the previous alignment step, sequences related to dog and human genomes have been added to the database. An example of the configuration file is reported in figure 5.

Figure 5

```
[DEFAULT]
quality control before trimming [yes/no] = no
quality control after trimming [yes/no] = no
number of threads [n] = 1
paired-end? [yes/no] = yes
filter out contaminant [yes/no] = no
adapter? = CTGTCTCTTATA
shotgun module [yes/no] = no
assembly denovo module [yes/no] = no
kraken2 confidence = 0.5
nucleic acid type [dna/rna] = DNA
k-mer [auto/21,33,55,77] = auto

[CUSTOM]
quality control before trimming [yes/no] = yes
quality control after trimming [yes/no] = yes
number of threads [n] = 16
path fastq = /root/HOME-BIO-master/fastq/
paired-end? [yes/no] = yes
path output = /root/HOME-BIO-master/output/
path host genome = /root/HOME-BIO-master/Genomes/dog/STAR_index/
filter out contaminant [yes/no] = no
path contaminant genome = /root/Scrivania/HOME-BIO-master/Genome/
adapter? = AGATCGGAAGAG
shotgun module [yes/no] = yes
assembly denovo module [yes/no] = no
path kraken2 & kaiju databases = /root/Scrivania/HOME-BIO-master/databases/
kraken2 confidence = 0.5
nucleic acid type [dna/rna] = RNA
k-mer [auto/21,33,55,77] = auto
```

Figure 5: example of the configuration file used in the analysis. The custom section was set in order to obtain quality control before and after cleaning of low-quality reads and host-related fragments. Furthermore, reads taxonomy annotation was set by fixing the shotgun module option with a confidence score of 0.5. 16 processors were selected using the "number of threads" option and no contaminant filtering (by alignment) was carried out.

4.2 RESULTS

The 13 dogs enrolled did not show any complications during and after the EMB procedures. In total, 19 EMBs were obtained from 13 samples (from Sample_002, Sample_004, Sample_007, Sample_008, Sample_009 and Sample_010, 2 biopsis were collected). Histological diagnosis of fibrosis was identified in 10/13 animals. Endocardial fibrosis was described in 8/13 dogs although in two of these dogs was observed also the presence of interstitial fibrosis (2/8); while in one dog only interstitial fibrosis was detected. Three dogs with fibrosis showed also the presence of inflammatory infiltrates, in particular, one dog (Sample_005) showed lymphoplasmacytic and macrophage infiltrates, while two dogs (Sample_002 and Sample_008) showed lymphocytic infiltrate (figure 6). Inflammatory infiltrate confirmed by using immunohistochemistry (data not shown), according to the Dallas criteria in absence of necrosis, confirm a diagnosis of mild form of myocarditis. In 8/13 dogs, signs of hypertrophy of cardiomyocytes were detected while necrosis was present only in 3/13 samples. Furthermore, 3/13 samples (Sample_003, Sample_005 and Sample_008) were positive to PCR assay for pathogen discovery. In detail, nucleic acid from canine parvovirus 2 was detected in 2 EMB samples (Sample_003 and Sample_005), while Sample_008 was positive for canine distemper virus.

Figure 6

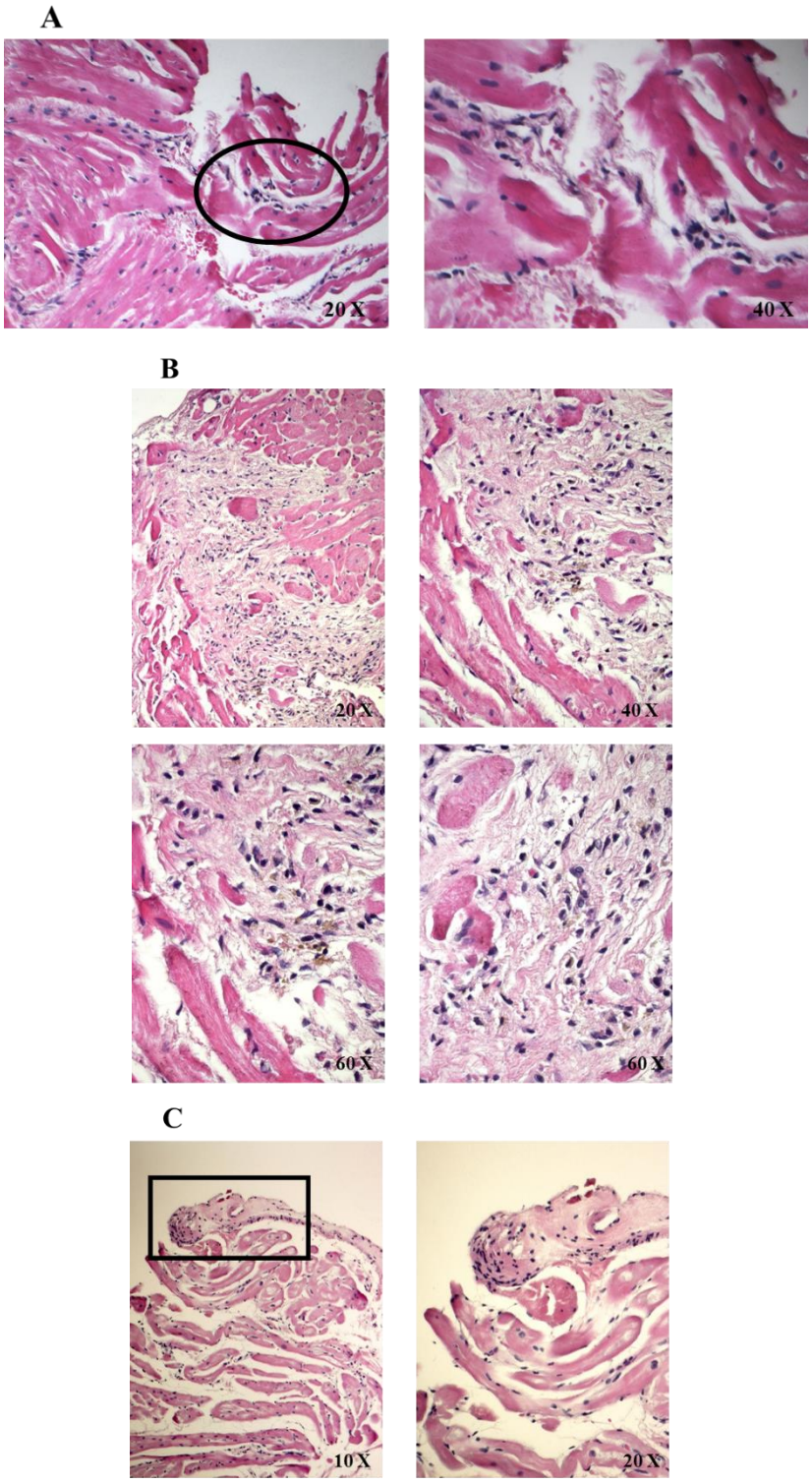


Figure 6: Photomicrographs of EMB samples from 3 dogs showing the presence of inflammatory infiltrates. H&E stain. Sample_002 (A) were associated with focal and minimal interstitial lymphocyte infiltrate; Sample_005 (B) characterized by interstitial fibrosis associated with moderate lympho-plasmacytic and macrophage infiltrate (with hemosiderosis); C: Sample_008 showed endocardial fibrosis associated with moderate lymphocytic infiltrate.

Total RNA extracted from the EMB samples presented a variable yield, going from a minimum of 28 ng to a maximum of 672 ng in total, depending on the size of the starting tissue. Furthermore, for some samples, the quantization of the extracted RNA at the Qubit fluorimeter was not possible as their concentration was below the minimum value (5 ng/ul) detectable by the instrument. However, the high sensitivity kit used for library preparation (SMARTer Stranded Total RNA-seq kit v2-pico Input Mammalian – Takara) allowed cloning even for poorly concentrated samples. On the basis of the quality control analysis of the RNA samples, the RIN (RNA Integrity Number) values ranged from values of 5.5 to 9.4, thus indicating in some cases a partial degradation of the sample, but on average integrity suitable for downstream applications. Successful libraries synthesis and amplification produced distinct curves spanning 200–1000 bp, with a local maximum at ~400–500 bp, in agreement with protocol guidelines. Comparison of the results with relative controls confirmed that EMB samples were all eligible for subsequent sequencing. The sequencing of NGS libraries from EMB samples was conducted on Illumina NextSeq500 and produced a total of 779,380,298.00 reads with an average of 41,020,015.68 reads per sample (range 21,853,604 - 104,856,322).

4.2.1 TESTING HOME-BIO

HOME-BIO processes NGS data to characterize the composition of host-associated microbial communities. The pipeline is implemented to be used by biologists with limited computational experience who intend to explore the pathogens' content in their samples. To test the performance of HOME-BIO, the public dataset of Mitra and collegues (2015) has been used, comprising 12 samples: 7 from symptomatic atherosclerotic plaques and 5 matched controls from asymptomatic atherosclerotic plaques. The analysis was performed running all three modules with default parameters and querying the databases provided with the pipeline [https://zenodo.org/record/4055180], comprising complete RefSeq bacterial, viral, and protozoa genome sequences and a subset of NCBI BLAST non-redundant database containing all proteins belonging to Archaea, Bacteria, and Viruses. To extract more confident sequences to be assigned to biological entities, the reads were mapped against the genome of the host (human genome assembly hg19). After running the Quality Control module, the number of unmapped reads, usable in the next computational step, ranged from 6,449,408 to 14,710,470, with an average value of 7,991,037.20 for Controls and 11,841,558.00 for Cases samples, representing 26.17% and 41.63% of the total reads in Control and Case, respectively. Subsequently, for each sample, the reads were analyzed with the Shotgun Metagenomic module, in order to obtain the taxonomy profile. HOME-BIO queried bacterial, viral, and

protozoal databases. The Quality Control and Metagenomic results are summarized in table 3 and figure 7. Running the Assembly de novo module, 567,379.57 contigs for Cases samples and 238,788.8 contigs for Control samples were obtained. HOME-BIO identified an average of 5,575.2 and 12,956.3 contigs classified as bacteria in Control and Cases samples, respectively. Among these, an average of 455.4 in Control and 816 in Cases were contigs classified as viruses, while 12,219.8 contigs in Control and 18,423.8 in Cases matched in the protozoal database. These results obtained on bacteria from the analysis of the 12 samples are in agreement with the ones described by Mitra and collaborators (2015).

Table 3

Sample-name	Raw reads	Non-hg19 %	HOME-BIO shotgun Bacteria reads classified (%)	HOME-BIO shotgun Virus reads classified (%)	HOME-BIO shotgun Protozoa reads classified (%)
SRR1205226	62,241,508	10.36	4,617 (0.07)	331 (0.01%)	177,713 (2.76)
SRR1205227	83,273,796	8.79	259,578 (3.54)	595 (0.01%)	237,811 (3.25)
SRR1205228	57,336,562	16.43	5,955 (0.06)	677 (0.01%)	218,329 (2.32)
SRR1205230	75,597,642	11.50	52,529 (0.6)	2,496 (0.03)	262,034 (3.01)
SRR1205231	76,089,590	10.60	5,316 (0.07)	697 (0.01)	217,939 (2.7)
SRR1205232	70,667,046	20.82	7,420 (0.05)	1,152 (0.01)	278,959 (1.9)
SRR1206003	76,651,484	12.06	139,334 (1.51)	736 (0.01)	264,054 (2.86)
SRR1206005	77,326,222	15.57	8,235 (0.07)	720 (0.01%)	264,974 (2.2)
SRR1206007	50,795,748	22.63	3,713 (0.03)	684 (0.01)	174,687 (1.52)
SRR1206009	49,963,912	21.03	4,393 (0.04)	491 (0.00)	202,508 (1.93)
SRR1206011	41,425,332	28.19	3,434 (0.03)	732 (0.01%)	196,978 (1.69)
SRR1206012	52,047,446	25.38	5,146 (0.04)	748 (0.01)	230,780 (1.75)

Table 3: Total number and percentage of reads assigned to Bacteria, Viruses, and Protozoa for each analyzed sample in the test dataset.

Figure 7

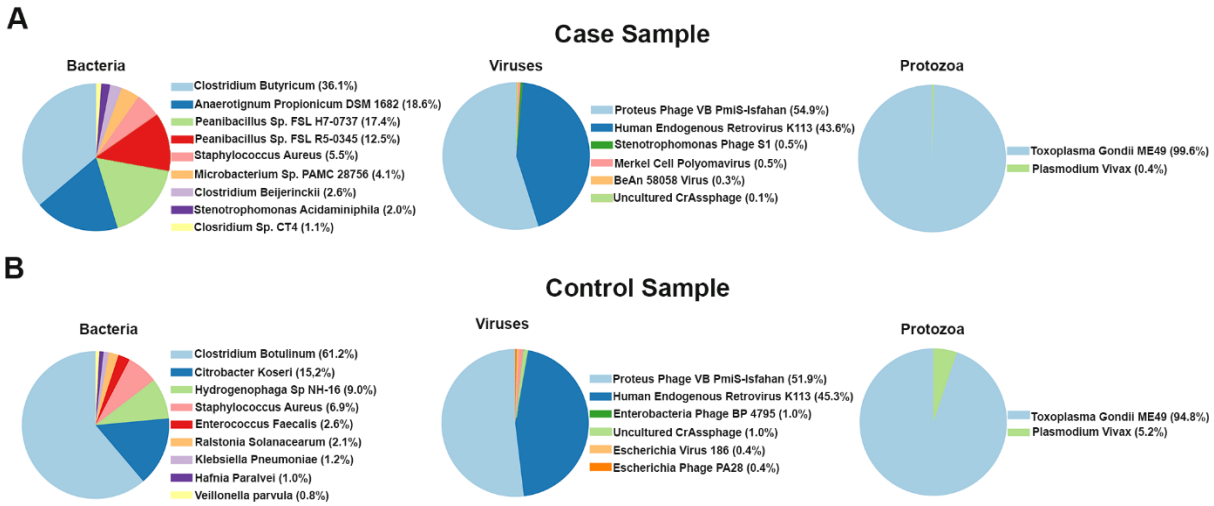


Figure 7: Pie charts examples summarizing the percentage of Bacteria, Viruses, and Protozoa detected in samples SRR1206003 (Case example) (A) and SRR1202227 (Control example) (B) obtained by data analysis with the HOME-BIO pipeline.

In addition, HOME-BIO was tested by performing a benchmark analysis using mock communities and comparing the results with those obtained with other tools often used in metagenomics investigation. In detail, the mock community dataset from Singer and coll. (2016) was downloaded and analyzed with HOME-BIO and Metaphlan2. As shown in figure 8, the results obtained with HOME-BIO are in agreement with the microbiota profiling described in the paper. In some cases, HOME-BIO was able to assign more reads to biological entities when compared with the results obtained using Metaphlan2. From the comparison of the results obtained from the two pipelines, HOME-BIO resulted more informative, because it was able to identify the lowest taxon-level as described in Singer and coll. (2016). This aspect is not covered by Metaphlan2.

About the taxonomic classification of contigs, using the mock community described above, HOME-BIO was able to classify 63.06% of the assembled contigs to the species-level, while CAMITAX classified only 1 contig to taxon level “superkingdom bacteria”.

Figure 8

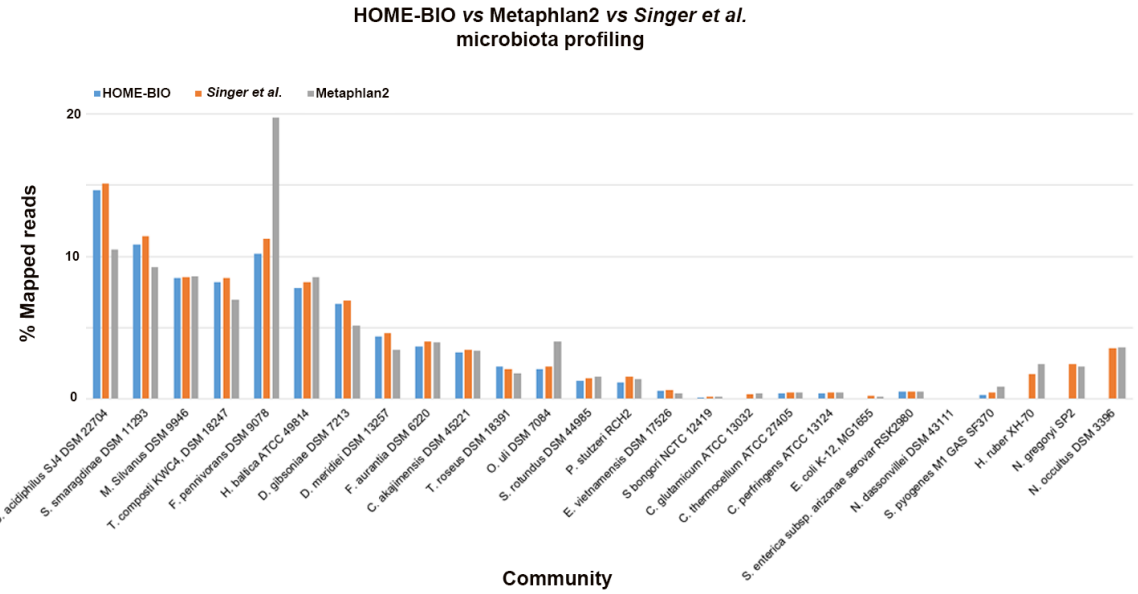


Figure 8: comparison of HOME-BIO and Metaphlan2 results in test dataset described in Singer and coll. (2016).

4.2.2 HOME-BIO ANALYSIS OF EMB SAMPLES

It has been investigated the possible relationship between cardiac muscle disorders and possible infective pathogens. Here, it has been performed taxonomy profiling analysis of EMB samples using the HOME-BIO pipeline, a useful tool consisting of public software and custom bioinformatics scripts. HOME-BIO is a modular pipeline, able to take in input raw sequence files (.fastq files), perform quality check of reads by filtering noninformative sequences and allow taxonomy classification using bacterial, viral and protozoal comprehensive databases. All EMB fastq files were imported in the pipeline by properly setting the configuration file, in order to perform "quality control module" before and after cleaning and to analyze resulting reads with "metagenomic shotgun module". The first step removed, from the dataset, reads below the quality value threshold of 15. According to results shown in the "SEQUENCING AND DIFFERENTIAL GENE EXPRESSION ANALYSIS" section, the filtering step in HOME-BIO resulted in 2.55% of reads removed from the whole EMB dataset due to low quality and sequencing adapter related fragments. The reports of the quality filtering step are shown in figure 9.

Figure 9

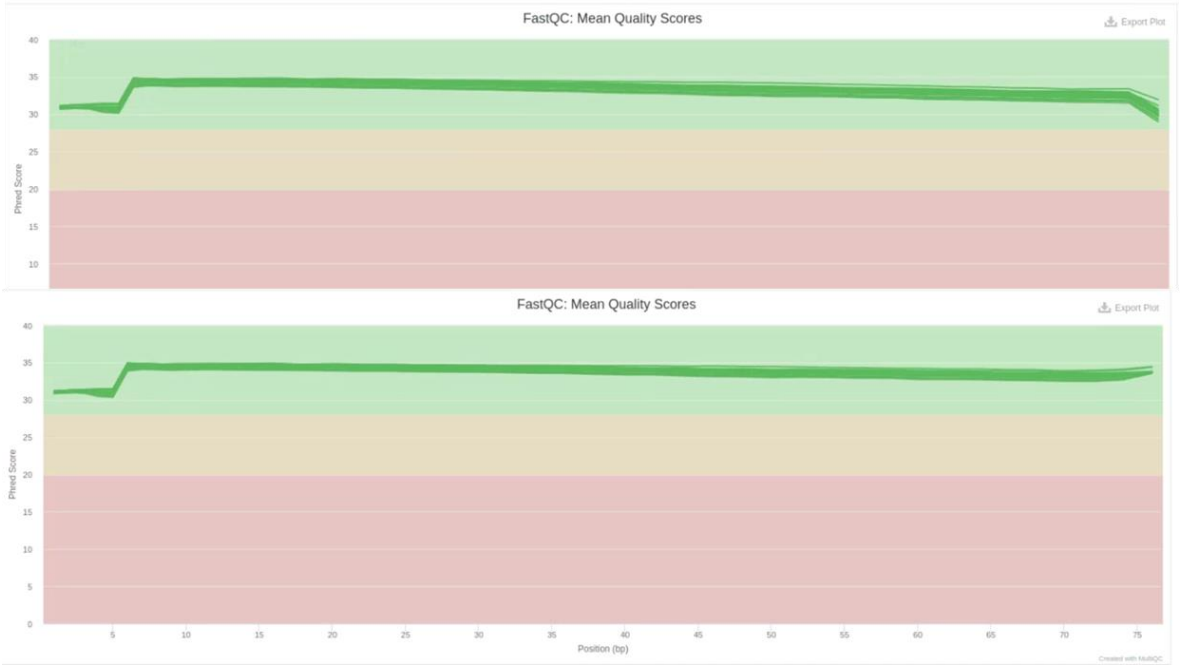


Figure 9: aggregated reports of quality filtering step obtained with HOME-BIO "quality control module". The quality reports of EMB samples before (on top) and after (on the bottom) cleaning are reported. The mean quality of all reads belonging to a sample is reported as green lines and each line represented a samples. As shown, after trimming, the mean quality of reads was notably increased (right portion of figures).

The "quality control module" also filtered out, all fragments matching to the dog reference genome (GCF_000002285.5) provided through the configuration file. The results of this alignment were in agreement with those obtained in the section "SEQUENCING AND DIFFERENTIAL GENE EXPRESSION ANALYSIS" with an average of mapping ~90%. After filtering of host sequences, ~45,4 million total reads were obtained for downstream analysis, about 6.05% of reads per sample, corresponding to an average of 2,628,194.842 high-quality reads per EMB sample. These unmapped reads were then processed for further investigation in order to perform taxonomy classification. On the entire dataset of EMB samples, HOME-BIO was able to classify ~29% of reads per sample, with an average of 714,805.68 reads per sample (range 160,656 – 2,247,868 reads) on the used database (figure 10).

Figure 10

Figure 10: histogram reporting the percentage (on the top) of classification in EMB samples. On the entire dataset ~29% of reads were classified (orange bars) and 71% resulted unclassified (green bars). On the bottom histogram, classification rates are reported as the number of reads with classified and unclassified reads in orange and green respectively.

The database provided for the analysis, consisting of NCBI RefSeq complete genomes/proteins information of bacteria, virus, protozoa and, in order to perform a further filtering step for

increased the accuracy of the analysis, sequences belonging to *Homo sapiens* and *Canis lupus familiaris* were added. HOME-BIO detected an average of 148,858.10 and 60,131.47 of reads per sample, assigned to *Homo sapiens* and *Canis lupus familiaris* respectively, corresponding to an average of 6.9% (for humans) and 2.9% (for dog) of reads imported in the taxonomy classification step. Regarding pathogens, bacteria were the most abundant biological entities detected in all EMBs dataset, while protozoa and viruses represented only a small portion of classified entities on all unmapped reads (0.003% and 0.0004% respectively). Bacteria were detected with an average of 358,556.95 reads (range 35,836-1,713,904 reads) that corresponded to 12.33% of all unmapped fragments on the reference genome (figure 11).

Figure 11

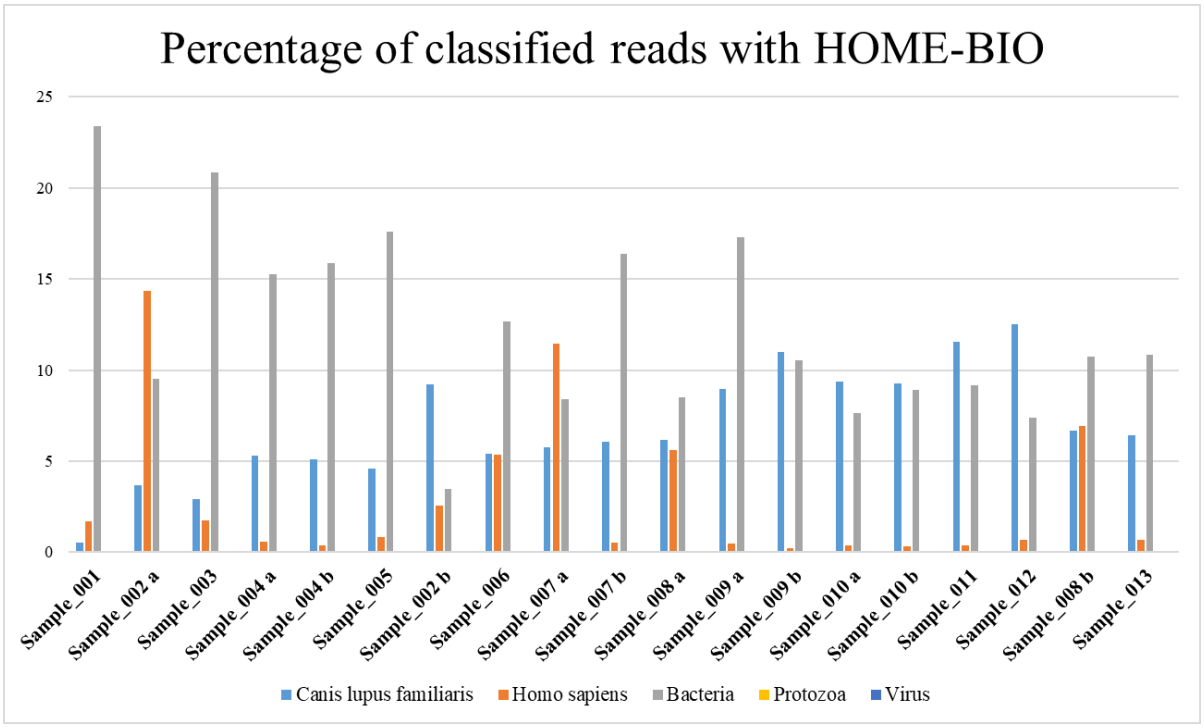


Figure 11: histogram reporting percentage of all unmapped reads classified as Canis Lupus familiaris (light blue) Homo sapiens (orange), bacteria (grey), protozoa (yellow) and virus (blue).

A total of 59 genera and 51 species of bacteria were detected with a read count >2 in at least 50% of dataset. The most abundant 20 genera identified in the analysis with a significant threshold were: *Agrobacterium*, *Cutibacterium*, *Burkholderia*, *Mesorhizobium*, *Acinetobacter*, *Rhodococcus*, *Paraburkholderia*, *Pseudomonas*, *Rhizobium*, *Micrococcus*, *Deinococcus*, *Brevundimonas*, *Streptococcus*, *Staphylococcus*, *Lactobacillus*, *Paracoccus*, *Cupriavidus*, *Corynebacterium*, *Stenotrophomonas* and *Dietzia* (figure 12).

Figure 12

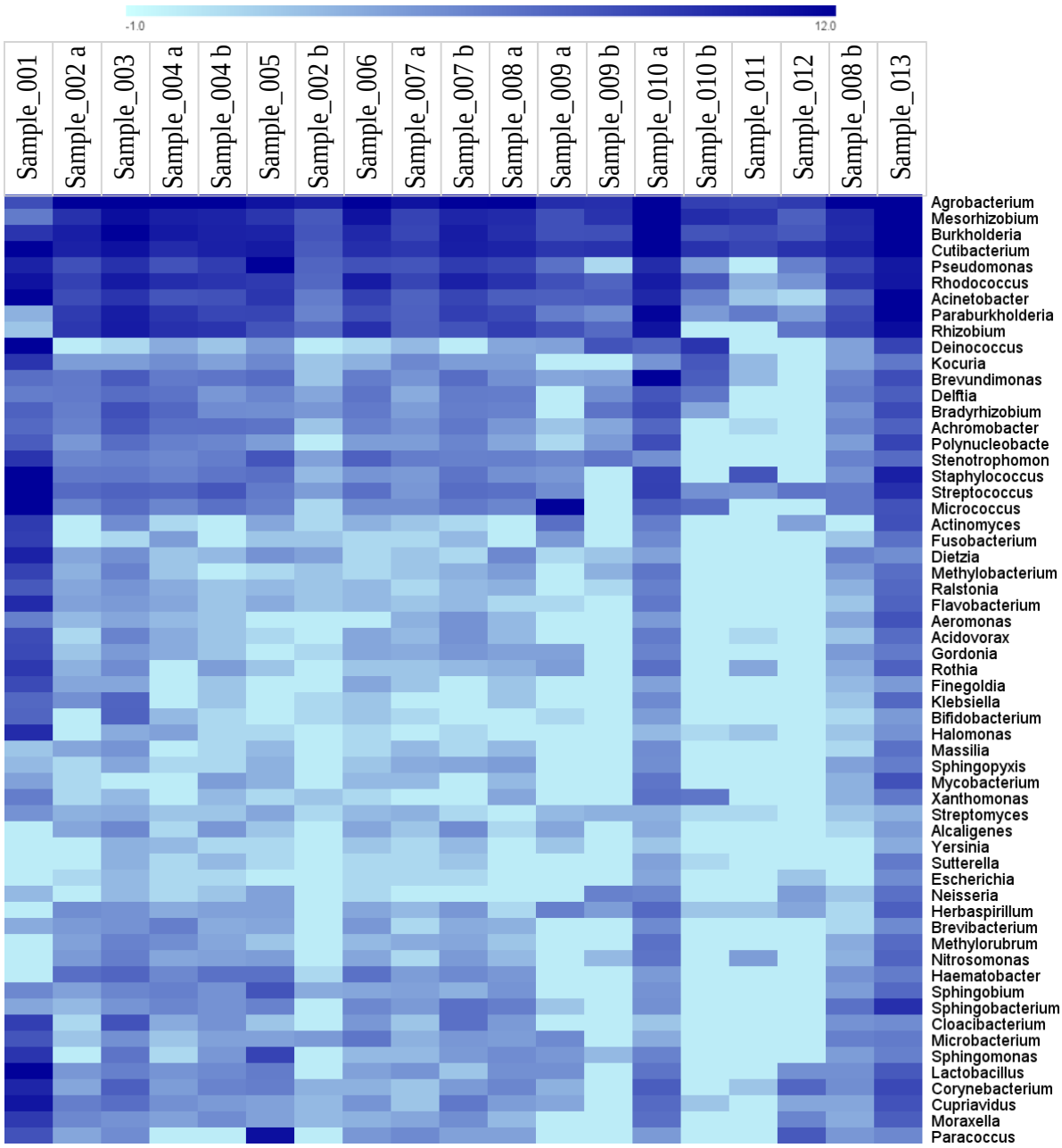


Figure 12: heatmap of 59 bacterial genera in EMBs dataset detected with HOME-BIO with >2 read count in at least 50% of analyzed samples. Counts are expressed as log2 values.

Concerning protozoal entities, which represented 0.003% of all unmapped reads, HOME-BIO allowed the identification of 13 genera and 20 species in all dataset. Detected protozoal genera were: *Phaeodactylum*, *Thalassiosira*, *Plasmodium*, *Theileria*, *Babesia*, *Toxoplasma*, *Cryptosporidium*, *Bigelowiella*, *Leishmania*, *Cryptomonas*, *Guillardia*, *Dictyostelium* and *Giardia*. Virus-related reads, classified by HOME-BIO, corresponded to 0.0004% of all reads

that did not find a match on the dog reference genome. Most abundant viruses detected were represented by *Betaretrovirus* (*Mason-Pfizer monkey virus*), *Mardivirus* (*Columbid alphaherpesvirus 1*) and *Pahexavirus* genus.

4.2.3 DIFFERENTIAL EXPRESSION ANALYSIS

In order to investigate the molecular bases of myocarditis, and the possible relationship to DCM, the transcriptomic landscape of EMBs of dogs affected with reported cardiac disorders was assessed with an RNAseq experiment. Gene expression profiling was carried out, comparing RNAseq data from sequenced EMBs (with an average of 41,020,015.68 reads per sample), with control samples downloaded from NCBI BioProject PRJNA78827, as described in the Materials and Methods section. The three control samples consisted in 61,373,812 of reads with an average of 20,457,937.33 of reads per samples (range 20,429,448 - 20,484,644 reads). In detail, the Differential Expression Analysis (DEA) was conducted between the 10 EMB samples from dogs diagnosed with DCM and downloaded controls. In the first step of quality check, after trimming of low-quality reads and adapters related fragments, a total of 814,484,304 reads (average 37,022,013.82 of reads per sample) were obtained, corresponding to ~ 96.9% of the entire dataset. These high-quality reads have been used in downstream analysis. Alignment on dog genome showed a high percentage of mapping reads. In detail, ~ 91.4% of all analyzed sequences matched on the dog reference genome, with an average of 36,205,727.47 and 17,765,470 of mapped reads for EMB and control samples respectively (table 4).

Table 4

	Total samples	***EMB samples***	***Control samples***
Total Raw sequences (mean)	840,754,110 (38,216,095.91)	779,380,298 (41,020,015.68)	61,373,812 (20,457,937.33)
Total high quality sequences (mean)	814,484,304 (37,022,013.82)	759,474,100 (39,972,321.05)	55,010,204 (18,336,734.67)
Total removed sequences (mean)	26,269,806 (1,194,082.09)	19,906,198 (1,047,694.63)	6,363,608 (2,121,202.67)
Total sequences mapped to dog reference genome (mean)	741,205,232 (33,691,146.91)	687,908,822 (36,205,727.47)	53,296,410 (17,765,470.00)
Percentage of mapping on dog reference genome	91.37 %	90.50 %	96.88 %

Table 4: number of reads before and after quality check and amount of sequences that aligned on the dog reference genome (GCF_000002285.5), results are reported as the sum of all reads and as mean.

Considering all alignment files (.bam), matrix count has been computed with FeatureCounts v2.0.0 by counting reads mapping to each transcript on the reference genome. The dog reference genome has more than 2.3 billion total sequence length, consists of 35,518 annotated genes distributed in 78 chromosomes (39 pairs). According to FeatureCounts results, 8,930 genes were detected with read count >10 in at least 80% of analyzed samples. The counts were imported in R statistical environment and normalized with the DESeq2 internal method. Based on the distribution of the Euclidean distance of analyzed samples, 3 biopsy cases (EMB_001, EMB_010 and EMB_014 corrisponding to Sample_001, Sample_007 and Sample_010 respectively) were excluded from the downstream analysis, because of their similiraty with control samples and were defined as outliers. Furthermore, Pearson correlation was computed and showed that the outlier samples clustered closely to the control samples, indicating bad relationship among variables of the same group, as described in figure 13. This filter removed one sample (Sample_001) from the downstream analysis, while for the other 2 outliers, the second biopsis remained.

Figure 13

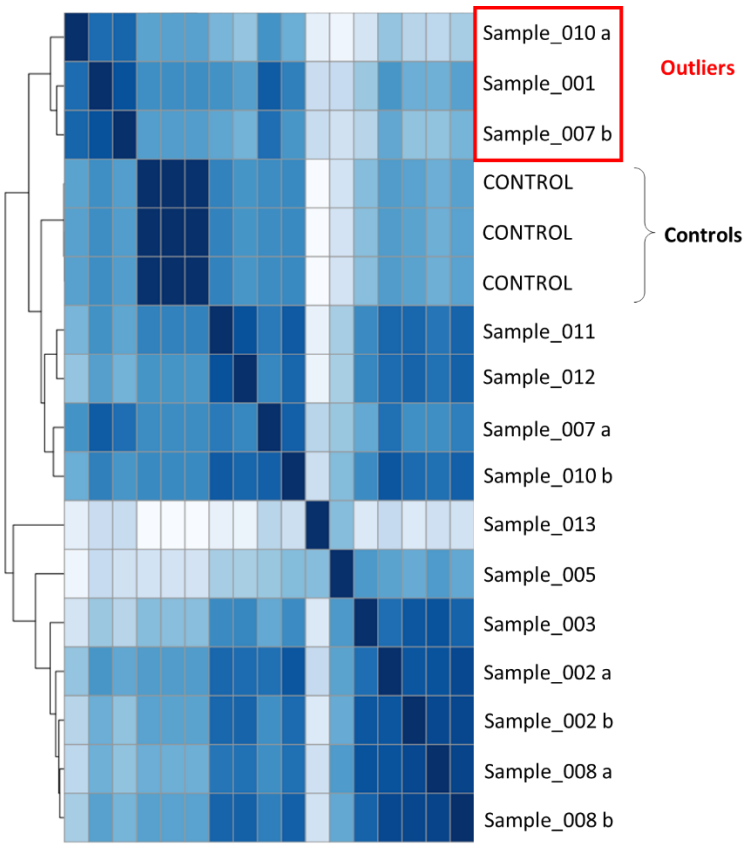


Figure 13: Heatmap of Euclidean distance showing similarity among samples. 3 biopsy cases were defined outliers (in red) because they clustered closely control samples (in black).

Only with the retained samples, PCA (Principal Component Analysis) and differential expression analysis (DEA) were performed. PCA results showed moderate clustering of samples as a function of the biological condition, according to PC1 (~ 47%) and a relatively good biological variability between components of the same type according to PC2 (~ 26%) as described in figure 14.

Figure 14

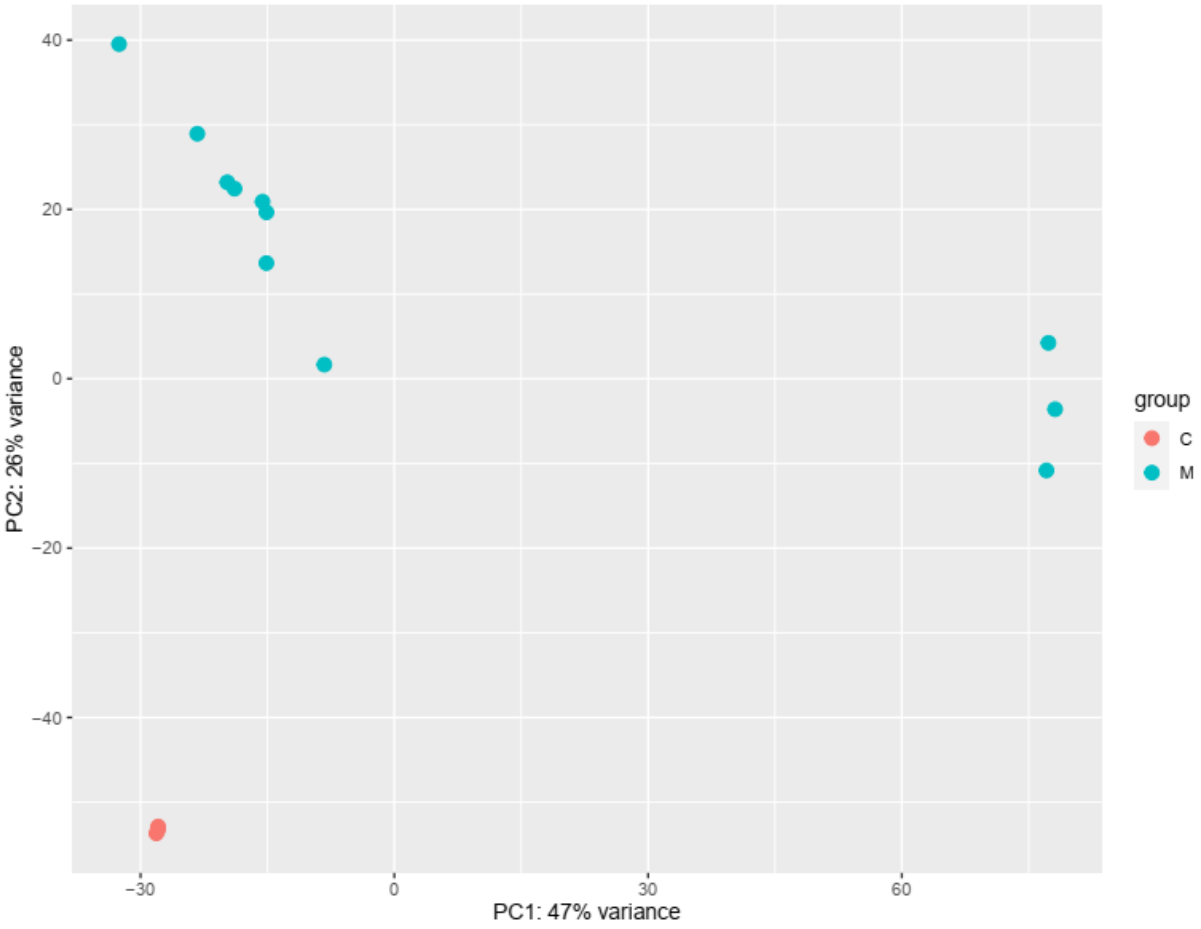


Figure 14: Principal Component Analysis (PCA) of RNAseq samples divided into two groups: in red Control samples (C), in light blue EMB samples (M). It has shown the good separation of control samples on the bottom of the graphic as a function of PC1, supported by a biological variability (PC2).

The expression profiles of EMB samples were compared with those of control samples to detect expression changes in the 2 tested groups. DESeq2 analysis allowed the identification of 2,442 differentially expressed (DE) genes in the EMBs group compared with controls, which were associated with significant adjusted *p* values (FDR) ≤ 0.05. Furthermore, among the 2,442 DE genes, ~ 87% (2,136 genes) belonged to protein-coding biotype, nonetheless, also non-coding RNAs and regulatory RNAs were detected. In detail, 137 genes coding for small RNA (snoRNA, snRNA, miRNA and misc RNA), 128 genes coding for long non-coding RNA (lncRNA), 19 genes were tRNA, 10 genes had a function of guide RNA, 6 genes were related to ribosomal RNA (rRNA), 4 genes were transcribed as pseudogenes, one gene coding for SRP RNA (a part of ribonucleoprotein complex that recognizes the signal peptide and binds to the

ribosome, blocking protein synthesis), and one gene was a ribonuclease (RNase P RNA), as described in figure 15.

Figure 15

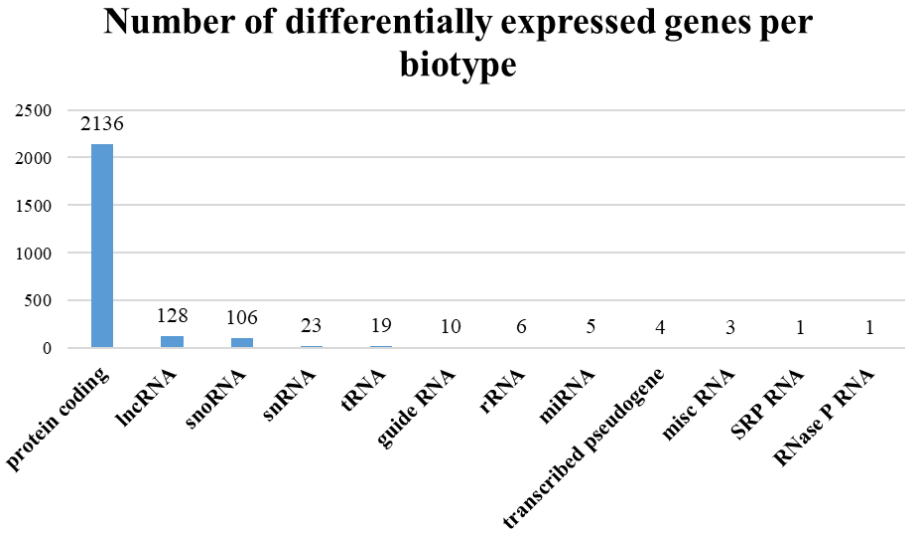


Figure 15: histogram with the number of differentially expressed genes divided per biotype.

In figure 16, differentially expressed genes are represented as a volcano plot. In order to deeper investigate differential gene expression profiles, functional Gene Ontology analysis with ShinyGO has been performed.

Figure 16

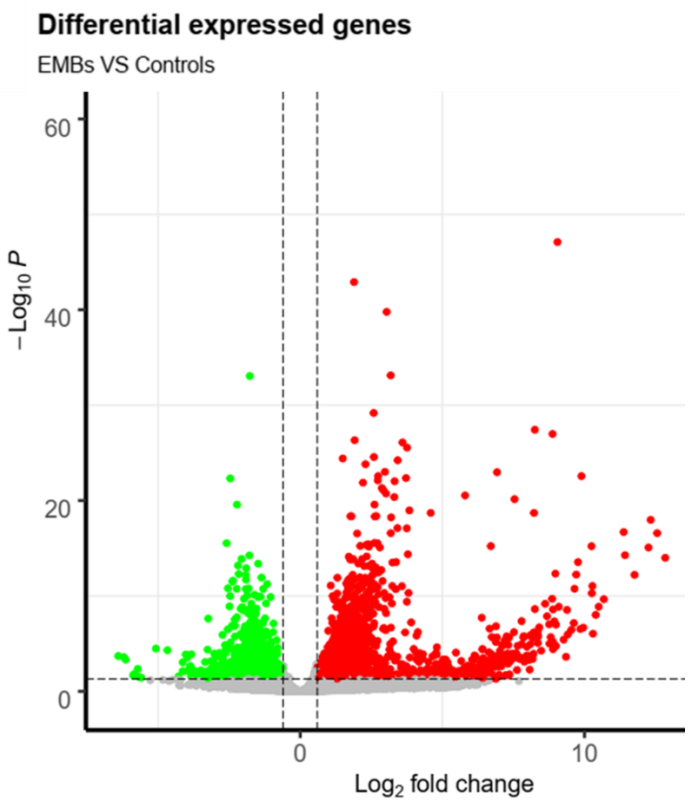


Figure 16: Volcano plot represented differentially expressed genes, up-regulated genes with a fold change ≥1.5 are shown in red and down regulated genes with fold change ≤-1.5 in green. Genes associated with not significant expression values, according to adjusted p values (FDR) threshold of 0.05, are reported in grey. On the x-axis and y-axis are reported Log2 fold change values and -Log10 p-values respectively.

The analysis for the identification of ontological processes enriched in all differentially expressed genes, showed the presence of several genetic pathways associated with significant FDR values ≤0.01, as summarized in figure 17.

Figure 17

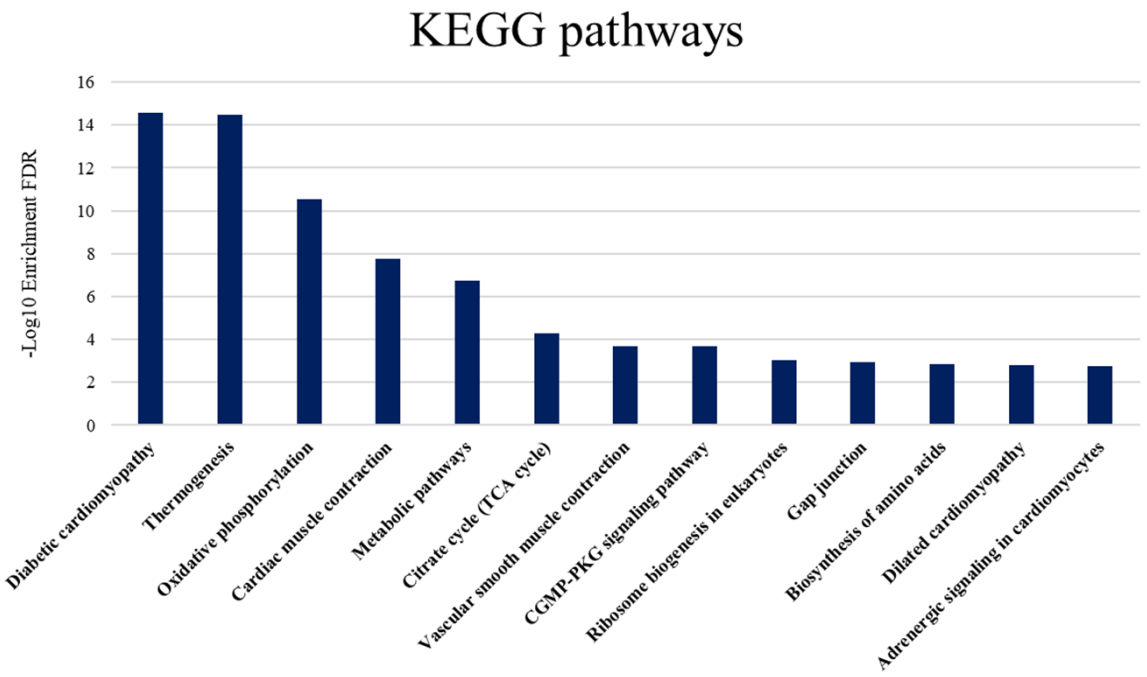
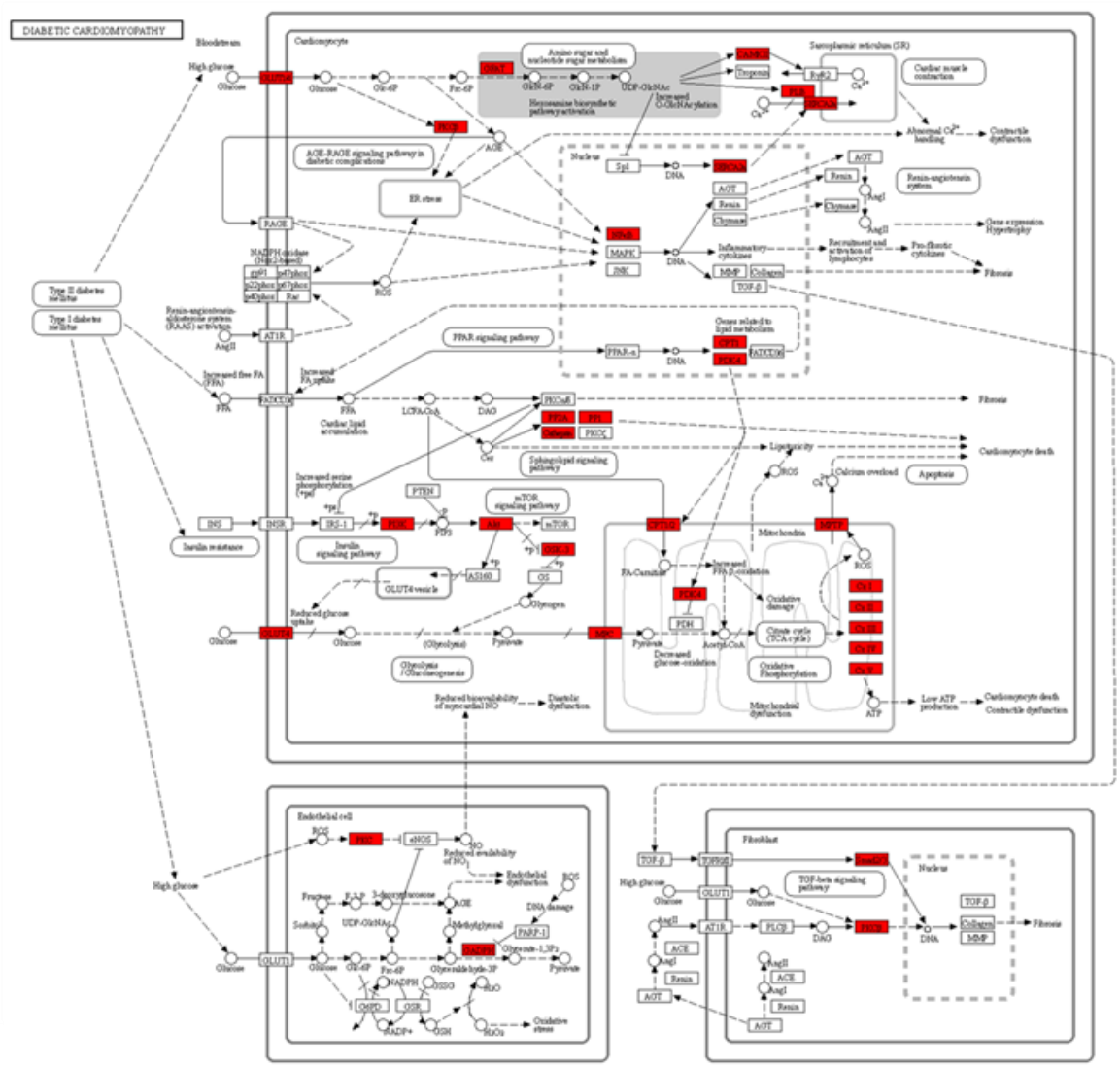


Figure 17: KEGG pathway detected with ShinyGO online software. On x-axis are reported KEGG pathways observed, on y-axis the values of enrichment, expressed as –log10 of Enriched FDR.

On the 2,442 total differentially expressed genes, the Kyoto Encyclopaedia of Genes and Genomes (KEGG) detected statistically significant correlations with distinct pathways involved in cardiac activity and heart disorders, as well as in tissue and cellular metabolism, such as "oxidative phosphorylation" pathway (associated with an enrichment FDR values=$2.95E^{-11}$ and 40 genes) and "citrate cycle (TCA cycle)" (enrichment FDR=$5.37E^{-05}$ and 12 genes). Focusing on GO terms related to heart functional processes or diseases, several genes are involved in cardiac muscle (enrichment FDR=$1.74E^{-08}$ and 28 genes) and vascular smooth muscle contraction (28 genes and enrichment FDR=0.0002), different forms of cardiomyopathies, as well as in pathway of signaling in cardiomyocytes, such as "CGMP-PKG signaling pathway" (that involved 32 genes with an enrichment FDR= 0.0002) and "adrenergic signaling in cardiomyocytes" GO term (enrichment FDR=0.002 and 28 genes); as described in figure 17. In the context of heart inflammation, GO results highlighted a good relationship between DE genes and 2 forms of myocardial disorders: diabetic and dilated cardiomyopathy, which are associated with significant enrichment FDR values of $2.58E^{-15}$ and 0.002 respectively. Interestingly, among DE genes, 61 and 21 were involved in the diabetic cardiomyopathy and the dilated cardiomyopathy pathways respectively as described in figure 18 (A and B), table 5 and table 6.

Figure 18

A


DIABETIC CARDIOMYOPATHY
Bloodstream
Cardiomyocyte
Sarcoplasmic reticulum (SR)
Endothelial cell
Fibroblast
Nucleus
Type II diabetes mellitus
Type I diabetes mellitus
High glucose
Glucose
GLUT1/4
HK
PFK
AGE-RAGE signaling pathway in diabetic complications
ER stress
RAGE
NADPH oxidase
ROS
AT1R
AngII
Increased free FA (FFA)
FABP/CD36
FFA
LCFA-CoA
DAG
Cardiac lipid accumulation
PPAR signaling pathway
PPAR-α
DNA
Genes related to lipid metabolism
CPT1
FABP/CD36
PDK4
PKCα/β
PKCδ
PKCε
PKCζ
Fibrosis
Lipotoxicity
Cardiomyocyte death
Sphingolipid signaling pathway
Increased serine phosphorylation (*pS/T)
PTEN
mTOR signaling pathway
INS
INSR
IRS-1
PI3K
PIP3
Akt
mTOR
Insulin resistance
Insulin signaling pathway
PDK-1
AS160
GLUT4 vesicle
GS
Glycogen
Reduced glucose uptake
Glycolysis (Glycolysis)
Pyruvate
Glycolysis / Gluconeogenesis
Reduced bioavailability of myocardial NO
Diastolic dysfunction
CPT1/2
Mitochondria
MPTP
FA-Carnitine
Increased FFA β-oxidation
PDK4
PDH
Acetyl-CoA
Citrate cycle (TCA cycle)
Decreased glucose oxidation
Oxidative Phosphorylation
Oxidative damage
Mitochondrial dysfunction
ATP
Low ATP production
Contractile dysfunction
Calcium overload
Apoptosis
Cx I
Cx II
Cx III
Cx IV
Cx V
Sarcoplasmic reticulum (SR)
Cardiac muscle contraction
Abnormal Ca2+ handling
Contractile dysfunction
RyR2
PLB
SERCA2a
Amino sugar and nucleotide sugar metabolism
GFAT
GlcN-6P
GlcN-1P
UDP-GlcNAc
Hexosamine biosynthetic pathway activation
Increased O-GlcNAcylation
Nucleus
Sp1
DNA
NFκB
MAPK
JNK
AOT
Renin
Chymase
AngI
ACE
Resin-angiotensin system
Gene expression Hypertrophy
Inflammatory cytokines
Recruitment and activation of lymphocytes
Pro-fibrotic cytokines
MMP
Collagens
TGF-β
Endothelial cell
RCS
PKC
eNOS
NO
Reduced availability of NO
Endothelial dysfunction
ROS
DNA damage
Fructose
F-6-P
3-deoxyglucosone
AGE
Sorbitol
UDP-GlcNAc
Methylglyoxal
PARP-1
GAPDH
Glyceraldehyde-3P
Glycerate-1,3P2
Pyruvate
NADPH
GSSG
H2O
GSH
NADP+
H2O2
Oxidative stress
Fibroblast
TGF-β
TGFβR
Smad2/3
TGF-beta signaling pathway
PLCβ
DAG
PKCδ
Fibrosis

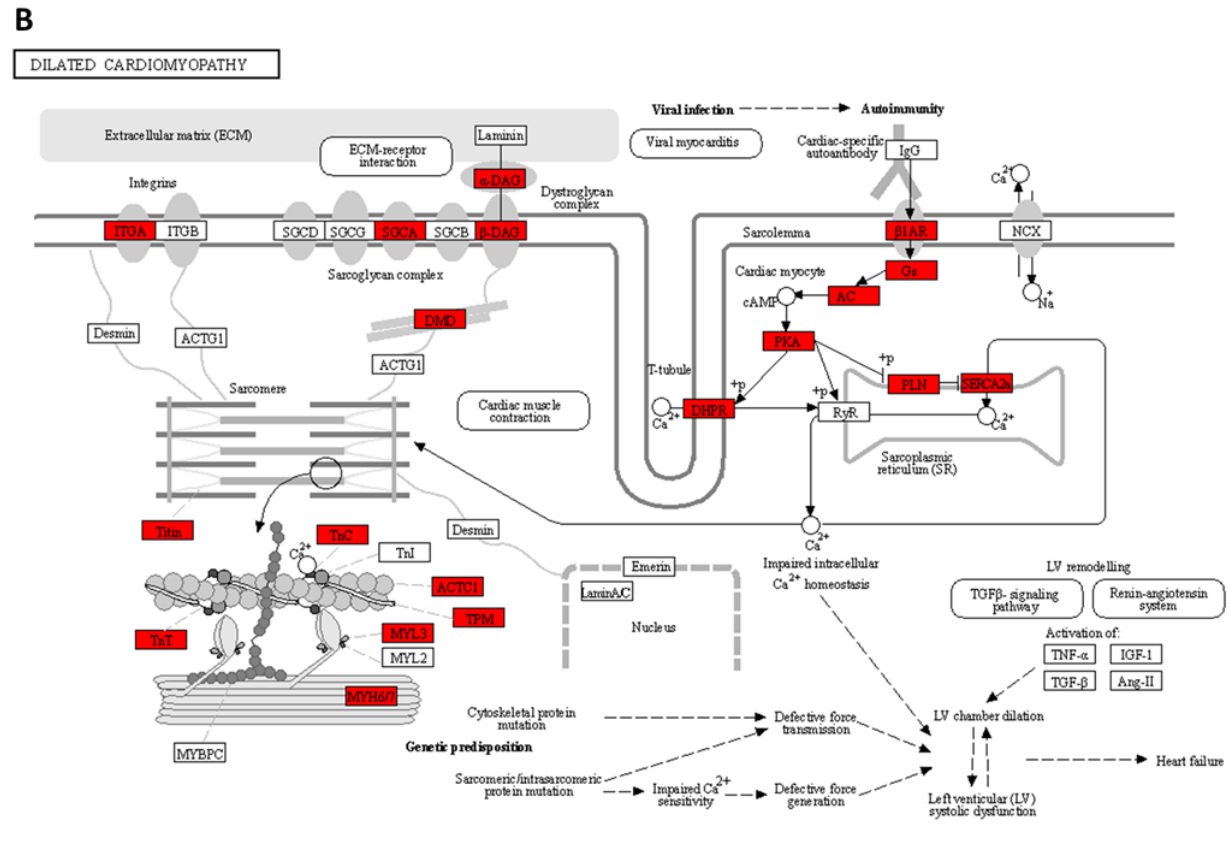


Figure 18: KEGG graph described diabetic (A) and dilated (B) cardiomyopathy pathways and the involved differentially expressed genes highlighted in red.

Table 5

Diabetic Cardiomyopathy genes	Pvalue	FDR	Diabetic Cardiomyopathy genes	pvalue	FDR
SMAD2	7.05E-07	2.07E-05	COX8A	0.0001	0.002
GSK3B	0.003	0.022	ATP2A2	4.81E-08	1.83E-06
PDK4	5.75E-08	2.15E-06	RELA	0.0001	0.002
PRKCB	1.32E-06	3.61E-05	COX6A2	2.97E-08	1.20E-06
GFPT1	0.002	0.01	VDAC1	0.0007	0.008
PIK3R1	0.001	0.014	ATP5F1A	3.94E-05	0.0007
AKT3	4.63E-05	0.0007	CYC1	1.28E-06	3.53E-05
GAPDH	1.60E-14	2.48E-12	NDUFV1	0.002	0.016
ATP5F1C	7.04E-13	7.40E-11	PDK2	2.34E-07	7.60E-06
ATP5F1D	2.21E-06	5.61E-05	NDUFS3	0.001	0.011
LOC100684983	1.66E-05	0.001	CPT1B	4.83E-07	1.46E-05
LOC609990	5.29E-05	0.0008685	NDUFS2	0.0046	0.0346
SLC25A6	0.001	0.009	SDHA	0.0007	0.008
UQCRC1	8.90E-07	2.54E-05	NDUFA11	1.55E-05	0.0003
ATP5F1B	5.61E-17	1.33E-14	PLN	1.37E-13	1.65E-11
CPT2	0.003	0.022	ATP5PO	6.33E-06	0.0001
PTPA	0.001	0.004	COX5A	5.99E-06	0.0001
SLC25A4	0.003	0.028	NDUFA9	7.78E-05	0.001
COX5B	3.09E-05	0.001	ATP5MC1	4.04E-10	2.38E-08
NDUFB11	4.01E-05	0.0007	UQCR11	0.0002	0.0031
SDHC	9.71E-19	2.95E-16	NDUFA2	1.35E-06	3.68E-05
NDUFS8	1.09E-06	3.03E-05	SLC2A4	4.55E-08	1.74E-06
NDUFV3	0.004	0.031	SDHB	1.01E-05	0.0002
UQCRH	0.0008	0.008	CAMK2D	0.0005	0.005
COX7A2L	0.006	0.041	NDUFB8	0.00011	0.0017
NDUFA4	0.0001	0.002	MPC1	0.003	0.026
NDUFC2	1.96E-06	5.13E-05	LOC612644	0.005	0.038
PPP1CA	1.31E-05	0.0003	ATP5MC2	1.59E-06	4.25E-05
NDUFA7	0.0001	0.0816	CTSD	0.0004	0.005
UQCRC2	4.68E-05	0.0007	NDUFS7	0.002	0.014
UQCRFS1	0.002	0.0197			

Table 5: the differentially expressed genes related to diabetic cardiomyopathy pathway and reported p-value < 0.01 and p-value adjusted (FDR) <0.05.

Table 6

Dilated cardiomyopathy related genes	pvalue	padj
CACNA2D3	0.004	0.0295
CACNA1S	0.0006	0.007
MYH7	1.68E-07	5.69E-06
TPM2	0.002	0.0204
TPM4	5.80E-05	0.0009
TNNT2	2.75E-10	1.71E-08
TTN	3.69E-33	6.59E-30
DMD	2.85E-14	4.17E-12
ITGA9	0.0023	0.0199
ACTC1	8.14E-09	3.73E-07
MYL3	0.003	0.025
PRKACB	0.005	0.038
TNNC1	0.006	0.044
ADCY4	0.0009	0.001
DAG1	0.002	0.02
GNAS	7.31E-15	1.20E-12
ATP2A2	4.81E-08	1.83E-06
PLN	1.37E-13	1.65E-11
ADCY5	5.96E-05	0.000956438
ADRB1	0.005	0.007
SGCA	8.74E-05	0.001

Table 6: the differentially expressed genes related to dilated cardiomyopathy pathway and reported p-value < 0.01 and p-value adjusted (FDR) <0.05.

Furthermore, the investigation was focused on the 3 cases diagnosed with a mild form of myocarditis, that showed lymphocytic interstitial infiltrates with inflammation signs. DEA was performed comparing transcriptome profiles of the 3 EMB samples with those of controls. The analysis revealed 6,474 differentially expressed genes associated with significant *p* value adjusted $\leq$ 0.05. Most of them, ~88%, belonged to "protein coding" biotype category (n=5,687), only 537 genes (~8%) were long non-coding RNA. The remaining genes were transcribed as snoRNA (n=123), snRNA (n=37), tRNA (n=28), miRNA (n=20), misc RNA (n=14), guide RNA (n=10), pseudogene (n=7), rRNA (n=7), C region (n=1), V segment (n=1), SRP RNA (n=1) and RNase P RNA (n=1) as described in figure 19. In figure 20, the total differentially expressed genes were reported as volcano plot.

Figure 19

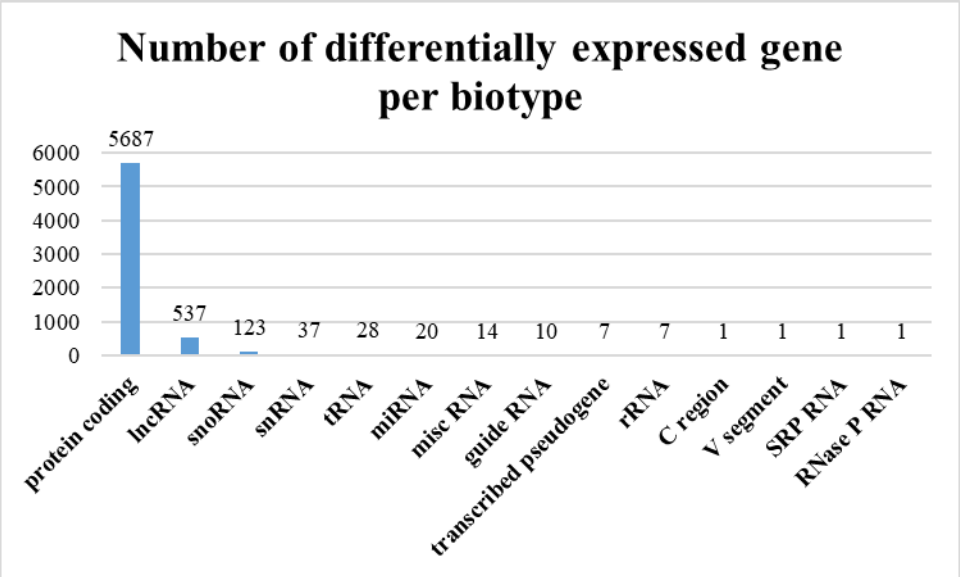


Figure 19: histogram with the number of differentially expressed genes per biotype.

Figure 20

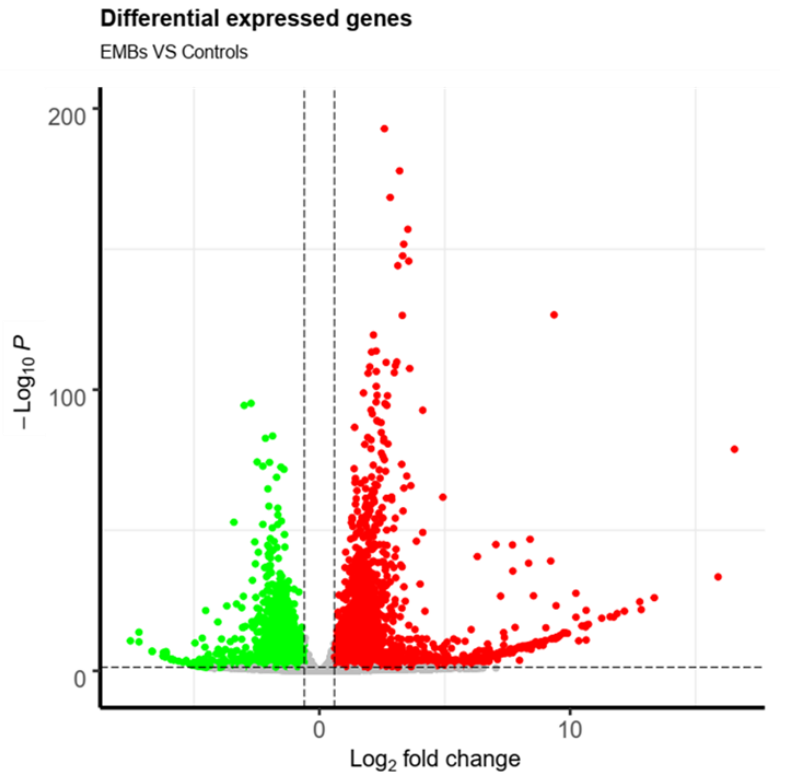


Figure 20: Volcano plot represented differentially expressed genes, up-regulated genes with a fold change ≥1.5 are shown in red and down regulated genes with fold change ≤-1.5 in green. Genes associated with not significant expression values, according to adjusted p values (FDR) threshold of 0.05, are reported in grey. On the x-axis and y-axis are reported Log2 fold change values and -Log10 p-values respectively.

In order to evaluate the functional pathways related to the 6,474 differentially expressed genes that characterized the 3 mild myocarditis samples, GO analysis was performed with ShinyGO software, resulting in significant enriched gene ontology terms associated with strong statistically relevant FDR values ≤0.01, as summarized in figure 21. This analysis revealed distinct pathways involved in cardiac activity and heart disorders, as well as in structural cell organization and metabolism such as "metabolic pathways" (436 genes) and "oxidative phosphorylation" (48 genes) associated with $5.74E^{-09}$ and 0.002 enrichment FDR value respectively. In particular, DE genes involved in reorganization of cell in cardiac tissue, highlighted pathways concerning "adherens junction" (28 genes, enriched FDR=0.008), "focal adhesion" (81 genes, enriched FDR=$2.37E^{-06}$), gap and tight junction (39 genes, enriched FDR=0.0002 and 57 genes, enriched FDR=0.004 respectively), "regulation of actin cytoskeleton" (73 genes and enriched FDR=0.003), "cell adhesion molecules" (46 genes and enriched FDR=0.009), "ECM-receptor interaction" (34 genes and enrichment FDR=0.004) and "platlet activation" (47 genes and enrichment FDR=0.001). Furthermore, the analysis revealed GO terms related to host responses and pathogen interaction during infections, in particular "leukocyte transendothelial migration" associated with 41 DE genes and enrichment FDR values of 0.005, "bacterial invasion of epithelial cells" (31 DE genes and enrichment FDR=0.008) and "B cell receptor signaling pathway", characterized by 27 significant genes and enriched FDR values of 0.008. Focusing on GO terms associated with heart related physiological processes, the analysis resulted in 35 genes involved in "cardiac muscle contraction" with an enrichment FDR value of 0.002, 56 genes implicated in "adrenergic signaling in cardiomyocytes" (enrichment FDR=0.0007), 60 and 75 genes for "CGMP-PKG signaling pathway" and "calcium signaling pathway" respectively (enrichment FDR=0.0005 and 0.009 respectively). Concerninng heart related diseases, 21 genes were linked to "viral myocarditis" and 82 genes to "diabetic cardiomyopathy" (enrichment FDR=0.006 and $1.62E^{-06}$ respectively).

Figure 21

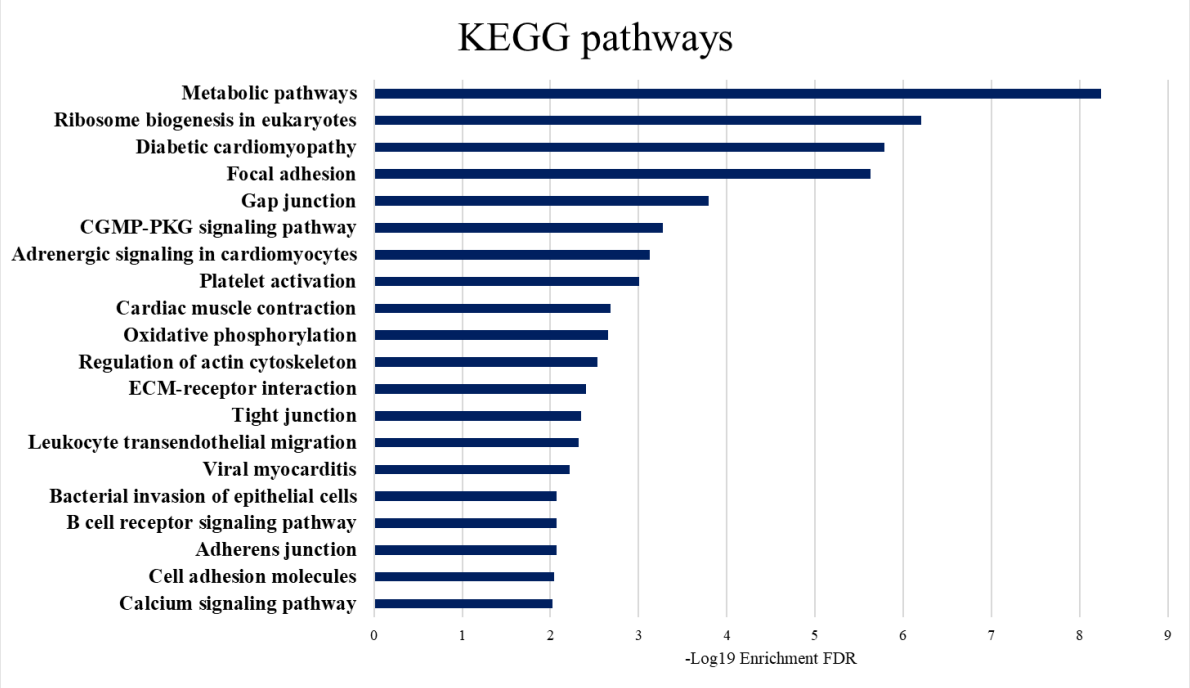


Figure 21: KEGG pathway detected with ShinyGO online software. On x-axis are reported KEGG pathways observed, on y-axis the values of enrichment, expressed as –log10 of Enriched FDR.

4.3 DISCUSSION

In this study endomyocardial biopsy technique used appeared feasible and did not show any complication episodes representing the gold standard to understand the etiological origin of myocardial disorders. A deep characterization of myocardial disorders was carried out by the use of NGS technology, which allowed the assay of the transcriptomic landscape of dogs affected with heart tissue inflammation. Furthermore, to explore the possible relationship between myocardial disorders and pathogens, a new software pipeline was developed and used as an advanced tool able to perform taxonomy classification of tested samples. Although the PCR assay for pathogen discovery showed the presence of canine distemper virus and parvovirus in 3 samples, these viral entities were not detected by metagenomic NGS experiment, probably due to intrinsic limitation of PCR-based methods. PCR assays, with specific pathogens panels, in fact, could provided an oriented selection of infective agents, that may exclude other possible agents of myocardial infection, and because of amplification step, this technique could reveal misleading results.

Generally, despite the large amount of data produced by NGS technology, oriented to pathogen discovery related to disease, there is a lack of a comprehensive and easy-use pipeline for data

analysis that avoids complicated bioinformatics steps and provides extensive information. With this aim, a new software, HOME-BIO, was developed. Testing results confirmed the good performance of the tool in all test datasets analyzed, leading, as expected, to the identification of several microorganisms independently from the nature and pre-treatment of the tissue samples. Thus, HOME-BIO emerged as a user-friendly pipeline based on a dockerized solution, designed for analyzing shotgun metagenomic data, avoiding time-consuming and error-prone installation and configuration steps. This modular pipeline provides a quality control step and the two main analysis approaches commonly used in metagenomic studies. By querying bacterial, viral, protozoal databases, HOME-BIO generates exhaustive taxonomic profiling of the biological entities in specimens. Consequently, this project focused on the exploration of the pathogen content of all 13 EMB samples considered in this study. Results showed that most of the reads related to microorganisms, belonged to bacteria kingdom (~12% of all high-quality reads analyzed), in agreement with enriched pathway "Bacterial invasion of epithelial cells" highlighted in differential expression analysis between the 3 samples diagnosed with mild form of myocarditis and control group. Interestingly, bacteria detected with HOME-BIO, are in agreement with the literature, especially for *Staphylococcus*, *Streptococcus* and *Corynebacterium* genus, which were described as cardiotropic pathogens in human-related with myocarditis and inflammatory cardiomyopathy [Calabrese et al., 2003]. In detail, in a recent study, Mahony and colleagues reported cases of Australian children affected with underlying cardiac disorders in which, *Staphylococcus aureus* predominated [Mahony et al., 2021]. Furthermore, *Staphylococcus* was also found in dogs. Janus and collaborators described the presence of *Staphylococcus aureus* within heart muscle in 1 dog with cardiac disorders, characterized by severe granulocytic infiltration in the left ventricle, occurred with mild diffuse lymphoplasmacytic infiltration and degeneration of cardiomyocyte [Janus et al., 2014]. Concerning the *Streptococcus* genus, it has been reported that myocarditis may develop after group A streptococcus (GAS) infection. According to HOME-BIO analysis, Aguirre and coll. (2015) reported some case series that could help clinicians to provide a diagnosis of GAS-induced myocarditis in young men with evidence of streptococcal tonsillitis or pharyngitis demonstrating the cardiotropism of this bacterial genus [Aguirre et al., 2015]. Furthermore, as described by Lamm an coll., streptococcal infection was associated with cardiac disorders in dogs. With their retrospective study, regarding 393 cases, they reported evidence of *Streptococcus* in animals affected with inflammation of the mitral valves [Lamm et al., 2010]. *Corynebacterium* genus, instead, represented by *C. diphtheriae*, due to bacterial exotoxin, could lead to heart inflammation and relative cardiomyopathy in humans. Nonetheless,

nowadays *C. diphtheria*-induced myocarditis is rare in Europe and America (except for Scandinavia cases in the mid-1980s) but remains a problem in developing countries [Calabrese et al., 2003].

Moreover, the parasites content of specimens was investigated. HOME-BIO was able to detect protozoa-related reads at low abundance levels. Although only a few percentage of high-quality reads was assigned to protozoal entities, 3 protozoal genera related to cardiac disorder were detected. Interestingly *Leishmania*, *Plasmodium* and *Toxoplasma* were described as infective agents with cardiotropism, which could lead to cardiac muscle inflammation and cardiomyopathy [Schultheiss et al., 2019]. Concerning the taxonomy profiling of EMB samples, no common pathogen associated with a high percentage of assigned RNA fragments was found. According to histological results, previously reported, the low abundance of pathogen detected with HOME-BIO is attributable to the possible clearance of the microorganisms.

Considering HOME-BIO findings, this study describes an evolved and chronic clinical picture that defines the progression to dilated cardiac phenotype. No presence of pathogens, associated with a high number of RNA fragments, was revealed by NGS evaluation on samples studied, however, a possible role of microorganisms to the genesis of histological and clinical presentation can not be excluded. In humans and experimental models, myocarditis, after the clearance of the infective agent, can recover the cardiac muscle in its normal functions but due to autoimmunity process promote a chronic scenario that may progress to DCM and end-stage heart failure [Caforio et al., 2013]. Dogs with DCM characterized the largest proportion of dogs in the present study, however since it has not been performed in the experimental condition, it is very difficult to correlate the tissue damage to a specific pathogen insult. Three animals selected in the present study showed the presence of inflammatory infiltrates, while the majority of dogs showed signs of fibrosis. With the exception of the presence of inflammatory infiltrate indicative of mild myocarditis, in the other samples the presence of fibrosis can not exclude to be a possible chronic evolution of a myocarditis process. Although is not directly evaluated, the presence of fibrosis, when detecetd in animals showing a clinical presentation of DCM seemed indicate a cronic evolution of a pathological process more than an idiopatic form. However, on the basis of these diagnostic elements, appear really difficult perform a correct diagnosis and discriminate a primary DCM from a chronic evolution of myocarditis.

In order to obtain more information about the clinical presentation of DCM and in particular DCM characterized by mild myocarditis, this research activity also focused on the transcriptomic landscape definition of EMBs of the 10 samples diagnosed with DCM. For transcriptome obtained from dogs diagnosed with ARVC and arrhythmias, no analytical assessment was made, since they showed heterogeneous pathological conditions. The data obtained, showed a clear alteration of the expression profiles of several genes and entire metabolic pathways in EMB samples, compared with normal heart tissues. In fact, a total of 2,442 differentially expressed genes were identified, most of them resulted as protein-coding and regulatory genes. The Gene Ontology (GO) analysis, performed on them, revealed significant enrichment in cardiac activity and heart disorders, as well as in tissue and cellular metabolism, demonstrating clear differences between sequenced samples and the control group. These results are strongly harmonized with both the diagnosis of DCM and the mild myocarditis recorded that characterized the dogs enrolled in the present study.

In canine primary DCM, different genes and their mutation have been studied, because related to the synthesis of proteins involved on pathologic process: dystrophin (*DMD*) in German short-haired pointers [Schatzberg et al., 1999], pyruvate dehydrogenase kinase isozyme 4 (*PDK4*) in Doberman Pinschers [Meurs et al., 2012], and striatin (*STRN*) in Boxers [Meurs et al., 2013], in addition to a locus on chromosome 5 in Doberman Pinschers [Mausberg et al., 2011]. Additional polymorphisms on chromosomes 1, 10, 15, 17, 21, and 37 have also been implicated in Irish Wolfhounds [Philipp et al., 2012]. These studies were focused on structural variations, mutations and polymorphisms of genes, obtained with the analysis of DNA sequences. Our data did not show any correlation with these genes, probably due to a different experimental design, indeed our study condition was based on RNA-seq data. This result is not easy to explain and compare between each other, and it could be considered that these genes are related to familiary forms of DCM, more than to transcriptomic responses to inflammatory process. Recently, RNA-seq analisis has been conducted 8 dogs naturally affected by DCM [Friedenberg et al., 2016]. The authors identified 86 protein-coding genes differentially regulated in dogs with DCM when compared with unaffected control dogs; of these genes, 52 and 34 were downregulated and upregulated respectivelly. Broadly, our findings were similar to Friedenberg and coll., demostrating the involment of the genes encoding for structural and functional cardiac proteins and a severe dysregulation of cellular energy metabolism in affected dogs. In detail, our results were in agreement with Friedenberg and coll. (2016), showing 17 overlapping significant protein-coding genes between the two studies (*ACSL1*, *APOE*,

ATP13A3, *DDAH1*, *DNAJB1*, *GNAO1*, *GNB3*, *HIPK3*, *KCNIP2*, *MMP11*, *OAT*, *PDCD4*, *PDE3A*, *PIK3R1*, *PPP1R13L*, *ROS1*, *RTF1*). These genes, more than an involvement on the sintesis of myocardial structural proteins seams to be related to ion trafficking and energy metabolism. In particular, acyl-CoA synthetase long-chain family member 1 (*ACSL1*); ATPase, Na+/K+ transporting, α 3 polypeptide (*ATP1A3*), and phosophoinositide-3-kinase regulatory subunit 1 (*PIK3R1*) were genes involved in energetic metabolism. Interestingly, *ACSL1* is one of the most important genes implicated in energy production. Encoding the protein long-chain acyl-CoA synthetase 1, *ACSL1* has a key function by promoting the first step in the intracellular metabolism of fatty acids and consequent energetic balance [Phillips et al., 2010]. *ATP1A3*, instead, is a member of P-type cation transport ATPases family, which is an integral membrane protein that manteins the correct electrochemical gradients of ions across the plasma membrane. In cardiac tissue, the role of *ATP1A3* is fundamental for the physiological muscle contractility and promotes glucose uptake into cells. Dysfunctions associated with dysregulation or mutation of this gene are also implicated in diabetic cardiomyopathy in experimental model [Belliard et al., 2013]. In this context, *PIK3R1* is another differentially expressed gene identified in both our investigation and Friedenberg's study. Its function regards glucose metabolism, in particular, it is involved in insulin signaling and the production of phosphatidylinositol 3,4,5-trisphosphate with activation of serine-threonine kinases. Furthermore, *PIK3R1* dysfunctions are associated with marked insulin resistence. Taken together, these findings, have been highlighted a severe cellular energy dysregulation in dogs with DCM and suggested that the origin of DCM could derive from an underlying imperfection in the regulation of energetic metabolism. This hypothesis is also supported by several research studies performed both on humans and dogs [Ashrafian et al., 2010, Sibbing et al., 2011, Zarrouk et al., 2012 and Meurs et al., 2012]. In addition, the diabetic cardiomyopathy pathway was detected in Gene Ontology analysis, it is a left ventricular dysfunction that generally arises among patients with diabetes. This pathology implies increased circulating fatty acids and hyperglycemia, along with insulin resistance, that promotes the generation of reactive oxygen species (ROS), mitochondrial dysfunction and impaired calcium metabolism. Together these processes trigger the cardiac fibrosis, hypertrophy, the cardiomyocyte death, with consequent contractile dysfunction and endothelial cell damage. The insulin resistance plays a key role in the development of type 2 diabetes, which in turn is a risk factor for several cardiovascular disease and may lead to heart muscle dysfunction. Although diabet is a disease related to the age, Mak and collegues, in their work, reviewed the correlations between this pathology and pediatric patients, according to young

dogs enrolled in this experiment. Furthermore, it has been reported that insulin resistance is also present in individuals with cardiomyopathy without diabetes, which suggests a potential bi-directional link between cardiomyopathy and insulin resistance as described by Mak and collaborators [Mak et al., 2021]. These results, raise the hypothesis that it might be interesting to evaluate, in the future, whether DCM affected animals may show insulin resistance. In addition, in canine species, changes in myocardial metabolism associated with heart disorders and failures have been described. In this context, the development of insulin resistence in myocardial tissue is part of the structural, functional, and metabolic remodeling in dogs affected with heart problems. Significant changes in the availability and production of the energy supply in the myocardium were noted in the later stages of myocardial disease, as described by Arkadievich (2021) [Arkadievich 2021]. In our study conditions, other genes involved in energetic metabolism and ion trafficking showed a differential expression between DCM and control dogs. Our results showed dysregulation of genes responsible for ion trafficking related to the contraction of cardiac muscle, a fundamental process compromised in DCM. In cardiomyocytes, in fact, the action potential (AP) induces cardiac muscle contractions that play a fundamental role in the pathogenesis of the disease. It is evident that cardiac ion channels affect the cardiac AP, in particular, Ca^{2+} ions, by directly binding to proteins of the sarcomere (the functional unit in the contraction step) and are responsible for the initiation of the myocyte contraction [Molina-Navarro et al., 2013]. Furthermore, in human DCM has also been described with arrhythmias, with mutations in genes encoding sodium and potassium channels. Changes in expression levels of genes related to ion channels, such as *CACNA2D3* (Calcium Voltage-Gated Channel Auxiliary Subunit Alpha2delta 3), *CACNA1S* (Calcium Voltage-Gated Channel Subunit Alpha1 S) may lead to defects of functional Ca^{2+} currents and consequently may trigger damages of the cardiac contraction process typical in DCM disease [Zhang et al., 2013]. In addition, the downregulated gene *ATP2A2* (ATPase Sarcoplasmic/Endoplasmic Reticulum Ca^{2+} Transporting 2), identified in this study, is an ATPase enzyme involved in DCM, particularly in the regulation of the level of Ca^{2+} during muscle contraction and relaxation. It has been reported that downregulation of *ATP2A2* also correlated with heart failure [Alimadadi et al., 2020]. *ATP2A2* was also involved in the "CGMP-PKG signaling pathway" that play a fundamental role in the maintenance of the heart physiological state [Gorbe et al., 2010].

Most of the differentially expressed genes associated with metabolic pathway and oxidative phosphorylation pathway, are related to energy metabolism. Examples of a dysregulated gene

involved in energy metabolism are represented by Phosphatidylinositol-4-Phosphate 3-Kinase Catalytic Subunit Type 2 Beta (*PIK3C2B*) that belong to PI3Ks family and genes that encode for components of an ATP synthase complex (*ATP5F1A*, *ATP5F1B*, *ATP5F1C* and *ATP5F1D*). As suggested by Friendeberg and collaborators (2016), genes related to the PI3Ks family and ATP synthase complex play a key role in energy balance in cardiomyocytes and dysregulation of these genes is strongly related to the pathogenesis of DCM in dogs [Friedenberg et al., 2016]. In this context, it is proposed that the origin of the processes that lead to DCM, could be determined by an underlying deficit in the energy metabolism of dogs affected with cardiac disorders [Meurs et al., 2012].

Regarding the genes involved in the sintesis of functional and structural proteins, our results did not show any correlation with other RNAseq studies in dog. However, most of them are associated to DCM in studies conducted in Human and canine DNA mutations. The genes differentially expressed in our investigation, are strongly related to DCM including Myosin Heavy Chain 7 (*MYH7)*, Troponin T2, Cardiac Type *(TNNT2*), Titin *(TTN*), Phospholamban *(PLN*), Actin Alpha Cardiac Muscle 1 (*ACTC1*) [Wiersma et al., 2007; Tayal et al., 2017, Kayvanpour et al., 2017, McNally et al., 2017]. All these genes are included in the 14 canine candidates selected by Wiersma and colleagues (2007), which are related to sarcomeric proteins synthesis and for this reason, are involved in DCM and the cardiac contraction muscle pathway. Other genes related to DCM in dogs and humans are reported in table 7. It is difficult, however, whether the dysregulation of these genes represents a primary or secondary effect in the pathogenesis of DCM. In our study condition, the results let us suppose that an important impulse in the development of DCM could be related to altered energetic metabolism more than structural damages. Although, it is also possible that the architecture of the cardiacmyocytes in dogs with DCM has been altered, leading to secondary dysfunction of energy regulation.

Table 7

Gene	*Associated with DCM in humans*	*Associated with canine DCM*	*DCM dogs (Breed)*	*Control dogs (Breed)*
ACTC1	Y	N	Doberman Pinscher, Irish Wolfhound, Newfoundland	mixed breeds, Irish Wolfhounds, Newfoundlands, unspecified dogs without overt heart disease
ACTN2	Y	N	Doberman Pinscher	unspecified dogs without overt heart disease
CAV1	N	N	Newfoundland	Newfoundlands
CSRP3	Y	N	Irish Wolfhound, Newfoundland, Doberman Pinscher	Irish Wolfhounds, Newfoundlands, Labradors, unspecified dogs without overt heart disease
DES	Y	N	Doberman Pinscher, Irish Wolfhound, Newfoundland	Doberman Pinscher, Irish Wolfhound, Newfoundland
DMD	Y	Y	German short-haired pointers	German short-haired pointers with reduced dystrophin
LDB3	Y	N	Newfoundland	Newfoundlands
LMNA	Y	N	Newfoundland, Doberman Pinscher	Newfoundland, Labradors
MYBPC3	Y	N	Doberman Pinscher	unspecified dogs without overt heart disease
MYH7	Y	N	Newfoundland, Doberman Pinscher	Newfoundlands, Labradors, unspecified dogs without overt heart disease
PDK4	N	Y	Doberman Pinscher	Doberman Pinschers + others from 11 breeds
PLN	Y	N	Doberman Pinscher, Newfoundland, great dane, Irish Wolfhound	Doberman Pinschers, Irish Wolfhounds, Newfoundlands
SGCD	Y	N	Doberman Pinscher, Irish Wolfhound, Newfoundland	Doberman Pinscher, Irish Wolfhound, Newfoundland
STRN	N	Y	Boxers	Boxers
TAZ	Y	N	Irish Wolfhound	Irish Wolfhounds
TCAP	Y	N	Doberman Pinscher, Newfoundland, Irish Wolfhound	Newfoundland, Irish Wolfhound, unspecified dogs without overt heart disease
TMOD	N	N	Irish Wolfhound	Irish Wolfhounds
TNNC1	Y	N	Doberman Pinscher	Labradors
TNNI3	Y	N	Newfoundlands	Newfoundlands
TNNT2	Y	N	Doberman Pinscher, Newfoundland	Labradors, Newfoundlands, dogs without overt heart disease
TPM1	Y	N	Newfoundland, Doberman Pinscher	Newfoundlands, dogs without overt heart disease
TTN	Y	N	Newfoundland, Doberman Pinscher	Newfoundlands, mixed breeds
VCL	Y	N	Doberman Pinscher, Newfoundland	Newfoundlands, dogs without overt heart disease

Table 7: genes investigated by Simpson and collaborators (2015) in relation with canine DCM [Simpson et al., 2015]. Y= yes, N= no.

Considering that, GO analisis confirmed the involvement of gene related to different pathological disorders (DCM, diabetic cardiomyopathy and metabolic disorders), the presence of inflammation infiltrate, indicative of mild myocarditis, in three of DCM dogs, drove us to perform DEA versus control group. Interestingly, DEA performed on these 3 samples highlighted ontology terms that were absent in the analysis involving all 10 cases. In detail, differentially expressed genes were associated to pathways that defined mild myocarditis condition and were related to structural cell organization, host responses to infection and pathogens' activity, raising the hypothesis that for these 3 dogs, the inflammatory process was more recent and the physiological recover of heart muscle did not completely occur. The inflammation of cardiac tissue, in fact, may lead to cells and tissues remodeling with consequent fibrosis [Molesan et al., 2019]. In addition, the leukocyte transendothelial migration from the blood into tissues is a key point for immune surveillance and inflammation that occur in myocarditis. The molecular bases of leukocytes migration (also known as leukocyte adhesion cascade or diapedesis) involved their binding to cell adhesion molecules of vascular endothelium. A leukocyte adherent to CAMs on the endothelial cells moves forward by leading-edge protrusion and retraction of its tail. This process is mediated by integrins which activate RhoA and the kinase p160ROCK. This event leads to the phosphorylation of myosin with consequent retraction of the actin cytoskeleton. Furthermore, leukocytes themselves, through endothelial cell signals, may stimulate cell retraction during the separation of the local cell junctions present in endothelium [van Buul et al., 2004]. To the best of our knowledge, this study performed for the first time RNA-seq analysis on heart tissues *in vivo* obtained from dogs affected with myocarditis. A similar approach was used by Takeuchi and coll. (2018) in human patients, nonetheless, this study did non explore the transcriptomic responses.

This study presents some limitations: it involved only dogs affected with myocardial disorders and no healthy control animals were enrolled; furthermore, the study cohort was characterized by different ages and canine breeds. In addition, the sequencing was performed with an average of 30 million of reads per sample, that could be considered a technical limitation for metagenomic/metatranscriptomic experiments, because pathogens at very low abundance could be not revealed from the analysis with HOME-BIO pipeline.

4.4 CONCLUSION

Thanks to the use of the minimally invasive diagnostic tool of endomyocardial biopsies, combined with the innovative NGS technology, the processes regarding cardiac tissue inflammation were deeply investigated. First, the experimental design and analysis were focused on the host transcriptomic responses to the inflammation. This approach allowed to highlight molecular and genetic events that could play an important role in the progression of myocarditis disease that, starting to heart tissue inflammation, may lead to DCM. On the other hand, the taxonomic characterization of endomyocardial biopsies, as a function of various pathological factors, represented the second step of this study. Experiments with NGS platforms have generated a huge amount of information that was a challenge for the management and data analysis. Thus, in order to investigate pathogens' content of samples, a new bioinformatic software (HOME-BIO) was developed. HOME-BIO results in a user-friendly pipeline designed for taxonomy profiling and metagenomic data. Pathogen characterization of endomyocardial biopsies revealed microorganisms related to cardiac disorders. Together these findings, provide new biological insight into myocarditis disease and represented a novel approach in the study of canine heart tissue inflammation. Moreover, when DCM when considered, our findings confirmed that altered cellular energy metabolism and gene of structural-functional cardiac protein are involved in its pathogenesis suggesting important pathways for future research. In conclusion, the present thesis represents a primary application of the NGS strategy on the study of cardiac tissues obtained in vivo. This "omic" technology showed a new horizon, improving the knowledge on the pathogenesis of myocarditis in the dog, contributing to developing a novel approach for the diagnosis and therapy of this pathology.

Aguirre JL, Jurado M, Porres-Aguilar M, Olivas-Chacon C, Porres-Muñoz M, Mukherjee D, Taveras J. Acute nonrheumatic streptococcal myocarditis resembling ST-elevation acute myocardial infarction in a young patient. Proc (Bayl Univ Med Cent). 2015 Apr;28(2):188-90. doi: 10.1080/08998280.2015.11929224. PMID: 25829649; PMCID: PMC4365115.

Alimadadi A, Munroe PB, Joe B, Cheng X. Meta-Analysis of Dilated Cardiomyopathy Using Cardiac RNA-Seq Transcriptomic Datasets. Genes (Basel). 2020 Jan 4;11(1):60. doi: 10.3390/genes11010060. PMID: 31948008; PMCID: PMC7017089.

Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data. 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Arkadievich OD. Metabolic markers of myocardium insulin resistance in dogs with heart failure. Open Vet J. 2021 Jan;10(4):363-370. doi: 10.4314/ovj.v10i4.2. Epub 2020 Oct 20. PMID: 33614430; PMCID: PMC7830177.

Ashrafian H, Docherty L, Leo V, Towlson C, Neilan M, Steeples V, Lygate CA, Hough T, Townsend S, Williams D, Wells S, Norris D, Glyn-Jones S, Land J, Barbaric I, Lalanne Z, Denny P, Szumska D, Bhattacharya S, Griffin JL, Hargreaves I, Fernandez-Fuentes N, Cheeseman M, Watkins H, Dear TN. A mutation in the mitochondrial fission gene Dnm1l leads to cardiomyopathy. PLoS Genet. 2010 Jun 24;6(6):e1001000. doi: 10.1371/journal.pgen.1001000. PMID: 20585624; PMCID: PMC2891719.

Belliard A, Sottejeau Y, Duan Q, Karabin JL, Pierre SV. Modulation of cardiac Na+,K+-ATPase cell surface abundance by simulated ischemia-reperfusion and ouabain preconditioning. Am J Physiol Heart Circ Physiol. 2013 Jan 1;304(1):H94-103. doi: 10.1152/ajpheart.00374.2012. Epub 2012 Oct 19. PMID: 23086991; PMCID: PMC3543687.

Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. Behav Brain Res. 2001 Nov 1;125(1-2):279-84. doi: 10.1016/s0166-4328(01)00297-2. PMID: 11682119.

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014 Aug 1;30(15):2114-20. doi: 10.1093/bioinformatics/btu170. Epub 2014 Apr 1. PMID: 24695404; PMCID: PMC4103590.

Caforio AL, Pankuweit S, Arbustini E, Basso C, Gimeno-Blanes J, Felix SB, Fu M, Heliö T, Heymans S, Jahns R, Klingel K, Linhart A, Maisch B, McKenna W, Mogensen J, Pinto YM,

Ristic A, Schultheiss HP, Seggewiss H, Tavazzi L, Thiene G, Yilmaz A, Charron P, Elliott PM; European Society of Cardiology Working Group on Myocardial and Pericardial Diseases. Current state of knowledge on aetiology, diagnosis, management, and therapy of myocarditis: a position statement of the European Society of Cardiology Working Group on Myocardial and Pericardial Diseases. Eur Heart J. 2013 Sep;34(33):2636-48, 2648a-2648d. doi: 10.1093/eurheartj/eht210. Epub 2013 Jul 3. PMID: 23824828.

Calabrese F, Thiene G. Myocarditis and inflammatory cardiomyopathy: microbiological and molecular biological aspects. Cardiovasc Res. 2003 Oct 15;60(1):11-25. doi: 10.1016/s0008-6363(03)00475-9. PMID: 14522403.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013 Jan 1;29(1):15-21. doi: 10.1093/bioinformatics/bts635. Epub 2012 Oct 25. PMID: 23104886; PMCID: PMC3530905.

Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016 Oct 1;32(19):3047-8. doi: 10.1093/bioinformatics/btw354. Epub 2016 Jun 16. PMID: 27312411; PMCID: PMC5039924.

Friedenberg SG, Chdid L, Keene B, Sherry B, Motsinger-Reif A, Meurs KM. Use of RNA-seq to identify cardiac genes and gene pathways differentially expressed between dogs with and without dilated cardiomyopathy. Am J Vet Res. 2016 Jul;77(7):693-9. doi: 10.2460/ajvr.77.7.693. PMID: 27347821; PMCID: PMC5511690.

Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics. 2020 Apr 15;36(8):2628-2629. doi: 10.1093/bioinformatics/btz931. PMID: 31882993; PMCID: PMC7178415.

Gorbe A, Giricz Z, Szunyog A, Csont T, Burley DS, Baxter GF, Ferdinandy P. Role of cGMP-PKG signaling in the protection of neonatal rat cardiac myocytes subjected to simulated ischemia/reoxygenation. Basic Res Cardiol. 2010 Sep;105(5):643-50. doi: 10.1007/s00395-010-0097-0. Epub 2010 Mar 28. PMID: 20349314.

Hoeppner MP, Lundquist A, Pirun M, Meadows JR, Zamani N, Johnson J, Sundström G, Cook A, FitzGerald MG, Swofford R, Mauceli E, Moghadam BT, Greka A, Alföldi J, Abouelleil A, Aftuck L, Bessette D, Berlin A, Brown A, Gearin G, Lui A, Macdonald JP, Priest M, Shea T, Turner-Maier J, Zimmer A, Lander ES, di Palma F, Lindblad-Toh K, Grabherr MG. An

improved canine genome and a comprehensive catalogue of coding genes and non-coding transcripts. PLoS One. 2014 Mar 13;9(3):e91172. doi: 10.1371/journal.pone.0091172. PMID: 24625832; PMCID: PMC3953330.

Janus I, Noszczyk-Nowak A, Nowak M, Cepiel A, Ciaputa R, Pasławska U, Dzięgiel P, Jabłońska K. Myocarditis in dogs: etiology, clinical and histopathological features (11 cases: 2007-2013). Ir Vet J. 2014 Dec 24;67(1):28. doi: 10.1186/s13620-014-0028-8. PMID: 25642323; PMCID: PMC4311452.

Kayvanpour E, Sedaghat-Hamedani F, Amr A, Lai A, Haas J, Holzer DB, Frese KS, Keller A, Jensen K, Katus HA, Meder B. Genotype-phenotype associations in dilated cardiomyopathy: meta-analysis on more than 8000 individuals. Clin Res Cardiol. 2017 Feb;106(2):127-139. doi: 10.1007/s00392-016-1033-6. Epub 2016 Aug 30. PMID: 27576561.

Lamm CG, Ferguson AC, Lehenbauer TW, Love BC. Streptococcal infection in dogs: a retrospective study of 393 cases. Vet Pathol. 2010 May;47(3):387-95. doi: 10.1177/0300985809359601. Epub 2009 Dec 31. PMID: 20382824.

Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012 Mar 4;9(4):357-9. doi: 10.1038/nmeth.1923. PMID: 22388286; PMCID: PMC3322381.

Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10(3):R25. doi: 10.1186/gb-2009-10-3-r25. Epub 2009 Mar 4. PMID: 19261174; PMCID: PMC2690996.

Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014 Apr 1;30(7):923-30. doi: 10.1093/bioinformatics/btt656. Epub 2013 Nov 13. PMID: 24227677.

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. doi: 10.1186/s13059-014-0550-8. PMID: 25516281; PMCID: PMC4302049.

Mahony M, Lean D, Pham L, Horvath R, Suna J, Ward C, Veerappan S, Versluis K, Nourse C. Infective Endocarditis in Children in Queensland, Australia: Epidemiology, Clinical Features and Outcome. Pediatr Infect Dis J. 2021 Jul 1;40(7):617-622. doi: 10.1097/INF.0000000000003110. PMID: 33902079.

Mak D, Ryan KA, Han JC. Review of Insulin Resistance in Dilated Cardiomyopathy and Implications for the Pediatric Patient Short Title: Insulin Resistance DCM and Pediatrics. Front Pediatr. 2021 Oct 28;9:756593. doi: 10.3389/fped.2021.756593. PMID: 34778146; PMCID: PMC8581153.

Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 2011;17(1):10.

Mausberg TB, Wess G, Simak J, Keller L, Drögemüller M, Drögemüller C, Webster MT, Stephenson H, Dukes-McEwan J, Leeb T. A locus on chromosome 5 is associated with dilated cardiomyopathy in Doberman Pinschers. PLoS One. 2011;6(5):e20042. doi: 10.1371/journal.pone.0020042. Epub 2011 May 20. PMID: 21625443; PMCID: PMC3098859.

McNally EM, Mestroni L. Dilated Cardiomyopathy: Genetic Determinants and Mechanisms. Circ Res. 2017 Sep 15;121(7):731-748. doi: 10.1161/CIRCRESAHA.116.309396. PMID: 28912180; PMCID: PMC5626020.

Menzel P, Ng KL, Krogh A. Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat Commun. 2016 Apr 13;7:11257. doi: 10.1038/ncomms11257. PMID: 27071849; PMCID: PMC4833860.

Merkel D. Docker: lightweight Linux containers for consistent development and deployment. Linux J. 2014;2014:239.

Meurs KM, Lahmers S, Keene BW, White SN, Oyama MA, Mauceli E, Lindblad-Toh K. A splice site mutation in a gene encoding for PDK4, a mitochondrial protein, is associated with the development of dilated cardiomyopathy in the Doberman pinscher. Hum Genet. 2012 Aug;131(8):1319-25. doi: 10.1007/s00439-012-1158-2. Epub 2012 Mar 25. PMID: 22447147.

Meurs KM, Stern JA, Sisson DD, Kittleson MD, Cunningham SM, Ames MK, Atkins CE, DeFrancesco T, Hodge TE, Keene BW, Reina Doreste Y, Leuthy M, Motsinger-Reif AA, Tou SP. Association of dilated cardiomyopathy with the striatin mutation genotype in boxer dogs. J Vet Intern Med. 2013 Nov-Dec;27(6):1437-40. doi: 10.1111/jvim.12163. Epub 2013 Aug 28. PMID: 24033487.

Mitra S, Drautz-Moses DI, Alhede M, Maw MT, Liu Y, Purbojati RW, Yap ZH, Kushwaha KK, Gheorghe AG, Bjarnsholt T, Hansen GM, Sillesen HH, Hougen HP, Hansen PR, Yang L,

Tolker-Nielsen T, Schuster SC, Givskov M. In silico analyses of metagenomes from human atherosclerotic plaque samples. Microbiome. 2015 Sep 3;3:38. doi: 10.1186/s40168-015-0100-y. PMID: 26334731; PMCID: PMC4559171.

Molesan A, Goodman L, Ford J, Lovering SJ, Kelly K. The Causes of Canine Myocarditis and Myocardial Fibrosis Are Elusive by Targeted Molecular Testing: Retrospective Analysis and Literature Review. Vet Pathol. 2019 Sep;56(5):761-777. doi: 10.1177/0300985819839241.

Molina-Navarro MM, Roselló-Lletí E, Ortega A, Tarazón E, Otero M, Martínez-Dolz L, Lago F, González-Juanatey JR, España F, García-Pavía P, Montero JA, Portolés M, Rivera M. Differential gene expression of cardiac ion channels in human dilated cardiomyopathy. PLoS One. 2013 Dec 5;8(12):e79792. doi: 10.1371/journal.pone.0079792. PMID: 24339868; PMCID: PMC3855055.

Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, et al. Assembling genomes and mini-metagenomes from highly chimeric reads. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer, Berlin; 2013. p. 158–70.

Ondov BD, Bergman NH, Phillippy AM. Interactive metagenomic visualization in a Web browser. BMC Bioinformatics. 2011 Sep 30;12:385. doi: 10.1186/1471-2105-12-385. PMID: 21961884; PMCID: PMC3190407.

Philipp U, Vollmar A, Häggström J, Thomas A, Distl O. Multiple Loci are associated with dilated cardiomyopathy in Irish wolfhounds. PLoS One. 2012;7(6):e36691. doi: 10.1371/journal.pone.0036691. Epub 2012 Jun 25. PMID: 22761652; PMCID: PMC3382626.

Phillips CM, Goumidi L, Bertrais S, Field MR, Cupples LA, Ordovas JM, Defoort C, Lovegrove JA, Drevon CA, Gibney MJ, Blaak EE, Kiec-Wilk B, Karlstrom B, Lopez-Miranda J, McManus R, Hercberg S, Lairon D, Planells R, Roche HM. Gene-nutrient interactions with dietary fat modulate the association between genetic variation of the ACSL1 gene and metabolic syndrome. J Lipid Res. 2010 Jul;51(7):1793-800. doi: 10.1194/jlr.M003046. Epub 2010 Feb 22. PMID: 20176858; PMCID: PMC2882737.

Schatzberg SJ, Olby NJ, Breen M, Anderson LV, Langford CF, Dickens HF, Wilton SD, Zeiss CJ, Binns MM, Kornegay JN, Morris GE, Sharp NJ. Molecular analysis of a spontaneous dystrophin 'knockout' dog. Neuromuscul Disord. 1999 Jul;9(5):289-95. doi: 10.1016/s0960-8966(99)00011-5. PMID: 10407848.

Schultheiss HP, Fairweather D, Caforio ALP, Escher F, Hershberger RE, Lipshultz SE, Liu PP, Matsumori A, Mazzanti A, McMurray J, Priori SG. Dilated cardiomyopathy. Nat Rev Dis Primers. 2019 May 9;5(1):32. doi: 10.1038/s41572-019-0084-1. PMID: 31073128; PMCID: PMC7096917.

Sibbing D, Pfeufer A, Perisic T, Mannes AM, Fritz-Wolf K, Unwin S, Sinner MF, Gieger C, Gloeckner CJ, Wichmann HE, Kremmer E, Schäfer Z, Walch A, Hinterseer M, Näbauer M, Kääb S, Kastrati A, Schömig A, Meitinger T, Bornkamm GW, Conrad M, von Beckerath N. Mutations in the mitochondrial thioredoxin reductase gene TXNRD2 cause dilated cardiomyopathy. Eur Heart J. 2011 May;32(9):1121-33. doi: 10.1093/eurheartj/ehq507. Epub 2011 Jan 18. PMID: 21247928.

Simpson S, Edwards J, Ferguson-Mignan TF, Cobb M, Mongan NP, Rutland CS. Genetics of Human and Canine Dilated Cardiomyopathy. Int J Genomics. 2015;2015:204823. doi: 10.1155/2015/204823. Epub 2015 Jul 22. PMID: 26266250; PMCID: PMC4525455.

Singer E, Andreopoulos B, Bowers RM, Lee J, Deshpande S, Chiniquy J, Ciobanu D, Klenk HP, Zane M, Daum C, Clum A, Cheng JF, Copeland A, Woyke T. Next generation sequencing data of a defined microbial mock community. Sci Data. 2016 Sep 27;3:160081. doi: 10.1038/sdata.2016.81. PMID: 27673566; PMCID: PMC5037974.

Takeuchi S, Kawada JI, Okuno Y, Horiba K, Suzuki T, Torii Y, Yasuda K, Numaguchi A, Kato T, Takahashi Y, Ito Y. Identification of potential pathogenic viruses in patients with acute myocarditis using next-generation sequencing. J Med Virol. 2018 Dec;90(12):1814-1821. doi: 10.1002/jmv.25263. Epub 2018 Aug 13. PMID: 30011073.

Tayal U, Prasad S, Cook SA. Genetics and genomics of dilated cardiomyopathy and systolic heart failure. Genome Med. 2017 Feb 22;9(1):20. doi: 10.1186/s13073-017-0410-8. PMID: 28228157; PMCID: PMC5322656.

Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP-a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome. 2018 Sep 15;6(1):158. doi: 10.1186/s40168-018-0541-1. PMID: 30219103; PMCID: PMC6138922.

van Buul JD, Hordijk PL. Signaling in leukocyte transendothelial migration. Arterioscler Thromb Vasc Biol. 2004 May;24(5):824-33. doi: 10.1161/01.ATV.0000122854.76267.5c. Epub 2004 Feb 19. PMID: 14976004.

Wiersma AC, Leegwater PA, van Oost BA, Ollier WE, Dukes-McEwan J. Canine candidate genes for dilated cardiomyopathy: annotation of and polymorphic markers for 14 genes. BMC Vet Res. 2007 Oct 19;3:28. doi: 10.1186/1746-6148-3-28. PMID: 17949487; PMCID: PMC2194671.

Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol. 2019 Nov 28;20(1):257. doi: 10.1186/s13059-019-1891-0. PMID: 31779668; PMCID: PMC6883579.

Zarrouk Mahjoub S, Mehri S, Ourda F, Finsterer J, Ben Arab S. Novel m.15434C>A (p.230L>I) Mitochondrial Cytb Gene Missense Mutation Associated with Dilated Cardiomyopathy. ISRN Cardiol. 2012;2012:251723. doi: 10.5402/2012/251723. Epub 2012 Jul 3. PMID: 22811935; PMCID: PMC3395144.

Zhang HB, Li RC, Xu M, Xu SM, Lai YS, Wu HD, Xie XJ, Gao W, Ye H, Zhang YY, Meng X, Wang SQ. Ultrastructural uncoupling between T-tubules and sarcoplasmic reticulum in human heart failure. Cardiovasc Res. 2013 May 1;98(2):269-76. doi: 10.1093/cvr/cvt030. Epub 2013 Feb 11. PMID: 23405000.

Zhu Z, Ren J, Michail S, Sun F. MicroPro: Using metagenomic unmapped reads to provide insights into human microbiota and disease associations. Genome Biol. 2019;20(1):154

La borsa di dottorato è stata cofinanziata con risorse del
Programma Operativo Nazionale Ricerca e Innovazione 2014-2020 (CCI 2014IT16M2OP005),
Fondo Sociale Europeo, Azione I.1 "Dottorati Innovativi con caratterizzazione Industriale"

UNIONE EUROPEA
Fondo Sociale Europeo

Ministero dell'Università e della Ricerca

PON RICERCA E INNOVAZIONE 2014 - 2020