UNIVERSITÀ POLITECNICA DELLE MARCHE

DIPARTIMENTO SCIENZE DELLA VITA E DELL'AMBIENTE

Corso di Laurea Magistrale

Biologia Molecolare e Applicata

Un approccio a livello di sistemi per decifrare gli effetti di alterazione endocrina dell'Aflatossina B1

A systems level approach to decipher the endocrine disrupting effects of Aflatoxin B1

Tesi di Laurea Magistrale
di:

Jacopo Umberto Verga

Relatore
Chiar.ma Prof.ssa
Oliana Carnevali

Correlatore:
Chiar.mo Prof.
Gary Hardiman

Sessione Straordinaria 22-26 Febbraio 2021

Anno Accademico 2019/2020

ABSTRACT

Aflatoxins are a family of secondary metabolites produced by different species of *Aspergillus sp*. Aflatoxin contamination in crops, foods and animal feed can occur in any time pre and post-harvest stage, making it difficult to control the contamination rate of foods. Despite the strict regulations, aflatoxins contaminations are still widespread and contaminated foods are causative of a large part of hepatocellular carcinoma (HCC) and aflatoxicosis in many countries. The most studied and toxic aflatoxin is Aflatoxin B1 (AFB1). The carcinogenic effects act through two main ways: forming alkylated DNA adducts and by oxidative stress. The first way act through the highly reactive metabolite AFB1-exo and -endo 8,9 epoxide (Epo-AFB1) that can intercalate with DNA and react with N 7 of Guanine. The second way the increased levels of reactive oxygen species (ROS) damages the DNA, proteins and lipids. Also, studies of differential gene expression concerning the AFB1 exposure on placental and testicular tissues suggested AFB1 as a potential endocrine disruptor. We hypothesise that prolonged exposures to Aflatoxin B1 and its metabolites can directly interact with Estrogen Receptors (ERs) α and β, impairing gene expression and pathways related to this class of receptors. We used an integrative

system biology approach that considered structural information from the receptors and aflatoxins, gene expression data and protein-protein interaction (PPI) network analysis to test the hypothesis. Molecular Docking simulations and interaction analysis suggest that AFB1 can enter the binding site. Differentially expressed genes and biological processes observed after AFB1 and ethinylestradiol exposures both in vitro and in vivo, highlighted key processes and expression of genes involved in liver diseases and carcinogenesis. Finally, the protein-protein interaction network highlighted the strong estrogenic components providing an additional hint of the estrogenic component in the mechanisms of action of aflatoxin.

Le aflatossine sono una famiglia di metaboliti secondari prodotti da diverse specie di *Aspergillus sp.* La contaminazione da aflatossine nelle colture, negli alimenti e nei mangimi per animali può verificarsi in qualsiasi momento prima e dopo la raccolta, rendendo difficile il controllo del tasso di contaminazione degli alimenti. Nonostante le rigide normative, le contaminazioni da aflatossine sono ancora molto diffuse e gli alimenti contaminati sono causa di gran parte dello sviluppo di carcinomi epatocellulare (HCC) e aflatossicosi in molti paesi. L'aflatossina più studiata e tossica è l'aflatossina B1 (AFB1). Gli effetti cancerogeni agiscono attraverso due modalità principali: formando addotti alchilati del DNA e lo stress ossidativo. Nel primo caso i metaboliti altamente reattivi AFB1-exo e -endo 8,9 epossido (Epo-AFB1) possono intercalarsi con

il DNA e reagire con N 7 di Guanina. Nel secondo caso, l'aumento dei livelli di specie reattive dell'ossigeno (ROS), causa danni a DNA, proteine e lipidi. Inoltre, studi di espressione genica differenziale riguardanti l'esposizione all'AFB1 sul tessuto placentare e testicolare, hanno suggerito l'AFB1 come un potenziale interferente endocrino. Si ipotizza che esposizioni prolungate all'Aflatossina B1 e ai suoi metaboliti possano interagire direttamente con i recettori degli estrogeni (ER) α e β, alterando l'espressione genica e le vie correlate a questa classe di recettori. In questo studio abbiamo utilizzato un approccio di biologia dei sistemi integrati che considera le informazioni strutturali dai recettori e dalle aflatossine, i dati di espressione genica e l'analisi di rete dell'interazione proteina-proteina (PPI) per testare l'ipotesi. Le simulazioni di docking molecolare e l'analisi d'interazione suggeriscono che AFB1 può entrare nel sito di legame. Geni espressi in modo differenziale e processi biologici osservati dopo esposizioni ad AFB1 ed etinilestradiolo, hanno evidenziato - sia *in vitro* che *in vivo* - l'attivazione di processi chiave ed espressione di geni coinvolti nelle malattie del fegato e nella carcinogenesi. Infine, la rete di interazione proteina-proteina ha messo in luce le forti proprietà estrogeniche suggerendo il coinvolgimento dei recettodi degli estrogeni nei meccanismi di azione delle aflatossine.

Contents

List of Figures **xi**

List of Tables **xv**

Abbreviations **xvi**

1 INTRODUCTION **1**

1.1 Aflatoxins 1

1.2 Endocrine Disruptors 8

1.2.1 Mycoestrogens 9

1.2.2 Ethinylestradiol 10

1.3 Molecular Docking 13

1.3.1 AutoDock 15

1.3.2 AutoDockVina 17

1.4 Microarray Technology 18

1.4.1 Expression Arrays 21

1.4.2 Comparative Genomic Hybridization (CGH) Array 21

1.4.3 ChIP-chip 22

1.4.4 SNP arrays 23

1.5 Protein-Protein Interaction (PPI) Networks 24

2 AIMS OF THE STUDY 27

3 MATERIALS and METHODS 31

3.1 Molecular Docking 31

3.1.1 Ligands 31

3.1.2 Receptors 32

3.1.3 Simulations 32

3.2 Expression analysis 34

3.2.1 *In vitro* samples 34

3.2.2 *In vivo* samples 35

3.2.3 Microarray analysis 35

3.2.4 Gene Ontology analysis 37

3.3 Protein-Protein Interaction (PPI) networks 38

3.3.1 Network construction 38

3.3.2 Identification of Downstream Estrogen Receptor Targets (DERT) 39

3.3.3 Network analysis 40

3.4 Graphs and Figures 41

4 RESULTS 43

4.1 Molecular Docking 44

4.1.1 AutoDockVina 45

4.1.2 AutoDock4 47

4.1.3 Poses and interactions 49
4.2 Expression Analysis 56
4.2.1 DE genes partially overlap after *in vitro* exposures to Aflatoxin B1 and Ethinylestradiol 56
4.2.2 Gene ontology (GO) annotation of *in vitro* DE genes 57
4.2.3 EE2 early exposure and AFB1 repeated exposures shows partial overlapping *in vivo* 62
4.2.4 GO analysis revealed estrogenic and cancerogenic markers were shared by the *in vivo* exposures to AFB1 and EE2 67
4.3 Protein-Protein Interaction (PPI) networks analysis 75
4.3.1 *In vitro* exposure 75
4.3.2 *In vivo* 3-day exposure 78
4.3.3 *In vivo* 5-day exposure 81

5 DISCUSSION 85
5.1 Docking and Interaction analysis 85
5.2 Expression analysis 87
5.2.1 *in vitro* 87
5.2.2 *in vivo* 90
5.3 PPI network analysis 95

6 CONCLUSION 99

7 ACKNOWLEDGMENTS 101

Bibliography **101**

Appendices **137**

A Interactions **139**

B BP terms **145**

List of Figures

1.1 Aspergillus sp - Conidiophores 1

1.2 Estradiol and Ethinyl estradiol molecules. 13

2.1 Pipeline to extract the estrogenic component of the Aflatoxin B1 mechanism of action. 29

4.1 Molecular structure of the compounds used for Molecular Docking simulations. 44

4.2 Binding poses on ERα after AutoDock4 simulations. . . . 51

4.3 Comparison of the interactions occurred among AFB1 and EE2 with ERα 51

4.4 Comparison of the interactons occurred among AFQ1 and EE2 with ERα 52

4.5 Binding poses on ERβ after AutoDock4 simulations. . . . 54

4.6 Comparison of the interactions occurred among AFB1 and EE2 with ERβ 55

4.7 Comparison of the interactions occurred among AFQ1 and EE2 with ERβ 56

4.8 Grapich of DE genes filtered by q-value in Human hepatocytes. 58

4.9 Venn diagram showing DE genes after AFB1 and EE2 *in vitro* exposure. 59

4.10 Venn diagram showing overlapping BP after *in vitro* exposures from iPathway Guide (Advaita). 61

4.11 Graphic of DE genes after AFB1 exposure in rats, filtered by q-value. 64

4.12 Graphic of DE genes after EE2 exposure in rats, filtered by q-value. 65

4.13 Venn diagram showing DE genes after AFB1 and EE2 *in vivo* exposures. 66

4.14 Venn diagram showing overlapping BP after *in vivo* exposures. 71

4.15 Comparison of the high centrality nodes in the PPI network from AFB1 *in vitro* exposure 77

4.16 Most central nodes and DERTs of the PPI network from AFB1 *in vitro* exposure 78

4.17 Comparison of the high centrality nodes in the PPI network from AFB1 *in vivo* 3-day exposure 80

4.18 Most central nodes and DERTs of the PPI network from AFB1 *in vivo* 3-day exposure 81

4.19 Comparison of the high centrality nodes in the PPI network from AFB1 *in vivo* 5-day exposure 83

4.20 Most central nodes and DERTs of the PPI network from the AFB1 *in vivo* 5-day exposure 84

5.1 Stacking interaction among aromatic ring of EST and *Phe404* of ERα. 87

5.2 Superimposition of ERα and β. 88

A.1 Interactions occurring among every ligand and ERα residues. 140

A.2 Interactions occurring among every ligand and ERα residues. 141

A.3 Interactions occurring among every ligand and ERβ residues. 142

A.4 Interactions occurring among every ligand and ERβ residues. 143

List of Tables

4.1 AutoDockVina binding affinities and K_Ds. 46

4.2 AutoDock4 binding affinities and K_Ds. 49

4.3 Number of DE genes filtered by q-value in Human hepatocytes. 58

4.4 ERs Log2FC and q-values from *in vitro* exposures 58

4.5 BP shared by AFB1 and EE2 *in vitro* exposures from iPathway Guide (Advaita) meta-analysis. 62

4.6 Number of DE genes after AFB1 exposure in Rats, filtered by q-value. 65

4.7 Number of DE genes after EE2 exposure in rats, filtered by q-value. 65

4.8 Overlapping DE genes among EE2 and AFB1 *in vivo* exposures. 67

4.9 Overlapping BP terms among EE2 and AFB1 *in vivo* exposures. 72

4.10 BP terms shared by AFB1 3-day and both EE2 *in vivo* exposures from iPathway Guide (Advaita) meta-analysis. . . 73

4.11 BP terms shared by AFB1 5-day and both EE2 *in vivo* exposures from iPathway Guide (Advaita) meta-analysis. . . 74
4.12 DERTs recognised by centrality analysis and q-values for AFB1 *in vitro* exposure. 77
4.13 DERTs recognised by centrality analysis and q-values for AFB1 *in vivo* 3-day exposure. 80
4.14 DERTs recognised by centrality analysis and q-values for the AFB1 *in vivo* 5-day exposure. 83

B.1 Unique BP signatures for *in vitro* exposures from iPathway Guide (Advaita). 167
B.2 Unique BP signatures for AFB1 *in vivo* exposures from iPathway Guide (Advaita). 180
B.3 Shared BP by AFB1 *in vivo* exposures from iPathway Guide (Advaita) 181
B.4 Unique BP signatures for EE2 *in vivo* exposures from iPathway Guide (Advaita). 195
B.5 Shared BP by EE2 *in vivo* exposures from iPathway Guide (Advaita). 199

Abbreviations

ADT AutoDockTools.

ERβ Estrogen Receptor β.

ERα Estrogen Receptor α.

DE Differentially Expressed.

AFB1 Aflatoxin B1.

Epo-AFB1 Aflatoxin B1-epoxide.

AFM1 Aflatoxin M1.

AFQ1 Aflatoxin Q1.

AFL Aflatoxicol.

EE2 Ethinylestradiol.

RIF Rifampin.

FDR False Discovery Rate, q-value.

adj.p-value p-value corrected with the Bonferroni method.

GO Gene Ontology.

BP Biological Process.

DERT Downstream Estrogen Receptor Target.

ENSP Ensemble protein IDs.

EST 17-β-Estradiol.

HCC Hepatocellular Carcinoma.

ROS Reactive Oxigen Species.

8-oxo-dG 7,8-dihydro-8-oxo-2 0 -deoxyguanosine.

GA Genetic Algorithms.

SNP Single Nucleotide Polimorphism.

DMD Digitial Micromirror Device.

CGH Comparative Genomic Hybridization.

aCGH Array-CGH.

ChIP-chip Chromatin Immuno-Precipitation on chip.

TF Transcription Factor.

Chapter 1

INTRODUCTION

1.1 Aflatoxins

Figure 1.1: Aspergillus sp - Conidiophores.

Aflatoxins are a family of secondary metabolites, also known as mycotoxins, produced by different species of *Aspergillus sp.* (Figure 1.1). The

Aspergillus species better known to produce Aflatoxins are *A. flavus* and *A. parasiticus*. The name Aflatoxins derives from "*A. flavus* toxins" after the turkey epidemic known as "Turkey X disease" that, in 1960, caused the death of more than 100,000 animals in England, and was linked to *A. flavus* feeds contamination [1]. *A. flavus* is a widespread fungus, that can grow ubiquitously in soils and can contaminate a wide range of crops and grins including rice, maize and peanuts [1]. While *A. parasiticus* is more adapted to grow in soils, its distribution is more limited than *A. flavus* and mainly produces aflatoxin B1 (AFB1), aflatoxin B2 (AFB2), aflatoxin (AFG1) and aflatoxin (AFG2), *A. flavus* better grows and contaminates aerial parts of the plants producing AFB1 and AFB2 and leading to more critical foods contaminations [2]. Many other *Aspergillus* species produces aflatoxins including species close to *A. flavus* (*A.minisclerotigenes, A. korhogoensis, A. aflatoxiformans and A. texensis*) and species closer to *A. parasiticus* (*A. novoparasiticus and A. arachidicola*) that in addition to producing G and B type aflatoxins produces Aflatoxicol and other toxins [2]. *Aspergillus* growth and the relative aflatoxin contamination in crops, foods and animal feed can occur in any time pre and post-harvest stage, making it difficult to control the contamination rate of foods. Moreover, high humidity and temperatures improve fungal growth making it challenging from countries with warm climates and third world countries, to safely storage harvested foods, and these environments more prone to aflatoxin contaminations and acute intoxications (aflatoxicosis) cases [1]

[2]. The most studied and toxic aflatoxin is Aflatoxin B1. Also, it is the most common aflatoxin in food contaminations and usually, it is present in other aflatoxins contaminations [2]. Aflatoxin B1 is documented well as the most cancerogenic aflatoxin; indeed it is classified in the group 1 of carcinogenic to humans by IARC and its exposure is a causative agent of Hepatocellular Carcinoma (HCC) development, principally in cooperation with hepatitis B and C viruses [1] [2]. Aflatoxin B1 exert various toxic effects, primarily on the liver, that are still not fully elucidated. The main cancerogenic and genotoxic effects act through two separate ways. The first and better known is the effects of the highly reactive metabolite AFB1-exo and -endo 8,9 epoxide (Epo-AFB1) produced by CYP450 enzymes. It is highly nucleophilic, can intercalate with DNA and react with N^7 of Guanine, forming alkylated DNA adducts that, due to the high instability of the adduct, can lead to apurinic sites or, opening the imidazole ring, AFB1-formamidopyrimidine [3]. Alkylated DNA adducts, apurinic sites and AFB1-formamidopyrimidine are the principal cancerogenic precursors of AFB1. Interestingly, reactions between Epo-AFB1 and guanine occur in the third guanine of the codon 249 on P53 tumour suppressor gene causing a G-to-T transversion and inactivating the gene [2] [3]. This mutation is considered to be a signature of aflatoxin induced HCC [2]. The second genotoxic effect act through oxidative stress, indeed AFB1 metabolism in the liver by CYP450 enzymes releases excessive amounts of reactive oxygen species (ROS). ROS can attack DNA causing

a wide variety of adducts, it was observed a significant increase of the most known DNA adduct, 7,8-dihydro-8-oxo-2 0 -deoxyguanosine (8-oxo-dG), after intraperitoneal injection of AFB1 to rats [3]. Similar to Epo-AFB1 adducts, 8-oxo-dG can cause the G-to-T transversion in a non-specific way, leading to various mutagenic and even tumorigenic effects [3]. Observing the genes deregulated by AFB1 and AFM1, the AFB1 metabolite found in the milk of animals exposed to contaminated feeds, some pathways inherent to steroid hormone synthesis, nuclear receptor signalling and estrogen signalling pathways were found, providing a suggestion regarding a potential estrogenic component in aflatoxin mechanism of action [4]. Other secondary effects of AFB1 toxicity include immunotoxicity, teratogenicity, growth impairment and malnutrition [3]. Aflatoxin B1 exerts different immunotoxic effects, high and prolonged exposures reduced both the innate and the adaptive immune response, increasing the susceptibility to infections (e.g. hepatitis) and decreased vaccination efficacy [2] [3]. On the other hand, low AFB1 exposures exert immunostimulatory effects upregulating the production of both the inflammatory and the anti-inflammatory cytokines, upregulation of TLR expression and stimulating T-cell proliferation through increasing he antigen-presenting capacity of dendritic cells [3]. Evidence, suggests that AFB1 and its metabolite Aflatoxicol (AFL), produced by an undefined NADPH reductase, can cross the placenta and be reconverted to AFB1 [2] [5]. Also, after a single dose of AFB1 5 mg/kg in gestating mice, the same DNA adducts (with N^7 of Guanine and AFB1-

formamidopyrimidine) were found both in mother and foetus [2]. The effects of AFB1 exposure in placental tissue caused hormone synthesis and homeostasis impairment, suggesting AFB1 as a potential endocrine disruptor [2] [5]. Other reported effects of AFB1 intake during pregnancy are related to foetal loss, foetal growth restriction and premature birth; the inverse relationship between the birth weight and AFB1 intake are well documented [1]. An additional hint regarding the endocrine-disrupting effects of AFB1 or reproductive tissue derives from mice testis exposed to AFB1 that showed the deregulation of Estrogen Receptor α (ERα) in addition to some proteins involved in cell cycle regulation including CDK-4, cyclin D1 and p21 [6]. Due its toxic effects, AFB1 is strictly regulated in many countries considering safe to human consumption AFB1 contamination levels in a range of 0 to 40 μg/kg and for animals up to 300 μg/kg. These strict regulations cause great losses for the agricultural industry. It was estimated that annual losses are up to USD 1.68 billion for AFB1 contaminated crops. Also, African exports estimate 64% of losses if European standards were applied [1]. Despite the strict regulations, aflatoxins contaminations are still widespread and contaminated foods are causative of a large part of HCC and aflatoxicosis in many countries [1]. Studies of AFB1 dietary intakes found in Swaziland and Mozambique a range of 14.3 to 183.7 ng/kg bw$\times$day, with a median of 58.85 ng/kg bw$\times$day [7]. Also, in China, the daily AFB1 intake was estimated in a range of 11.7 to 2027.4 ng/kg bw$\times$day with a median value of 397.25 4 ng/kg bw$\times$day [7]. Due to

the great efforts to control contamination levels in foods, European intakes are significantly lower. A study conducted on the Belgian population's dietary habits, considered 5,299 food samples from different countries to model two different scenarios, the first with all data available and the second only the samples with AFB1 contamination level below legal limits. The first scenario estimated an AFB1 daily intake of 7.99 ng/kg bw$\times$day and the second a daily intake of only 0.028 0.028 [8]. Despite these optimistic estimations, a study conducted on urine and serum AFB1 levels from five European nations, uncovered intakes of 5.61 ng/kg bw$\times$day [9], significantly above the intakes with only food samples that respected legal AFB1 contamination limits [8]. Dietary intakes levels reflect HCC occurrences in different countries and continents. Asia and particularly China accounts the highest HCC incidence rates. Also, the highest HCC related deaths were observed in Easter Asia, followed by Asian-Pacific countries [10]. Africa was the second continent for HCC incidence rate, and the main cause of cancer-related deaths remain the HCC mortality [10]. European HCC incidence is lower than in Asia and Africa but showed distinct rates within their countries. Indeed, the Western Europe region exhibited high HCC incident and mortality rates, so much to locate it respectively third and fourth on a global scale [10]. *Aspergillus sp.* fungi better grow in high humidity and temperature conditions, preferring tropical and subtropical areas. Thus, until early 2000 it was not a concern for primary production temperate areas including Europe. In recent years several afla-

toxin outbreaks were recorded in South-European regions including Italy in 2003 and Serbia in 2012 [2] [11]. These events suggest that the climate change conditions including the increase of medium temperatures and humidity could enlarge the *Aspergillus sp.* habitats and adding European nations in high-risk contamination areas. Two predictive models based in +2°and +5°C climate changes scenarios estimated the increase of maize and wheat contamination risk in numerous European nations [11]. Particularly, the +2°C climate change scenario estimated a high risk of aflatoxins contamination of maize harvest in South-European nations while the +5°C scenario estimated medium to high risk of contamination of maize harvest in South-European nations [11]. Additionally, the +5°C scenario estimated the increased risk of contamination to low and medium for both maize and wheat harvests in northern nations including Germany, Poland, Scandinavian countries and the south of the United Kingdom [11]. Another study considered water activity (a_W) × temperature × elevated CO_2 as key climate change factors to evaluate AFB1 production regulators activity (*aflD and aflR*) in *A. flavus*. These results were used to estimate the fold changes in AFB1 production in 34°and 37°C conditions. For the 34°C condition the increasing to AFB1 production increased up to 2.6 fold change, the 37°C reported a possible increasing on AFB1 production up to 79.2 fold change, depending on CO_2 and a_W situations [12]. These studies suggest that both the *A. flavus* habitat and AFB1 production could significantly increase in the coming years due to climate change, increasing the risk of

dietary intakes and related disease. This makes aflatoxin an important topic that must be studied to elucidate its mechanisms of action and prevents the populations to be dangerously exposed to the toxins.

1.2 Endocrine Disruptors

Endocrine-disrupting chemicals (EDCs) are naturally found or man-made chemicals that can interfere with regulation, production and signalling of endogenous hormones *in vivo* in both humans and wildlife. Disruption of androgens, estrogens and thyroid hormones can lead to various adverse outcomes including altered reproductive function, abnormal growth patterns, immunity, cancer, and neurodevelopmental delays in children [13]. The mechanism of action of the EDCs can act altering hormone metabolism, biosynthesis and regulation blocking the interactions between the hormone and the receptors, binding the hormone receptor inhibiting binding of the endogenous hormone, also acting as an agonist or antagonist of the receptor (anti-estrogens and anti-androgens) or at post-receptor levels disrupting hormone signaling pathways [13] [14]. EDCs alterations include genomic effects modifying gene expression via epigenetic mechanisms, effects that were observed in progeny [14]. EDCs are ubiquitous and can derive from natural resources (hormones, phytoestrogens, mycoestrogens) or from chemicals including pesticides, dioxins, pharmaceuticals such as ethinylestradiol or methylandrogen, polychlorinated biphenyls, and plasticizers such as Bisphenol A (BPA) [13]. EDCs can be persistent in

the environment, caccumulating to high concentrations in wildlife and humans. Others are less persistent in the environment due to shorter half-lives but they can still be of concern due to continuous exposure [14].

1.2.1 Mycoestrogens

In nature various EDCs are synthesized from plants (phytoestrogens) and fungi (mycoestrogens). One of the most widespread mycoestrogens is Zearalenone (ZEA), which, despite its non-steroidal structure, can bind both the ERα and ERβ acting as competitive agonists/antagonist on both the receptors [15]. It is produced by some *Fusarium* species that can contaminate wheat, barley, rice, maize, and some other crops, the contamination can occur pre and post-harvesting and it is reported on almost all continents [15]. Due to the high number of crops that *Fusarium sp.* can infest, ZEA is one of the most common mycoestrogens in foods and feeds [15]. In animals, ZEA exposures caused several reproductive conditions including fibrosis of the uterus, persistent estrus and sterility. The effects were observed in pigs, ruminants and poultry leading to hyperestrogenism [15]. Studies to elucidate its role in cancer development and progression highlighted the proliferation-stimulating effects on human cell lines [14]. Also, ZEA exerts anti-apoptotic effects and alters the expression of proteins involved in cell-cycle [14]. Interestingly, the fact that lower doses of ZEA can increase proliferation and higher doses reduce the viability of cells mirrors estrogenic function, where low doses are needed to mantain

physiological processes and estrogenic imbalance may lead to carcinogenesis [14].

1.2.2 Ethinylestradiol

Another well-known endocrine disruptor is the synthetic hormone 17α-ethinylestradiol (EE2) used in contraceptive formulations, treatment of osteoporosis, hormone replacement therapy and hormone-sensitive cancers. EE2 is a synthetic hormone deriving from estradiol (EST) with an ethynyl substitution in C17α position (Figure 1.2). Since the global industrialization, estrogens revealed to be an emerging problem due to their rapidly increasing concentrations in soils and waters, and their endocrine-disrupting properties. It is estimated that a human population of 7 billion can discharge 30,000 kg/yr of natural estrogens. Regarding the synthetic estrogen EE2, the estimated annual discharges are approximately 700 kg/yr. only from birth control pills. Moreover, considering the estrogens discharged by livestock manufacture the amount of hormones released in the environment rises to 83,000 kg/yr [16]. EE2 environmental concentrations were found up to 5.6 ng/L in Sewage Treatment Plants (STPs) and 4.67 ng/L in Sea waters showing different proportions with other estrogens depending on the site of examination. Indeed, river waters from Brazil, Mediterranean countries including Italy, Israel and Far East countries including China, Thailand and Indonesia, reported high rates of EE2 pollution compared to other natural estrogens[16]. EE2, if compared to natural estrogens, is a

compound with a longer half-life that can reach 81 days in aquifer material under aerobic conditions, in other aquatic environments, the half-life could decrease due to the breakdown by photocatalysis. Moreover, it was observed that EE2 is slowly degraded by anaerobic bacteria and under reducing condition the dissipation period could exceed 1000 days adding additional concerns about the possibility of accumulation [16]. Several studies asserted the deleterious effects that natural and synthetic estrogens exert on the wildlife. EE2 is the most potent estrogen with a Predicted No Effect Concentration (PNEC) of 0.1 ng/L in aquatic repeated exposures [16]. Elevated concentrations of estrogens in waters cause feminization in male fishes, reduced reproductive fitness, lower sperm count, increased vitellogenin production and other reproductive characteristics. Other than reproductive issues EE2 reduced fish biomass and interrupted the aquatic food chain [16]. Studies on Estrogen (E2) and EE2 effects on liver reported opposite outcomes and suggested that the two hormones could exert protective or damaging effects, depending on the dosage, hormonal equilibrium, and sex of the treated animals. Indeed, studies on EE2 treatment on ovariectomized rats showed reduced lung metastasis of HCC [17]. Moreover, menopausal women taking menopause hormone therapy resulted in a lower risk of developing HCC and increased survival expectations for HCC patients [18]. On the other hand, several studies have highlighted how xenoestrogen exposures and hormone deregulation leads to several hepatic pathological conditions. Exposure to EE2 5 mg/kg to

male rats for 5 days caused liver damage, severe bile duct hyperplasia and induced hepatic morphologic changes [19]. Models of inflammation on rat hepatocytes highlighted how EE2 can activate several inflammatory pathways, cause cell necrosis, decreased glutathione (GSH) and superoxide dismutase (SOD) levels in liver and intestine [19][20]. The mechanism for cholestasis is still unclear but is supposed to involve estrogen receptor signalling pathways [20]. Similar to rat, fish (*Cnesterodon decemmaculatus*)treated with EE2 showed hepatoxicity leading to liver alterations and damage, including single-cell necrosis, macrovesicular steatosis with cystic degeneration [21]. Also, global gene expression studies in primary rainbow trout (*Oncorhynchus mykiss*) hepatocytes treated with EE2 uncovered a large amount of transcriptional and estrogen biomarkers changes [22]. Microarray and Gene Ontology (GO) enrichment analysis highlighted various toxicity pathways, differentially affected by the concentration of exposure. Also, lipid metabolism alteration with the deregulation of PPARs expression (*pparα and pparγ*, impaired cellular growth and proliferation, and several cancer-related pathways where observed [22]. Suggesting that the potential toxic mechanism of EE2 affects lipid metabolism, ER signalling and cellular responses mediated by ERs. As chronic liver disease and liver cancer are much more common in males, those with cirrhosis and HCC have elevated levels of estrogen, the effects of EST on male and female zebrafish were studied. EST treatment increased liver growth and size, induced cell proliferation through activation of G protein-coupled es-

trogen receptor 1 (GPER1) and the activation of PI3K-mTOR signalling, suggesting the involvement of GPER1 in estrogen-induced carcinogenesis. Moreover, the growth and proliferation-inducing effects were more marked in males, suggesting higher tolerance of females to EST exposures [23].

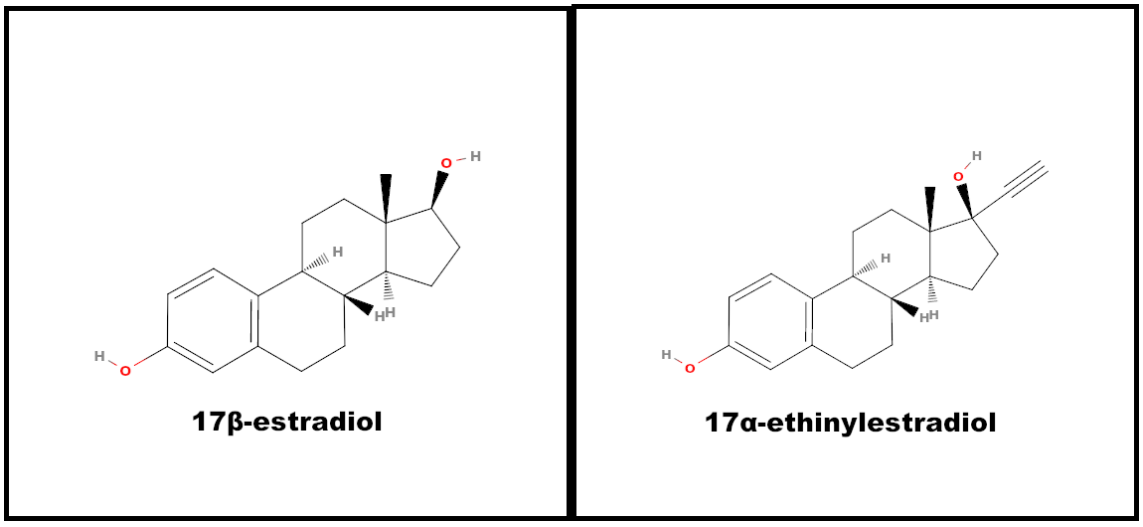

Figure 1.2: EE2 is a synthetic hormone deriving from estradiol with an ethynyl substitution in C17α position. Structure images were obtained with http://molview.org/.

1.3 Molecular Docking

Molecular Docking is a computational approach used to identify and estimate the interaction between two molecules depending on the structure, shape and features of the two. Its main use is protein-drug molecular docking, for studying interactions among a macromolecule, usually a protein, and a small ligand, tipically a lead drug candidate. Interestingly, with this technique, other molecular binding interactions can be studied, including protein-protein, nucleic acid-protein, nucleic acid-nucleic acid [24].

Since traditional *in vitro* research, aiming to elucidate the effectiveness and binding mode of a single molecule, can take several years and high costs (around 800 million USD), computational approaches that can investigate thousands of ligands significantly diminish these costs and times [24]. Since 1980s Molecular Docking approaches have become one of the most common techniques in drug discovery and design. In parallel to the increase in computational resources, both the scoring functions used to estimate the best binding poses and the algorithms used to calculate the interactions formed among the two molecules have grown in complexity and reliability [24]. The first software initially considered both the structures as rigid shapes. An example is DOCK, one of the pioneering search algorithms, it explores possible orientations of the ligand inside the binding pocket calculating the geometric distance among the ligand atoms and the macromolecule residues considered as spheres [25]. This approach is called ”rigid docking”, it considers heuristically all most favorite poses, moving by six degrees of freedom the rigid ligand, and identifes the best matches that show optimal binding energies. The energies are estimated using the Van der Waals interactions between ligand and macromolecule [24]. Rigid docking algorithms were innovative and paved the way for numerous improvements, DOCK was enhanced with new functions and conferring flexibility to the ligand. This technique was very useful to enhance the drug discovery process of HIV-1 protease [26]. Later, many other approaches evolved to increase the precision and reliability of the al-

gorithms. Flexible docking can better imitate the behaviour of molecules in a biological context, it attempts to evaluate conformational changes of the ligand and rotations on the backbone and side-chains of the macromolecule. Today there are two main differences in approaches to integrate flexibility in docking simulations. The first one is the systematic incremental construction algorithms, where the pieces of the fragmented ligand are docked on the binding pocket as rigid anchors, the incremental construction joins all the fragments rebuilding the fragment adapted to the binding pocket. This approach is used by numerous search algorithms including DOCK [27], Glide [28] and Hammerhead[29] [24]. The second approach uses stochastic algorithms to cross the most probable conformations of the ligand. With this technique, the ligand is taken as a chromosome and all the fragments as genes, through the generations the individual fragments can mutate independently, every mutation consists of a different conformation. This method, called genetic algorithms (GA), is integrated with Lamarkian GA in AutoDock. Other software also integrates other stochastic methods including Monte Carlo and Metropolis, where the conformations are explored focusing on minimum energies and choosing the most thermodynamical stables poses [24].

1.3.1 AutoDock

AutoDock, released in 1990, with subsequent versions developed later, are some of the most common platforms used to perform Molecular Docking

and Virtual Screenings [30]. It uses the Metropolis method and Lamarkian GAs to perform conformational searching on a population of trial conformations [30][31]. The successive generations "mutate" by changing the torsional angles and executing translations on the rigid body. The energy of interaction of the new conformations and binding poses are evaluated by a semi-empiric force field based on Amber [30][31]. Those conformations that show lower binding energy than the parent are automatically accepted, the other conformations are treated probabilistically: a random number between 0 and 1 is generated, if the number is lower than the formula below 1.1, the conformation is accepted [31].

$$X < e^{\frac{-\Delta E}{k_B T}} \tag{1.1}$$

ΔE = Difference in energy between the new and the previous pose (Kcal/mol)

k_B = Boltzmann's constant 1.380649×10^{-23}

T = Temperature (°K)

After a certain number of cycles, the process is stopped resulting in a large number of conformations and poses in local minima [31]. Interestingly, to allow rapid evaluations of the binding energy, a grid-based method is used. A grid that includes the protein or the binding pocket is built, then for every grid point, the interaction energy is computed and saved for every atom type. The grid scores are used to quickly calculate the binding energies, permitting reduction in computational times[30]. AutoDock allows optimization of several parameter settings including temperature,

search algorithm selection, energy and the number of Genetic Aglorithm (GA) steps. Also, it is possible to select residues from the macromolecule and fragments of the ligand to perform full flexible docking analysis[30]. Despite the computational optimization, AutoDock does not operate with multi-thread processing, making it difficult to work with large numbers of ligands. Therefore, newer faster-calculating versions have been developed such as AutoDockVina [32].

1.3.2 AutoDockVina

AutoDockVina [32] was developed after almost 20 years by the same laboratory of developed AutoDock. Similarly to AutoDock, it uses a stochastic method optimized with the Metropolis algorithm, but it differs for some aspects. The genetic algorithms, simulated annealing and other algorithms were optimized to allow faster simulations[32]. Also, it implemented a quasi-Newton optimization method to take care of the gradient of binding energy among the generations, allowing a more precise search of global minima [32]. The scoring function is inspired by X-score function [33], and the optimization for multi-threading processing made AutoDockVina one of the most common tools used in Virtual Screenings due to its fast.scomputing properties.

1.4 Microarray Technology

The DNA and RNA microarray technology relies on solid support with several thousands of specific DNA sequences (probes) attached in it. Using samples of labelled nucleic acids (such as genomic DNA or cDNA from mRNA) it is possible to hybridize the mixture of target sequences with the probes and, after washing the excess of the sample, quantify those complementary sequences to each probe measuring the fluorescence of the related spot. DNA microarray technology derives from previous procedures initially developed to isolate plasmids containing cloned DNA of interest from *E. coli* cultures [34]. Here the *E. coli* colonies firstly were transferred to a nitrocellulose filter, the DNA is anchored to the filter through cell lysis, DNA denaturation and fixation. The hybridization of radiolabelled probes allows the identification of the genes of interest [34]. This approach underwent numerous improvements, the nitrocellulose filters were substituted with microtiters, allowing to test thousands of colonies. Also, the sensibility of the procedures increased, allowing in the 1980s to identify Single Nucleotide Polymorphisms (SNPs), differentially expressed genes and physical mapping [35]. Since the 1990s to the early 2000s several improvements, including robotics for the creation of ever-smaller spots containing the probes, fluorescent probes detection and the increasing knowledge of genomics due to the sequencing projects of the human genome and transcriptome, made it possible to develop the microarrays known to-

day [35]. The first microarray technology was released in 1995, this first array was generated by a homemade robot that released probes of cDNA derived from 48 *A. thaliana* genes. This first array permitted to study the expression pattern of a small sub-set of *A. thaliana* genes [36]. Advances in the density capability of spotting permitted to assay simultaneously tens of thousand mRNA transcripts and to cover all the human expressed coding genes [36]. Over time three main array production processes were developed. Spotted arrays is the first published method that allowed a high-density DNA array. It is based on robotic spotter that, using slotted pins, releases tiny drops of a DNA solution containing only one specific probe on a glass substrate coated with poly-lysine to allow a good binding substrate for DNA probes [35]. The second method is the *in-situ* synthesized arrays, it uses nucleotides blocked in 3' position with a photolabile protecting group, attached to the substrate, selectively degrading the protective groups it is possible to build custom probes in selected positions. The main advantage of using *in-situ* synthesized arrays is that only a small collection of reagents is needed to build custom array. The main disadvantages of the initial method were that for every increase of probe length, there were needed 4 photolithographic masks, one for each base (ACTG), resulting in the use of a large number of masks to fabricate single custom arrays. Moreover, the masks allowed the construction of relatively big spots, decreasing the number of assayed probes and genes. Later, the masks photolithographic masks were substituted with Digital Micromirror

Devices (DMD) to precisely direct the light beams, increasing the number of available spots and reducing the number of needed masks. This technology is patented by Affymetrix ®that produces the most common genomics microarray chips including the Affymetrix® Human Genome U133 Plus 2.0 Array that contain 1,300,000 unique oligonucleotide features and covers more than 47,000 human transcripts [35]. Another *in-situ* synthesized arrays production method is the so-called inkjet printing technology. Here, similarly to the spotted array, the jets move over the substrate releasing single drops of solution in specific positions. The main difference with spotted arrays is that each jet is filled with only one different nucleotide, allowing the precise *in-situ* synthesis of custom probes. The inkjet printing technology, licensed by Agilent Technologies® shares the same advantage of the Affymetrix chips, there are needed only a few of reagents, and to produce low volumes of custom chip only requires a different set of information [35]. The last method, licensed by Illumina®is based on synthesizing the DNA probes on polystyrene beads, each bead contains only one highly repeated specific probe. The chips are etched with several wells that contain one bead. This kind of arrays are extremely flexible allowing a large number of different beads on a single array, and to test the array before it uses [35]. Microarrays have paved the way for high-throughput genetic studies, allowing to assay entire transcriptomes or a huge number of genes from several samples[37][38][39][40][41][42]. Over years many different procedures were developed from microarray technology, here some examples:

1.4.1 Expression Arrays

One of the most common applications of microarrays is the gene expression analysis. The RNA extracted from the sample is converted in cDNA and labelled with a fluorophore, then the cDNA is hybridized to a genomic array. After washing the array from the excess of sample is possible to measure the fluorescence at each spot and obtain the expression levels for each probes gene[35]. Due to the extensive amount of available gene expression data and the related profound analysis, it was possible to develop numerous diagnostic test based on expression patterns of the patient. The expression-based clinical test can be useful to classify, distinct and select patients for specific treatments. Some clinical test available are the MammaPrint®(Agendia®), used as a prognostic test to estimate the breast cancer recurrence, Decipher®(GenomeDx®) used to predict the recurrence of prostate cancer in postoperative follow-up and ThyraMIR^{TM} (Interpace Diagnostics Group), a microRNA classifier to evaluate if thyroid nodules are benign or malignant and precisely target surgeries [43].

1.4.2 Comparative Genomic Hybridization (CGH) Array

This approach derives from CGH analysis carried out on metaphase chromosomes. This method was based on specific labelled probes annealed to the chromosomes that were used to represent the genome. The fluorescence measured along the chromosomes provides information about copy

number changes. The main disadvantages of using CGH directly on chromosomes are the low resolution of 20Mb and the difficulties in obtaining good chromosome metaphase spreads [36][43]. The Array-CGH (aCGH) significantly increased the resolution of comparative hybridization up to 1 kb [43], here the genomic DNA from the test and control sample are labelled with different colour and hybridized on the array. Colour differences of the spots indicate copy number variant, this method was used to map several genetic abnormalities of cancers and other chromosome changes in other genetic diseases [43]. Array-CGH suffers from some limitations: it can detect only unbalanced copy variants (losses or gains), while all the balanced rearrangements (including translocations and inversions) cannot be detected by the method because the resulting number of genes is the same of the control [43]. Also, aCGH provides information only on the number of copy number changes rather than giving information about chromosomes aberrations [43]. Repetitive DNA regions including satellites and telomeres should be omitted from the analysis to reduce the background noise [43].

1.4.3 ChIP-chip

Chromatin Immuno-Precipitation on chip (ChIP-chip) is a technique that allows determining the binding sites of transcription factors (TF). To perform this type of analysis, before fragmentation, DNA is cross-linked with the proteins bound on it with formaldehyde. After the fragmentation, the

TFs with the DNA bound are purified by affinity (Antibodies, FLAGs, HIS-tags). After the purification, TF and DNA are separated, then the DNA is labelled and hybridized to the array to highlight the genomic regions with regulatory properties. This method allowed to elucidate several features of regulatory and intergenic regions both in prokaryotes and eukaryotes[35].

1.4.4 SNP arrays

SNPs are the most common variations in genomes and can be present 10 million SNPs in one human genome. SNPs are used as markers for cancer, genetic diseases and susceptibility loci. Since the high number of SNPs and their function as biomarkers, several microarray approaches were developed to exert genotyping, pattern identification and risk association analyses. Once a set of SNPs inherent to a specific condition are mapped on the genome, or some specific SNPs are associated with cancer, it is possible to build a custom array to profile the patient and guide the therapies in a more personalized way [43]. The most used SNP array technologies derive from Affymetrix and Illumina tests rely on the identification of the SNPs using PCR or DNA elongation and the measurement is based on fluorescence. These technologies can detect more than 1 million of human SNPs with high reproducibility (more than 99.5%) [35].

1.5 Protein-Protein Interaction (PPI) Networks

According to the central dogma of biology, proteins, derived from the translation of RNA, act as main effectors in biological systems. Each protein, exert its function by physically interacting with other biological components including other proteins, lipids, DNA and RNA. Thus, protein-protein interaction mediates several cellular functions including signals transduction, regulation of gene expression, and response to environmental changes [44]. Thus, in the last few decades, uncovering the intricate interactome became a primary target to understand several biological and medical points of interests. Firstly, starting in the 1930s, several findings permitted identification of proteins as the ”effectors” of the cells. Later, during the second half of the 20th-century technologies such as affinity purifications, yeast two hybrid, mass spectrometry and high-throughput analysis, significantly increased the knowledge about protein interactions and roles in the organization of cellular systems [44]. Finally, the advent of the genomic revolution provided by sequencing, NGS technologies and significant projects such as the Human Genome Project, supplied a vast amount of data, including which gene encodes each protein and high-throughput gene expression regulation data. This big-data linked with the interactions provided by mass spectrometry, and gave birth to interactome networks. In 2000 this was released as the first protein-protein interaction map, followed by other large scale studies. These global interconnectivity networks were

studied with bioinformatics approaches and contributed important insights into the organization of biological systems[44]. Two class of nodes (proteins) were definied, those that share few edges (interactions) and the other, highly connected nodes that play a central role in regulation and disease development, called hubs. During the last two decades, PPI networks have found utility in several fields, including the classification of tumours, medical target identification and development of modulator of gene expression [44]. PPI networks already have contributed in numerous biological and medical topics. Moreover, the increasing knowledge of genetic regulation and refinement in high-throughput analysis will provide further interest in and utility for PPI network analysis.

Chapter 2

AIMS OF THE STUDY

Aspergillus sp fungi and aflatoxins foods and feeds contamination is a rising concern due to climate change. It is estimated the increasing of aflatoxin contaminations from low to medium European countries with an increase of 2°C, a fact also reported by recent *A. flavus* outbreaks and increased Aflatoxin B1 levels in maize found in Serbia and Italy [2]. Also, evidence suggests that climate change factors as the increase of temperature, CO_2 levels in the air and water stress (a_W) can cause a switch from fumonisins to aflatoxins contamination and significantly increase gene expression of key regulators of Aflatoxin B1 production (*aflD* and *aflR*). At 34°C the expression of both the regulators doubled, and at 37°C, the expression levels increased from 15 to 79.2 fold change [12].Moreover, Aflatoxin B1 mechanism of action is not completely elucidated, mainly regarding oxidative stress, metabolic and synthesis pathways [3]. Additionally, evidence regarding possible endocrine disrupting properties of the toxin were provided by the ability of the selective estrogen receptor modulator Tamoxifen to prevent AFB1 interacting with enzymes involved in placenta

hormone production (CYP19A1) and decrease accumulation of its metabolite Aflatoxicol [5]. The literature reported ambiguous results of estradiol (EST) on the development of Hepatocellular Carcinoma (HCC), suggesting a protective role for the hormone at physiological exposures [45] but deregulation of the hormone levels and exposure to other synthetic estrogens such as ethinylestradiol (EE2) resulted in increased chance for HCC development [20] [46].

We hypothesise that prolonged exposures to Aflatoxin B1 and its metabolites can directly interact with nuclear and membrane Estrogen Receptors (ERs) α and β, impairing gene expression and pathways related to this class of receptors. We used an integrative system biology approach that considered structural information from the receptors and aflatoxins, gene expression data and protein-protein interaction (PPI) network analysis to pursue the following objective and test the hypothesis (Figure 2.1):

- Molecular Docking simulation to estimate the binding affinity of Aflatoxin B1 and its metabolite with ERs to evaluate the possibility of interaction;
- *In vitro* and *in vivo* Aflatoxin B1 and EE2 exposure differential gene expressions and pathways deregulation analysis to overlapping effects on gene expression;
- PPI network analysis to estimate the involvement of Downstream Estrogen Receptors Targets (DERTs) in Aflatoxin exposure.

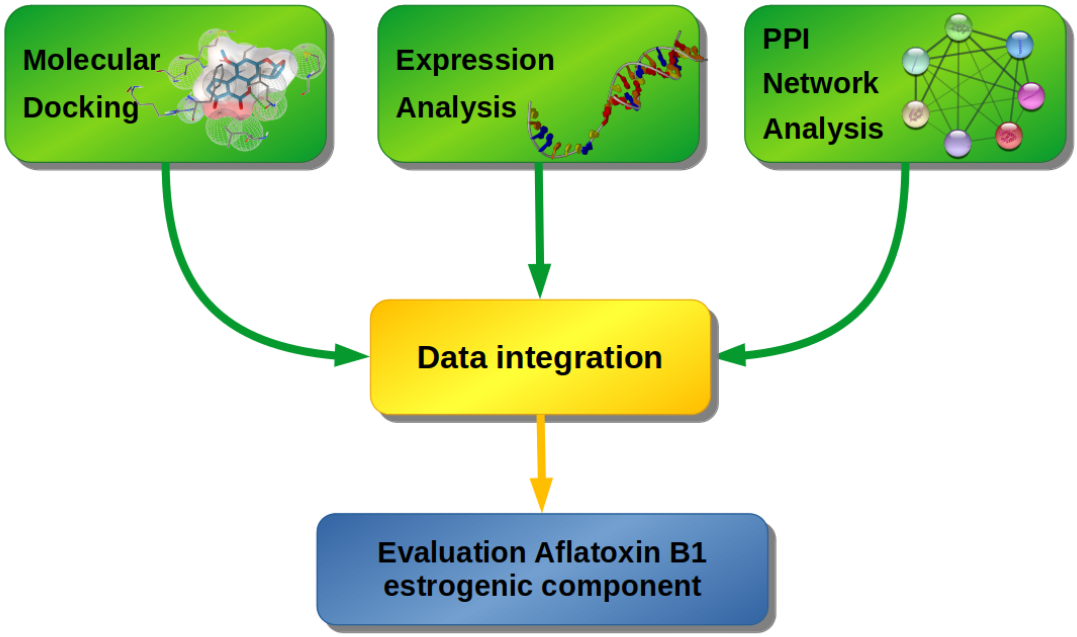

Figure 2.1: Pipeline to extract the estrogenic component of the Aflatoxin B1 mechanism of action.

Chapter 3

MATERIALS and METHODS

3.1 Molecular Docking

3.1.1 Ligands

Ligand selection was guided by the literature [1] [2] [4] [5]; Compounds deriving from first step of metabolization through oxidation (Cytochrome P450 enzymes) and reduction were considered. Aflatoxin B1 (AFB1), Aflatoxin B1-epoxide (Epo-AFB1), Aflatoxin M1 (AFM1) and Aflatoxin Q1 (AFQ1) have been selected as first step metabolites from oxidative metabolism, Aflatoxicol (AFL) has been selected from reductive metabolism of Aflatoxin B1. Zearalenone (ZEA), due its already known xenoestrogenic properties [14], and Ethinylestradiol (EE2) were included as positive controls. Ultimately Rifampin (RIF) was selected as a negative control due its different mass and chemical properties from previous compounds. From PubChem [47] 3D and 2D (only for RIF) structures were downloaded in SDF format. 3D structure generation, Hydrogen adding and conversion in

PDBQT file format has been performed using Open Babel chemical toolbox (version 2.4.1) [48]. Figure 4.1 shows molecular structures of selected compounds.

3.1.2 Receptors

From Protein Data Bank (PDB) molecular structures of Human Estrogen Receptor α (ERα) [49], ID:1ERE, and Human Estrogen Receptor β (ERβ) [50], ID:2NV7, were downloaded as templates for docking simulations. Receptors preparation, consisted of the selection of chain A, hydrogen addition and conversion to PDBQR file format. This was performed using UCSF Chimera (version 1.14) [51] and AutoDockTools(ADT) (version 1.5.6) [30].

3.1.3 Simulations

The two most widely used docking programs were used to run molecular docking simulations, AutoDockVina (version 1.1.2) [32] and AutoDock4 (version 4.2.6)[30]. The two programs use different scoring function, AutoDock4 is based on AMBER force fields and Vina uses hybrid function based on X-score function [33] (considering VdW interactions, deformation penality, H-Bond and hydrophobic effect) implemented with empirical and knowledge based parameters. The two programs also differ in termis of simplicity of parametrization, AutoDock4 requires algorithm settings, number of run and energy evaluation, Vina require setting on exhaustiveness value

and amount of poses to be recorded. Also, Vina is significantly faster due to the parallel computation optimization and outputs results with shorter simulation times, which is particularly useful studying datasets with large numbers of ligands. Despite these foundamental differences, it is not clear which one of the two programs is more accurate and even for similar proteins the precision between the two scoring functions can vary signigicantly [52]; Additionally, binding affinities and poses resulting from the simulations could be different among the functions [53]. Thus, to better validate the affinities and estimated constants of the dissociations we decided to use both programs and compare the results. Every previously selected ligand has been docked to ERα and ERβ with AutoDockVina (version 1.1.2) [32] and AutoDock4 (version 4.2.6)[30]. AutoDockVina grid box, centred on the binding site of each receptor, and was set as a cube with a side of 24 Å, an exhaustiveness value of 200. Poses for the 100 lowest energies were collected. In AutoDock4 simulations the cubical grid box, centred as AutoDockVina parametrization, was set with a side of 22.5 Å, 60 points $\times$ 0.375 Å spacing. AutoDock4 was run using the Lamarckian Genetic Algorithm (LGA) for 200 trials with a maximum of 1,000,000 energy evaluations. Poses from the cluster with higher populations and best affinity were considered. The equation 3.1 used to calculate the constant of dissociation (K_D) from binding affinity (Kcal/mol) implemented in

AutoDock4 was used to estimate K_D form AutoDockVina results.

$$K_D(M) = e^{\frac{\Delta G}{RT}} \tag{3.1}$$

ΔG = Binding Affinity (Kcal/mol)

R = Gas Constant 1.987×10^{-3}

T = 298.15 °K

Interactions

The best poses deriving from the program which obtained better results for positive controls, thus were considered more precise for simulating the bindings of cyclic compounds on ERs, and were used to generate structural models containing receptor and docked ligands using ADT and Chimera. The structural models were processed with *LigPlot*$^+$ (version 2.1) [54] to highlight interacting residues and H-Bonds among ligands and receptors.

3.2 Expression analysis

3.2.1 *In vitro* samples

Human hepatocytes were cultured and treated as described previously [55], in Open TG-GATEs (https://toxico.nibiohn.go.jp/english/index.html) CEL files concerning 24 hours AFB1 and EE2 exposures at 0.8 and 3 μM, respectively were downloaded. A total of 6 samples were obtained for every

compound, 3 treated samples and 3 control samples exposed only to the vehicle, DMSO for Aflatoxin B1 exposure and corn oil for ethinylestradiol exposure.

3.2.2 *In vivo* samples

Rat *(Rattus norvegicus) in vivo* samples were obtained from Open TG-GATEs and GEO datasets (www.ncbi.nlm.nih.gov/gds):
EE2 *in vivo* treatment in Open TG-GATEs was as previously described [55]. Samples from 6 and 9 hours for 3 mg/kg exposure were considered for downstream analysis.
AFB1 *in vivo* sample datasets were obtained from NCBI GEO (Experiment ID: GSE57815) [56], samples obtained after 3 and 5 days of 0.3 mg/kg bw$\times$day. A total of 6 samples were obtained for every compound, 3 treated samples and 3 control samples exposed only to the vehicle, carboxymethyl cellulose for Aflatoxin B1 oral gavage and corn oil for ethinylestradiol oral gavage.

3.2.3 Microarray analysis

All CEL files obtained from public databases were analysed using R (version 3.6.1) [57], R studio (version 1.1.463) [58] and packages present in Bioconductor (version 3.11) [59] [60]. CEL file upload, pre-processing and normalization was done using *affy* (version 1.66.0) [61], Differential Expression analysis from normalized expression was done using *limma*

(version 3.44.3) [62] and Biobase (version 2.48.0) [63] from Bioconductor (www.bioconductor.org). Human hepatocytes samples were analysed with gene chip Affymetrix-Affymetrix® Human Genome U133 Plus 2.0 Array and *In vivo* samples were analysed with gene chip Affymetrix® Rat Genome 230 2.0 Array for both the exposures (AFB1 and EE2). Probe ID to Gene Symbols lists for both the chips were obtained from Affymetrix® website. Microarray files from hepatocytes were directly processed using the human gene list from Affymetrix website; Rat *in vivo* microarray data, before being processed with Bioconductor packages, were appended with human gene IDs using the orthology search (g:Orth) function on g:Profiler [64] to link probe IDs to human gene symbols. This resulted in fewer gene symbols in the limma output but the human gene annotation is more precise and was selected to enable more accurate downstream analyses including Gene ontology (GO) and pathway level analysis. In an effort to select a considerable number of Differentially Expressed (DE) genes and keep the list as statistically significant as possible the outputs were filtered by three different settings of adj.p-value, or q-value (0.05, 0.1 and 0.4 respectively). The q-value corresponds to the Benjamin & Hochberg correction of p-value or False Discovery Rate (FDR), this procedure consists in the correction of p-values by population (genes) size. We selected the most stringent FDR value that provides informative numbers of genes for downstream systems analysis.

3.2.4 Gene Ontology analysis

Gene ontology (GO) is the most comprehensive resource concerning gene functions. Released in 1998, in more than 20 years it has defined more than 45,000 terms that cover three aspects of gene functions: biological processes, molecular functions and cellular components. The GO terms (classes) are linked together with specific relationships (134,000), that can be represented as a graph of different terms that are or are part of other classes. The same relationships can be expressed with logical definitions, it makes GO annotations available to computational analysis. The support for computational approaches made GO one of the most used tools in biomedical research [65]. GO annotation, Biological Process (BP) enrichment and meta-analysis of differential gene expression output from limma was analysed using iPathwayGuide (Advaita Bio) (https://www.advaitabio.com/ipathwayguide) [66]. The *in vitro* samples were filtered using a q-value threshold of 0.05 and an absolute log2FC value of 0.6. The *in vivo* samples were filtered using a q-value of 0.4, this less stringent threshold was used to provide greater number of genes from all the exposures for system level analyses. GO BPs were considered after Elim pruning [67], Elim pruning incorporated Bonferroni correction with a filtering method based on GO term redundancy, allowing more precise results than pure statistical methods such as FDR alone. We will refer to Elim pruned GO terms with adj.p-value statistically significance.

3.3 Protein-Protein Interaction (PPI) networks

With the aim of performing protein-protein interaction network analysis, isolating the most central nodes and comparing them with known downstream estrogen receptor targets (DERTs) was the direction pursued. This allowed the estimation of magnitude of the estrogenic component of the AFB1 mechanism of action.

3.3.1 Network construction

In an effort to highlight estrogen receptors involvement in modulating gene expression from AFB1 exposures, PPI networks analysis was explored. The gene lists obtained from the AFB1 treatments used in the GO analysis (*in vitro* AFB1 0.8 μM: genes filtered by q-value threshold of 0.5 and log2FC of 0.6, *in vivo* AFB1 0.3 mg/kg bw$\times$day: genes filtered by q-value threshold of 0.4) were added with ERα (*ESR1*), ERβ (*ESR2*) and G coupled-protein Estrogen Receptor 1 (*GPER1*) gene symbols. The gene lists were inputted to the String Database [68] to construct PPI networks based on a "combined score" deriving from text-mining, experiments, curated databases (e.g. KEGG), co-expression, neighbourhood, gene fusion and co-occurrence evidences. The networks obtained were subsequently imported into Cytoscape (version 3.8.0) [69] for downstream analysis.

3.3.2 Identification of Downstream Estrogen Receptor Targets (DERT)

The list of DERT was obtained using the ”consensome” function from Ominer on https://signalingpathways.org/ [70]. The function “consensome” works in a large-scale meta-analysis of transcriptomics datasets available in different databases (e.g. GEO) involving deregulation of specific pathways. The differential expression analysis compares logFC and p-values of target genes among the experiments to obtain a ranked list of genes most consistently impacted by a given pathway and bearing features such as discovery rate, consensus p-value and percentile. The research was focused on the consensome linked to estrogen receptors signalling pathway. Targets on the 90th percentile (frequency of differential expression of a given target relative to others in a specific consensome) were selected to maintain a high confidence level. The obtained list of 2,184 HGNC terms was considered as the total amount of most impacted DERTs. To facilitate the comparison with the proteins in the PPI networks built with AFB1 exposures constructed with STRING, the DERT list was inputted in g:Convert of g:Profiler [64] to convert the HGNC terms to Ensemble protein IDs (ENSPs) that resulted in a list of 12,365 redundant ENSP. For every network we considered only the DERT effectively present in the network, this step removed redundancy and focused the attention only on gene expression resulting from AFB1 exposures.

3.3.3 Network analysis

Essential proteins are the proteins whose removal can cause lethal defects [71]. To highlight the most important nodes in the networks the procedure for identification of essential proteins previously described was modified [72]. Essentially the list of essential proteins was substituted with the DERT list, allowing the comparison of estrogen receptor targets with the core nodes of the network. Network analysis was done using Cytoscape (version 3.8.0) [69] and CytoNCA (version 2.1.6) [72]. This enabled analysis and calculation of several kinds of centrality and compared the network with a list of selected proteins. The centrality measures selected to highlight key nodes of every network were:

- 'Betweenness centrality' a measure of the shortest paths passing through each node, that can highlight nodes lying on communication paths and can control information flow.

- 'Closeness centrality' the average length of the shortest paths between the node and all other nodes in the graph. It highlights nodes that are close to all the other nodes as a measure of how fast the flow of information is through this node.

- 'Subgraph centrality' characterizes the participation of each node in all subgraphs in a network as a measure of the ability of the node to link separate motif and clusters of the network.

After having calculated the different centrality measures the top central nodes for every centrality were selected in the same amount as those DERTs present in the network. To test the independency of the frequency of DERTs in the top nodes from the total amount of DERTs in the complete network Pearson's chi-squared test and post hoc testing corrected by FDR were carried out using R (version 3.6.1) [57], R studio (version 1.1.463) [58] and the package *rmngb* (version 0.6-1) [73]. The overlapping nodes among the centrality measures were used to build a new subnetwork and compared with the DERTs to find the estrogen receptors targets and their direct neighbours and observe the amount of the estrogenic component in the core network.

3.4 Graphs and Figures

Graphics and figures were generated using the Moderndive R package (version 0.4.0) [74], the GNU Image Manipulator Program (version 2.10.18) [75], ADT (version 1.5.6) [30] and Cytoscape (version 3.8.0) [69].

Chapter 4

RESULTS

4.1 Molecular Docking

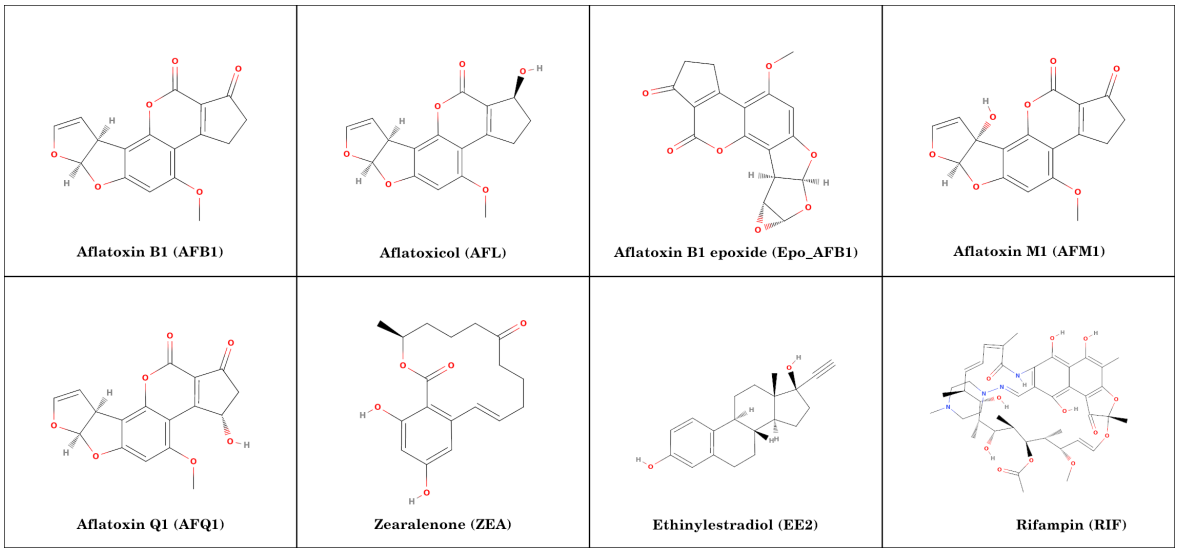

Figure 4.1: Molecular structure of the compounds used for Molecular Docking simulations. Is possible to observe the presence of an aromatic ring in every compound, apart from ZEA and RIF all the compounds are composed of four to five linked rings, a feature that increases the structures rigidity. Moreover, the presence of carbons in sp2 hybridization, in addition to the aromatic ring and the closed conformation suggests a possible planar conformation of AFB1 and its metabolites. Mycotoxins molecular weights range from 312.3 to 328.3 g/mol. EE2 molecular weight is equal to 296.4 g/mol. RIF, the negative control's molecular weight is equal to 822.9 g/mol. Structure images were obtained with http://molview.org/.

4.1.1 AutoDockVina

ERα Vina docking simulations on the ERα showed high affinity of EE2 for the receptor (-10 Kcal/mol) and an estimated constant of dissociation in the order of nanomolar (50nM). The second positive control ZEA demonstrated lower affinity (-8.1 Kcal/mol) leading to an estimated K_D of 1.15 μM, 23 times higher than EE2. Aflatoxin affinities ranged from -7.4 to -8 Kcal/mol (AFB1: -8, AFL: -7.4, Epo-AFB1: -7.6, AFM1: -7.8 and AFQ1: -7.9) suggesting active concentrations on the order of μmol. The similar affinities among AFB1 and ZEA, and an estimated K_D that differs only in 0.22 μM (AFB1: 1.37 μM, ZEA: 1.15 μM), suggests the possibility of interaction with ERα at similar concentrations, implying that the biological effects of AFB1 contamination could have an estrogenic component. RIF included as a negative control, had a binding energy of 63.9 Kcal/mol, making unlikely the possibility of effective interactions with the receptor (Table 4.1).

ERβ Simulations with ERβ revealed lower average affinities, EE2 included as the first positive control bound the receptor with a binding affinity equal to -8.3 Kcal/mol, increasing the estimated K_D to 0.82 μM. ZEA included as the second positive control showed a poor affinity value (-4.4 Kcal/mol), leading to a K_D that make improbable efficient interaction with the receptor in biological conditions (594.9 μM). This result suggests that ZEA acts only through ERα or that Vina was not able to correctly estimate the binding energies with ERβ. Excluding AFL (-7.7 Kcal/mol), all

other aflatoxins showed worse affinities than with ERα (AFB1: -7.7, Epo-AFB1: -6.9, AFM1: -7.3 and AFQ1: -7.6 Kcal/mol) leading to higher K_Ds. Aflatoxins with higher affinity were AFB1 and AFL with the K_Ds equal to 2.27 μM, a value lower than ZEA K_D, but significantly higher than the ERα results. RIF, similarly with ERα resulted in binding the ER receptor with a poor affinity (Table 4.1).

The significantly lower affinities for ERβ among all the ligands, especially the positive control ZEA, raised the concern that the Vina scoring function could have approximated results with low precision. Forward evaluations with AutoDock4 were used to provide clarity and suggest which results could be considered for downstream analysis.

	ERα		**ERβ**	
	Binding Affinity (Kcal/mol)	**K_D (μM)**	**Binding Affinity (Kcal/mol)**	**K_D (μM)**
EE2	-10	0.05	-8.3	0.82
AFB1	-8	1.37	-7.7	2.27
AFL	-7.4	3.76	-7.7	2.27
Epo-AFB1	-7.6	2.68	-6.9	8.74
AFM1	-7.8	1.91	-7.3	4.45
AFQ1	-7.9	1.62	-7.6	2.68
ZEA	-8.1	1.15	-4.4	594.90
RIF	63.9		70.6	

Table 4.1: Results from AutoDockVina simulations, constants of dissociation (K_D) were obtained from binding affinity using equation 3.1. ERα) EE2 showed significantly better affinity and K_D than other molecules, nevertheless AFB1 and AFQ1 showed similar affinity to the mycoestrogen, and positive control, ZEA; Also, the AFB1 calculated K_D results were only 0.22 μM higher than the mycoestrogen. ERβ) Also for Estrogen Receptor β EE2 showed better results than the other ligands, the second positive control ZEA was the worst ligand excluding the negative control. RIF as a negative control revealed for both receptors affinity values too high to be considered effective interactions.

4.1.2 AutoDock4

ERα Simulations with ERα showed results like the Vina simulations on the same receptor. EE2 was the best ligand with a binding affinity of -10.98 Kcal/mol and a K_D of 10 nM. The second positive control ZEA yielded better results than the Vina simulations, with a binding affinity of -9.6 Kcal/mol, thus a K_D on the order of nanomolar (90 nM). Aflatoxins showed affinities lower than both the positive controls (AFB1: -7.99, AFL: -7.45, Epo-AFB1: -7.82, AFM1: -7.72 and AFQ1: -8.06) and keeping the K_Ds were like the previous analysis. The two aflatoxins that resulted binding the receptor with better affinities were AFB1 and AFQ1 with K_Ds of 1.38 and 1.25 μM respectively. Contrary to the Vina results, these data suggest that AFB1 contamination could have estrogenic components at higher concentrations than ZEA or long duration exposures, leading to the impairment of ERs after the accumulation of AFB1 and its metabolites. The RIF binding energy (340.71 Kcal/mol) made unlikely any interaction with the receptor (Table 4.2).

ERβ AutDock4 simulations on the ERβ showed greater affinities for the positive controls than Vina (EE2: -9.72, ZEA: -7.7 Kcal/mol). The EE2 results, the affinity and the resultant K_D of 70 nM suggested high activity at low hormone concentrations. ZEA, despite an affinity lower than with ERα, had a resulting estimated K_D equal to 2.26 μM and suggests the possibility of interaction at low concentrations and indicates that AutoDock4 approximated more precisely the affinities for this receptor

than Vina. Also, for some aflatoxins affinities with ERβ were slightly better than with ERα: AFB1 Binding affinity of -8.1 Kcal/mol and K_D of 1.16 μM suggested active concentrations at lower levels than ZEA. AFQ1 with an affinity of -8.24 Kcal/mol was the only one aflatoxin with an estimated K_D lower than 1 μM (0.9 μM), this result gave one more suggestion of the possibility of an estrogenic component to aflatoxins contamination. The principal genotoxic metabolite of AFB1, Epo-AFB1, also resulted in a Binding affinity higher than ZEA (-7.83 Kcal/mol, K_D: 1.81 μM). Other aflatoxins resulted in lower affinities (AFL: -7.53, AFM1: -7.54) than ZEA and K_Ds higher than 2.9 μM. RIF included as a negative control yielded high binding energy and no possibility to effectively interact with the binding pocket of ERβ (Table 4.2).

More robust results for the positive controls resulting from AutoDock4 made it possible to consider this program more precise in affinity estimations for the ERs and cyclic compounds. Thus the set of poses from the AutoDock4 simulations were chosen for the interaction analysis.

	ERα		ERβ	
	Binding Affinity (Kcal/mol)	K_D (μM)	**Binding Affinity (Kcal/mol)**	K_D (μM)
EE2	-10.98	0.01	-9.72	0.07
AFB1	-7.99	1.38	-8.1	1.16
AFL	-7.45	3.48	-7.53	3.04
Epo-AFB1	-7.82	1.84	-7.83	1.81
AFM1	-7.72	2.18	-7.54	2.96
AFQ1	-8.06	1.25	-8.24	0.9
ZEA	-9.6	0.09	-7.7	2.26
RIF	340.71		540.44	

Table 4.2: Results from AutoDock4 simulations, binding affinities and K_Ds were obtained directly from the AutoDock4 output. ERα) The Vina results demonstrated that EE2 showed the best affinity and K_D among all the ligands, ZEA showed better affinity than the previous simulations; Aflatoxins showed results similar to the Vina simulations, AFB1 and AFQ1 are the best ligands among the aflatoxins. ERβ) EE2 showed the best affinity and K_D values, ZEA showed significant better affinity than Vina results but still lower than AFB1 and AFQ1. AFQ1 was only one aflatoxin that showed an estimate K_D lower than μM concentration. RIF affinities were too high to consider probable interactions with both the receptors.

4.1.3 Poses and interactions

ERα Observing the binding poses with ADT it is possible to observe that all the ligands entered the binding site of ERα. Focusing the attention on oxygens and positioning of the rings, I observed that AFB1, AFL, AFM1 and Epo-AFB1 shared similar binding poses. AFQ1 showed a pose more like EE2 than the other aflatoxins, especially the two six-carbon rings and *O5* of AFQ1 which appeared in a similar position to the first aromatic ring, the second ring and *O1* of EE2. Interestingly ZEA and EE2, the two positive controls do not appear to bind the receptor similarly, only *O1* and *O4* of ZEA showed positions close to *O1* of EE2 (Figure 4.2). Using *LigPlot*$^+$

on a structural model of ERα bound with ligands was possible to observe that interactions occurred. In the comparison among AFB1 and EE2 on ERα, *O4* of AFB1 and *O2* of EE2 could form a H-Bond with the same Arginine residue (*Arg349*). Also,*Glu353* can form an H-Bond with the *O2* of EE2 and weak interactions with AFB1. Several residues that interact with EE2 also interact with AFB1 excluding *Leu346, Thr347, Leu 349 and Phe404* for AFB1 and *Met421, Met424, Gly521 and His524* for EE2. Notably, *Gly521* forms a H-Bond with *O1* of EE2 improving the specificity of binding (Figure 4.3). Excluding AFQ1, the other aflatoxins bound the receptor similarly to AFB1 (not shown, images in appendix A). Interestingly, observing the comparison among AFQ1 and EE2, despite the good affinity value of AFBQ1 for ERα and the presence of two H-Bonds, one formed among *O5* and *His524* and the other among *O6* and *Leu346*, no one appears to occur with the same residues that form H-Bonds with EE2. Nevertheless, it is possible to observe that two of the residues that forms H-Bonds with EE2 appear to form weak interactions with AFQ1 (*Glu353 and Gly 521*). Additionally, several residues that form weak interactions are shared between the two toxins (Figure 4.4). These results suggest that AFB1 could bind with good affinity and specificity with ERα; AFQ1 results revealed similar affinity to AFB1 but a quite different mechanism to bind the receptor. Taken together one can to speculate that the two aflatoxins could bind at similar concentrations to the ERα but leading to different states of activation due to the different interaction specificity.

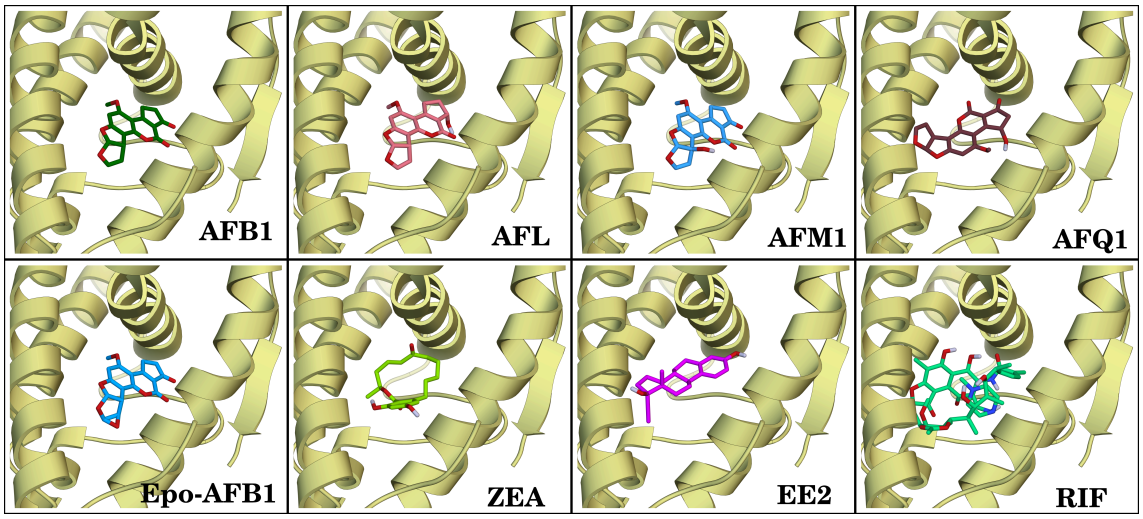

Figure 4.2: Binding poses on ERα after AutoDock4 simulations, obtained with ADT. The 11th α-helix was hidden to present a clearer visualization of the poses. It can be observed that every ligand entered the binding site; Aflatoxins showed a similar binding pose among them excluding AFQ1, which also showed higher affinity of ERα and resulted in a pose more similar to EE2. Despite the high affinity of both EE2 and ZEA, the two positive controls, for ERα they do not share a similar binding pose. Also RIF, the negative control, entered the binding site, suggesting this ligand has low affinity for ERα.

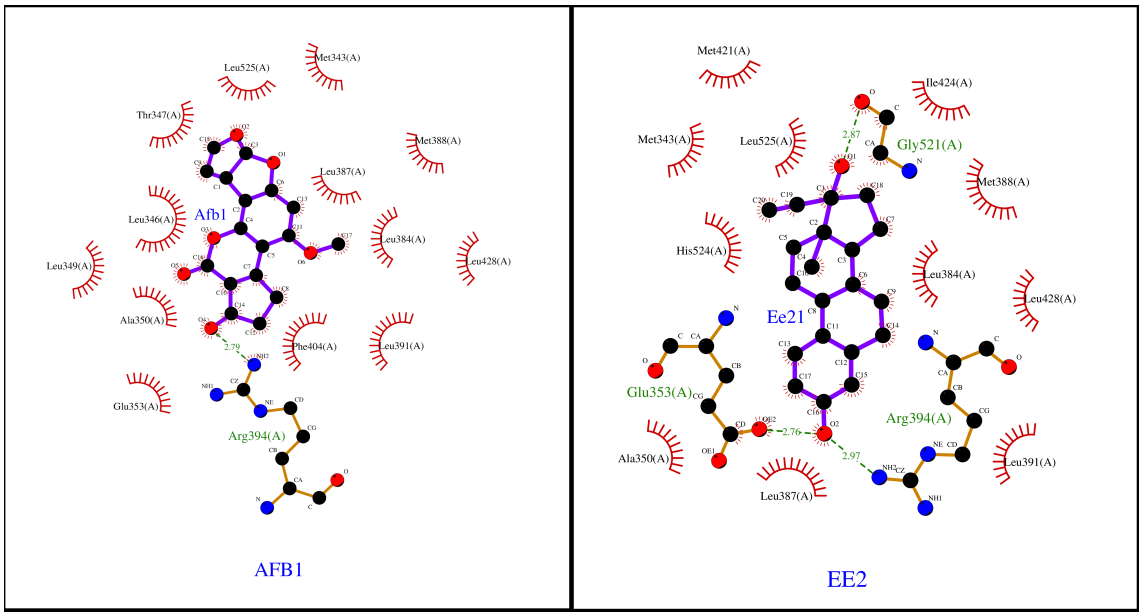

Figure 4.3: Comparison of the interactions occurred among AFB1 and EE2 with ERα residues obtained with *LigPlot*$^{+}$. Notably both the ligands share a H-bond and several weak interactions with the same residues. Excluding AFQ1, other aflatoxins interact with ERα similarly to AFB1 (not shown, images in appendix A).

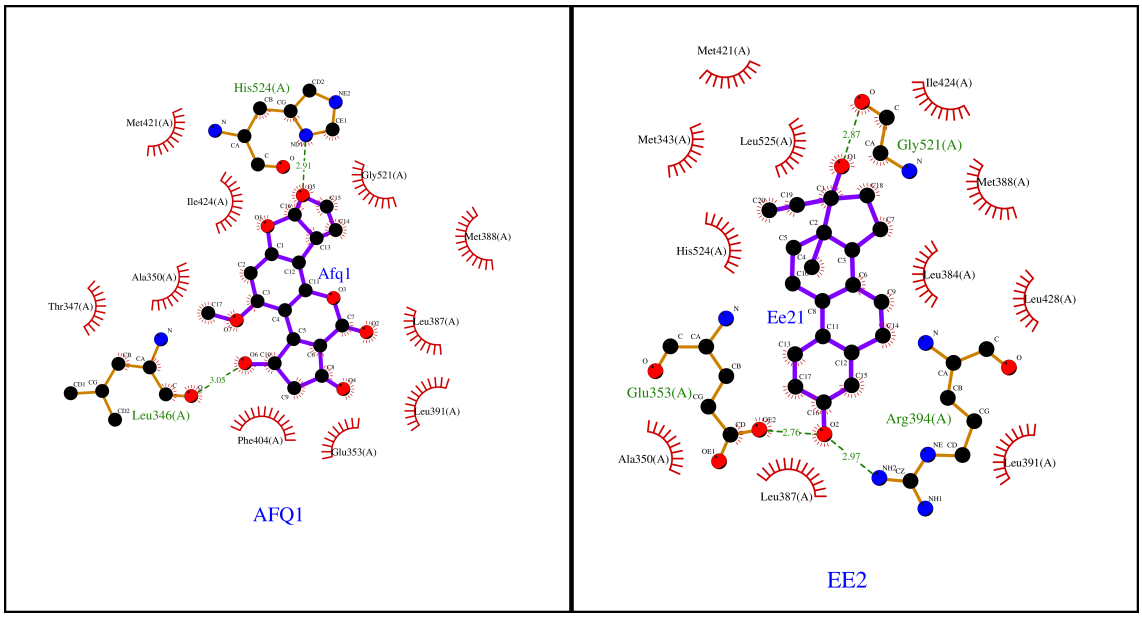

Figure 4.4: Comparison of the interactions occurred among AFQ1 and EE2 with ERα residues obtained with *LigPlot*$^{+}$. AFQ1 appears to interact with ERα establishing two H-Bonds with different residues than EE2. The smaller number of shared H-Bonds but quite good affinity suggests a different way for AFQ1 to interact with ERα.

ERβ Binding poses in the ERβ binding pocket were less conserved among aflatoxins. AFB1 and AFM1 appear to share a very similar binding pose. Additionally AFQ1 aromatic rings appear positioned close to Epo-AFB1 ones. AFL appears to bind the receptor in an inverted mode compared to AFB1. Observing the aromatic ring of EE2 similarity was found close to the first aromatic ring of AFB1 and AFM1, *O4* of both the aflatoxins appears to place close to *O1* of EE2 and finally *O2* of AFB1 and *O5* of AFM1 appears close to *O2* of EE2 (Figure 4.5). Similarly to the ERα results, ZEA showed a different binding pose to EE2 except for the *O2* of ZEA and *O1* of EE2, a high degree of freedom due to the smaller number of intramolecular rings and sp2 hybridized carbons is observed in

ZEA. RIF enters the binding site but shows a low affinity for the receptor. Probing deeper in the study of interactions among AFB1 and EE2, despite the good affinity, no one H-Bond is established among the AFB1 and ERβ residues. Interestingly the two residues that establish H-Bonds with the *O2* of EE2, *Glu305 and Arg346*, appear to be involved in weak interactions with AFB1. Several other interacting residues are shared among the two substances, except for *Thr299* for AFB1 and *Gly472 and His 475* for EE2 (Figure 4.6). As observed in Figure 4.5, AFM1 interactions are like AFB1 (not shown, images in appendix A). In a similar fashion to ERα, AFQ1 can establish two H-Bonds with the ERβ, the first among *O5* of AFQ1 and *His475* and the second among *O6* and *Leu298*. As with the previous results, both do not appear to interact with the same residues that form H-Bonds with *O2* in EE2. Despite this difference both of the residues implied in H-Bonds with AFB1 could form weak interactions also with EE2 and several interacting residues are shared among the two substances, with the exception of *Thr299 and Leu 491* for AFQ1 and *Arg 346 and Phe377* for EE2 (Figure 4.7). Logically with ERα interactions, these results could suggest a different action of AFQ1 toward the ERβ activation state. It is also important to highlight that AFL, among *O5* and *Glu305*, is the only one aflatoxin that establishes an H-Bond with the same residues implied with EE2. Also, as suggested by similar positioning of Epo-AFB1 and AFQ1, epoxidic oxygen *O1* can make an H-Bond with *His475*, similarly to AFQ1 (not shown, images in appendix A).

Despite being a mycoestrogen and positive control ZEA showed high affinities, the interaction studies indicate that it does not appear to establish H-Bonds with ERs (not shown, images in Appendix A).

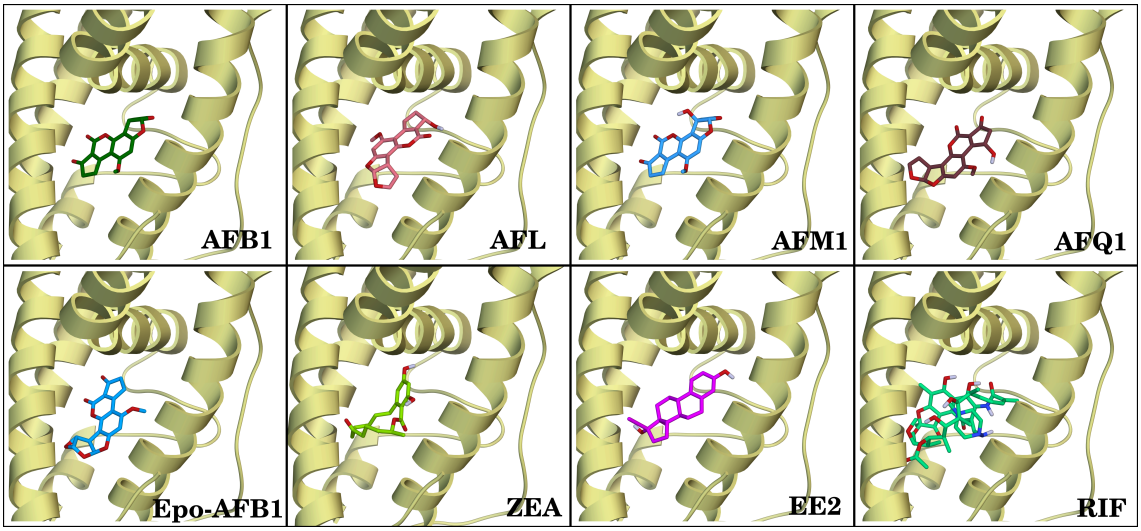

Figure 4.5: Binding poses on ERβ after AutoDock4 simulations, obtained with ADT. The 10th α-helix was hidden to provide clearer visualization of the poses. All the ligands entered the ERβ binding site, aflatoxins poses were less conserved than with ERα. AFB1 and AFM1 share similar binding poses, Epo-AFB1 and AFQ1 aromatic rings appear to be positioned in the receptor binding pocket in a very similar way. ZEA resulted in a different pose than EE2. RIF entered the binding site, suggesting low affinity for the receptor.

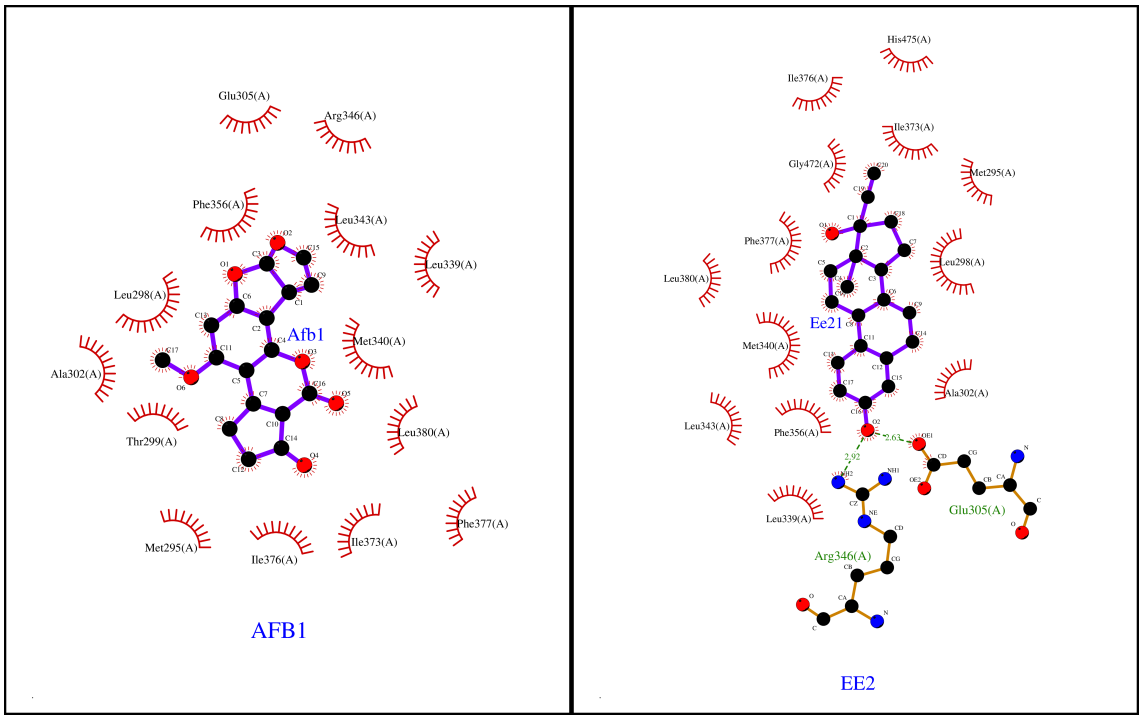

Figure 4.6: Comparison of the interactions occurred among AFB1 and EE2 with ERβ residues obtained with *LigPlot*$^{+}$. Despite the affinity no H-Bonds were established among the AFB1 and ERβ residues, similar interactions among AFB1 and AFM1 with ERβ residues are coherent with the binding poses in 4.5 (not shown, images in A).

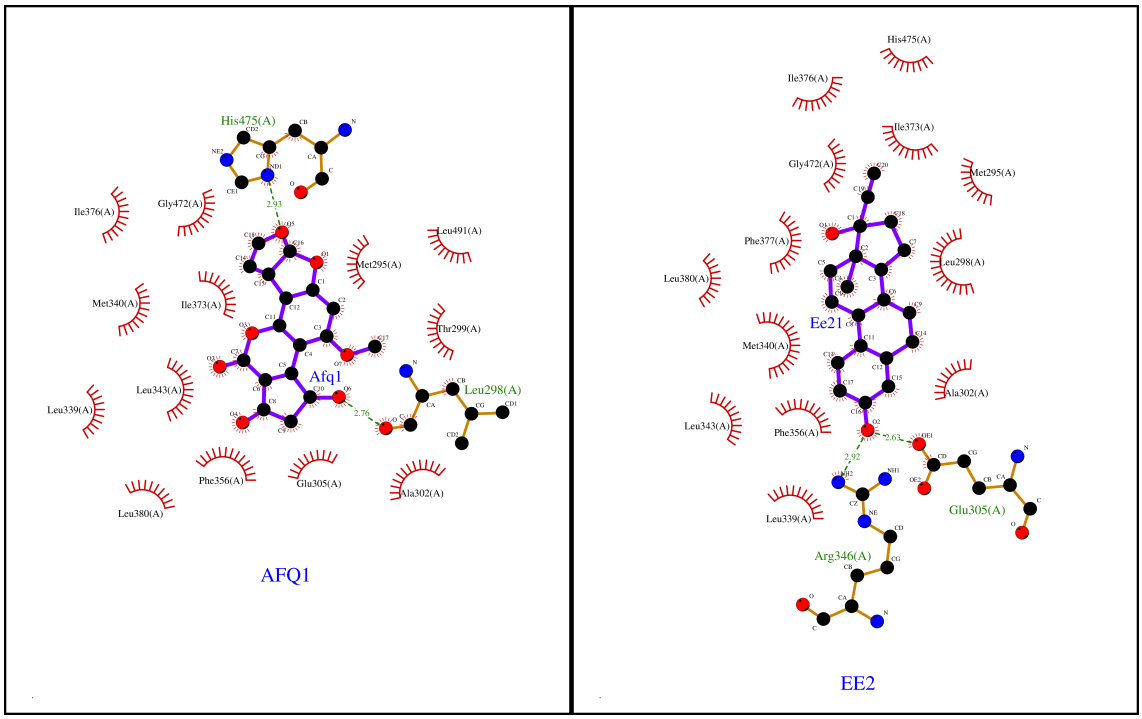

Figure 4.7: Comparison of the interactions occurred among AFQ1 and EE2 with ERβ residues obtained with *LigPlot*$^{+}$. Similarly to the ERα results AFQ1 appears to establish two H-Bonds with different residues than EE2, and when considered with the ERα results, these features potentially suggest a different state of activation of ERβ after AFQ1 binding.

4.2 Expression Analysis

4.2.1 DE genes partially overlap after *in vitro* exposures to Aflatoxin B1 and Ethinylestradiol

After expression analysis and FDR filtering it was observed that for both the AFB1 0.8 μM and EE2 3 μM exposures, a similar number of DE genes was uncovered when the same q-value threshold was used (Figure 4.8). A high number of genes were obtained with lower stringent FDR values,

0.4 FDR revealed 13,000 genes for both compounds (AFB1: 1,305, EE2: 13,294). Using an FDR value of 0.1 more than 7,000 genes were obtained (AFB1: 7,439, EE2 7,602). Also, using the most stringent threshold we obtained a consistent number of genes for both substances (AFB1: 5,931, EE2: 6,055). To consider only significant DE genes and reduce background noise the most stringent FDR was selected for downstream analysis (Table 4.3). In order to retain relevant biological information a log2FC equal or higher than 0.6 was selected as the threshold. Interestingly ERα (*ESR1*) lacked both the significance and a biologically relevant fold change value for both exposures to be considered in the downstream analysis; Also ERβ (*ESR2*) was not significant in the in AFB1 exposure and lacked a relevant fold change for both exposures to be considered in the downstream analysis (Table 4.4). To compare the number of DE genes resulting from both exposures a Venn Diagram with filtered genes was generated (Figure 4.9). The total amount of DE genes obtained from both exposures was 2,477, 1,349 for EE2 and 1128 for AFB1, of them 166 genes were shared by all exposures.

4.2.2 Gene ontology (GO) annotation of *in vitro* DE genes

In an effort to extract significant biological information, the gene lists with all the features present in the *limma* output were inputted into iPathwayGuide (Advaita Bio). The meta-analysis focused on Biological Processes (BP) was performed by filtering the terms using Elim pruning method

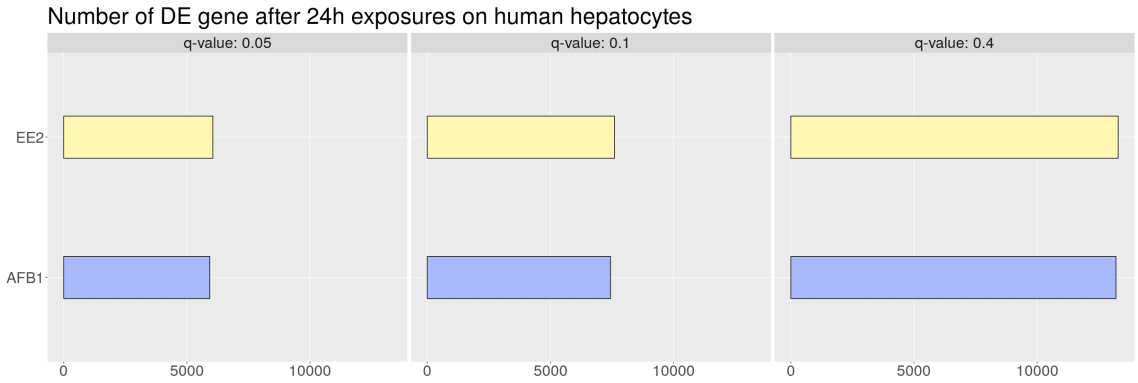

Figure 4.8: Graphic of Differentially Expressed genes after 24h exposure of AFB1 0.8 μM and EE2 3 μM to human hepatocytes. Similar amounts of DE genes resulted at similar q-values.

	DE genes		
q-value	0.05	0.1	0.4
AFB1	5,931	7,439	13,205
EE2	6,055	7,602	13,294

Table 4.3: Number of Differentially Expressed genes after 24h exposure of AFB1 and EE2 on human hepatocytes. Greater than 5,000 genes resulted from filtering using a q-value of < 0.05. The most stringent False Discovery Rate (FDR) vuale was utilized as a threshold for subsequent analysis.

	ERα		**ERβ**	
	Log2Fc	**q-value**	**Log2Fc**	**q-value**
AFB1	0.06	0.81	0.11	0.57
EE2	-0.13	0.51	0.29	0.02

Table 4.4: From the expression analysis after both exposures none of the ERs showed a significant Log2FC and/or q-values to be considered in GO analysis.

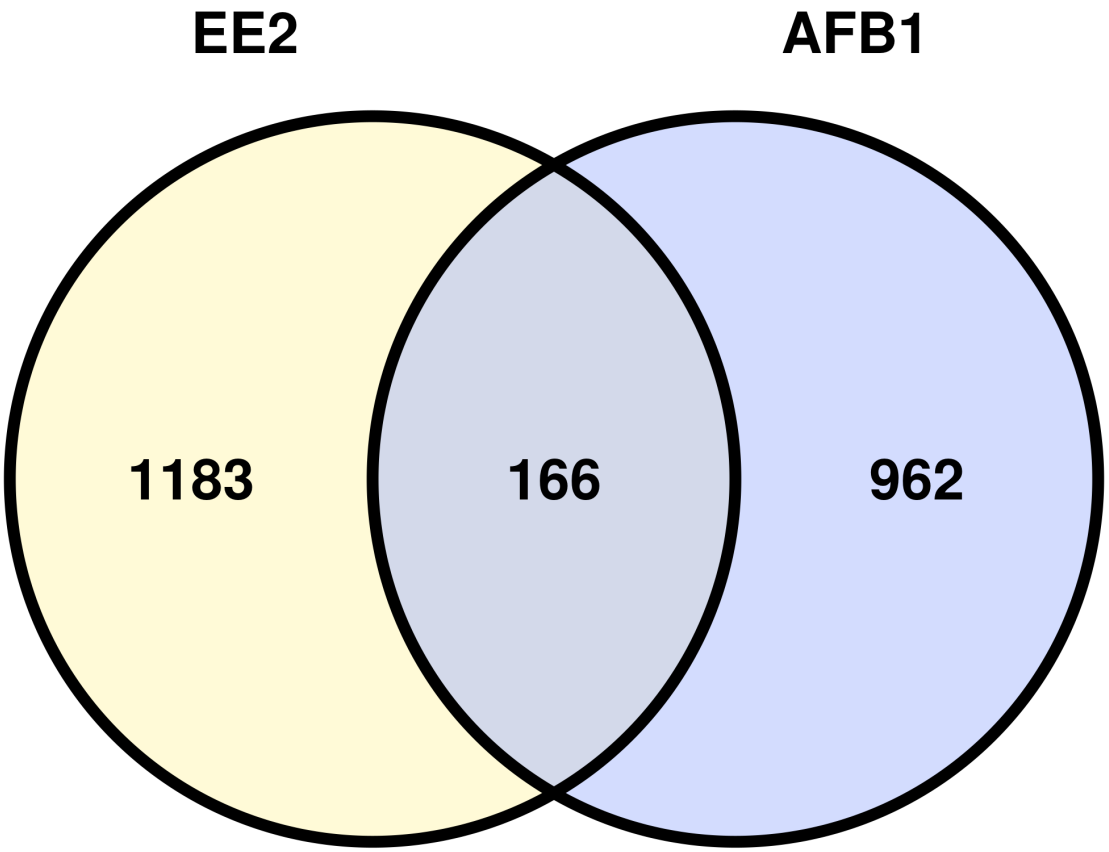

Figure 4.9: Overlap of DE genes after AFB1 0.8 μM and EE2 3 μM exposure on human hepatocytes, AFB1 and EE2 showed unique signatures of 962 and 1,183 genes respectively. A total of 166 genes were shared among the two exposures. Both gene lists were filtered using a Log2FC ≥ 0.6 and $q < 0.05$.

which includes Bonferroni p-value correction (adj.p-value). GO analysis resulted in a partial overlap of BP terms among the exposures. Unique signatures for the AFB1 and EE2 exposures corresponded to 198 and 390 GO terms respectively (not shown, Table B.1 in Appendix B), the overlapping terms are 24 (Figure 4.10). GO terms shared among the two substances (Table 4.5) revealed alterations in apoptotic and cytolytic processes also in response to DNA damage (6 BP terms), impaired equilibrium of differentiation and cell proliferation (6 BP terms) and upregulation of the immune response and inflammation(4 BP terms). Moreover, Translesion synthesis (adj.p-value 0.03584), a process responding to DNA damage and replication, indicates the possibility of mutations due to the error-prone DNA-pol involved. Interestingly Negative regulation of cell-cell adhesion mediated by cadherin (adj.p-value 0.03411), appears to be directly involved in the estrogenic response [76]. Adhesion deregulation appears to be strictly involved in tumour transformation and metastasis [77].

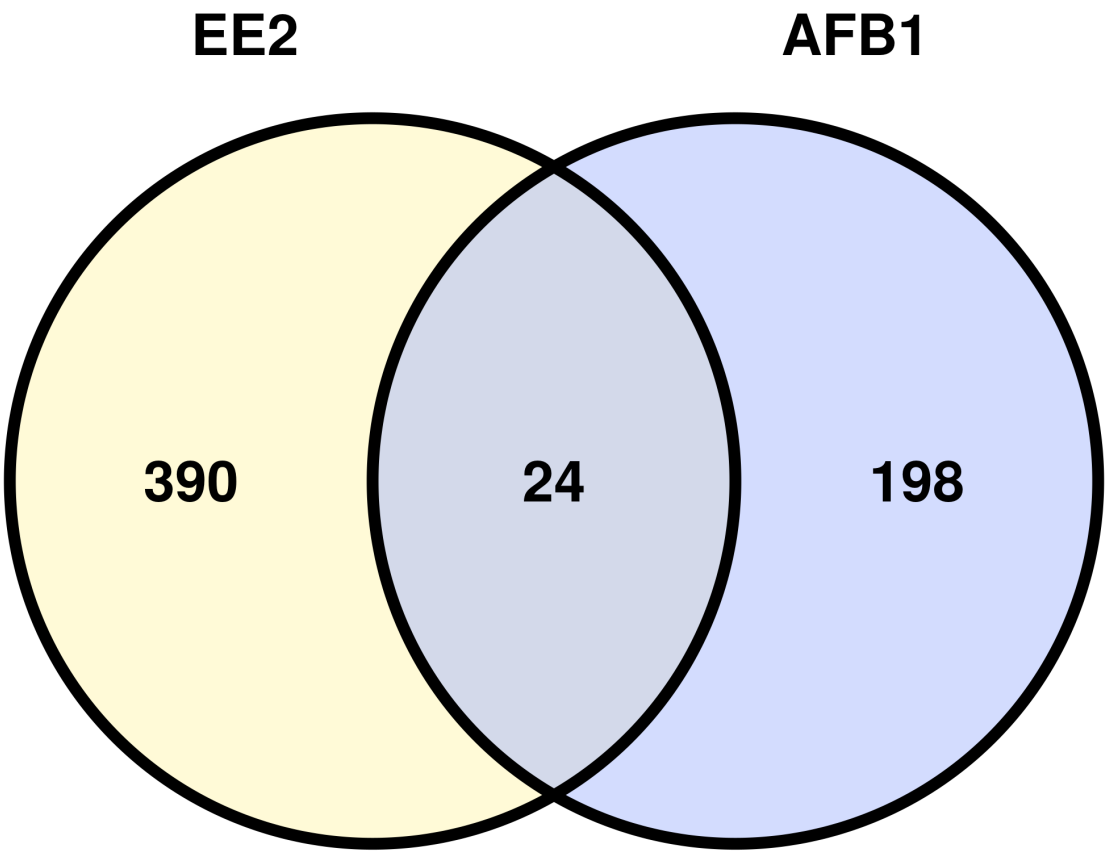

Figure 4.10: iPathway Guide (Advaita) meta-analysis among AFB1 and EE2 expressions indicates unique signatures for both substance and a partial overlap of 24 BP terms.

GO:Biological Process	AFB1 adj.p-value	EE2 adj.p-value
Amino-acid betaine catabolic process	0.01113	0.01653
Atrial septum primum morphogenesis	0.03411	0.0497
Atrioventricular valve morphogenesis	0.00292	0.03206
Choline catabolic process	0.04908	0.00741
Extrinsic apoptotic signaling pathway	0.03615	0.03544
Intrinsic apoptotic signaling pathway in response to dna damage	0.02196	0.03004
Negative regulation of apoptotic process	0.04496	0.02079
Negative regulation of cell proliferation	0.03034	0.03624
Negative regulation of cell-cell adhesion mediated by cadherin	0.03411	0.00393
Negative regulation of monocyte chemotaxis	0.03411	0.0497
Negative regulation of protein localization to microtubule	0.03411	0.0497
Negative regulation of protein phosphorylation	0.03719	0.00835
Negative regulation of transcription by transcription factor localization	0.01113	0.00044
Positive regulation of apoptotic process	0.01396	0.02678
Positive regulation of cytolysis	0.02721	0.04587
Positive regulation of i-kappab kinase/nf-kappab signaling	0.02958	0.02766
Positive regulation of neutrophil mediated killing of symbiont cell	0.00387	0.0058
Regulation of endodermal cell differentiation	0.03411	0.0497
Regulation of muscle cell differentiation	0.04243	0.03912
Regulation of neutrophil mediated killing of gram-negative bacterium	0.00387	0.0058
Regulation of sulfur amino acid metabolic process	0.01113	0.01653
Response to gamma radiation	0.0057	0.01989
Translesion synthesis	0.03584	0.00078
Traversing start control point of mitotic cell cycle	0.02134	0.03139

Table 4.5: BP shared by AFB1 and EE2 *in vitro* exposures from iPathway Guide (Advaita) meta-analysis. The presence of several terms implied in apoptotic, proliferative and inflammatory condition suggests common cancerogenic pathways leading from exposure of both substances. Also, Negative regulation of cell-cell adhesion mediated by cadherin, responding to ERs activity, csuggests estrogen receptors involvement in the AFB1 mechanism of action.

4.2.3 EE2 early exposure and AFB1 repeated exposures shows partial overlapping *in vivo*

In contrast to the *in vitro* experiments, expression analysis and FDR filtering yielded lower amounts of DE genes for AFB1 0.3 mg/kg bw$\times$day and EE2 3 mg/kg exposures. From the AFB1 exposures low numbers of DE genes were obtained with a 0.05 q-value filter (3-day: 30, 5-day: 0) and

0.1 (3-day: 141, 5-day: 0). Thus, in an effort to consider an amount of DE genes useful for system level GO analysis the 0.4 q-value threshold was selected obtaining 1,456 genes for 3-day exposure and 605 genes for the 5-day exposure (Figure 4.11 and Table 4.6). EE2 exposures showed higher numbers of significant DE genes, the 0.05 q-value threshold resulted in 1,717 genes for 6 hours exposure and 3,291 genes for 9 hours. The 0.1 cut-off threshold yielded 2,442 genes from the 6 hours exposure and 4,138 genes from the 9 hours exposure. Despite the higher numbers of genes with more stringent q-values the same threshold for AFB1 exposure was selected (Figure 4.12 and Table 4.7). Due to the lower number of DE genes obtained, we decided not to apply a Log2FC threshold. To compare the expression shared by different exposures a Venn Diagram was generated (Figure 4.13).

Within Comparison within expression patterns from the same substance show partial overlaps. AFB1 samples at 3- and 5-day shared only 106 genes of 1,898 DE genes obtained from 3- and 5-day AFB1 exposures. Also, unique signatures of the two different time points differed significantly, 852 genes for the 3-day exposure and 333 genes for the 5-day exposure, probably due to the different amount of DE genes obtained with q-value filtering. EE2 exposure at 6 and 9 hours showed partial overlaps, of the total of 3,653 DE genes obtained from both the EE2 exposures, only 1,212 were shared by the two-time points. Moreover, 997 of these 1,212 differentially expressed genes were not involved in the AFB1 expression

signature. Similarly, to the AFB1 expression, the EE2 unique signatures differed significantly, 372 genes were expressed only after exposure to EE2 after 6 hours and 1,636 after 9 hours.

Between Comparison between the two compounds shows that AFB1 at 3- and 5-day exposures differs for the amount of DE genes shared with the EE2 exposures. Expression to AFB1 after a 3-day exposure shared 243 differentially expressed genes with EE2 expression after 6 hours and 450 genes with EE2 expression after 9 hours of exposure. AFB1 5-day exposure shared 89 differentially expressed genes in common with EE2 at 6 hours of exposure and 161 genes with EE2 at 9 hours of exposure (Table 4.8). The 5-day AFB1 exposure appeared to share fewer genes with AFB1 at 3-day than with EE2 exposures.

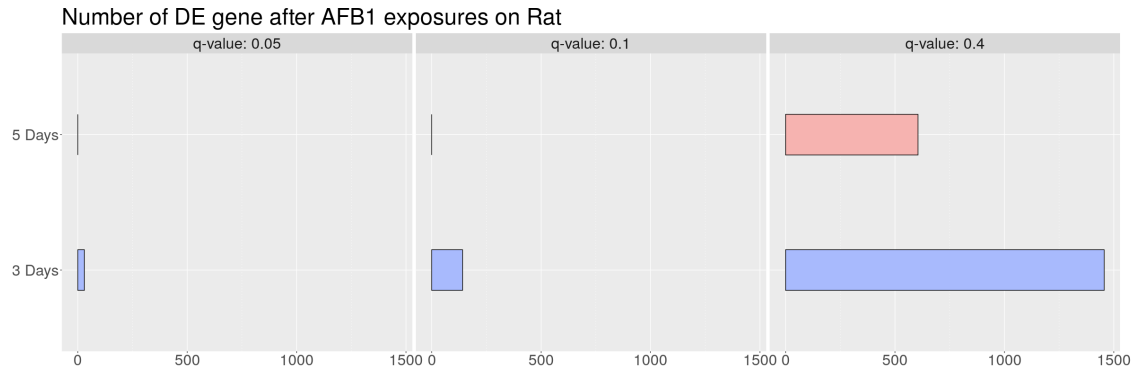

Figure 4.11: Graphic of DE genes after AFB1 (0.3 mg/kg bw×day) exposure in rats, filtered by q-value. Low significantly DE genes at stringent q-value were obtained from GEO datasets AFB1 *in vivo* exposures.

	DE genes		
q-value	0.05	0.1	0.4
3-Day	30	141	1,456
5-Day	0	0	605

Table 4.6: Number of Differentially Expressed genes after 3- and 5-day exposure of AFB1 (0.3 mg/kg bw$\times$day per rat). In an effort to consider a number of DE genes that permit system level GO analysis the less stringent q-value was selected.

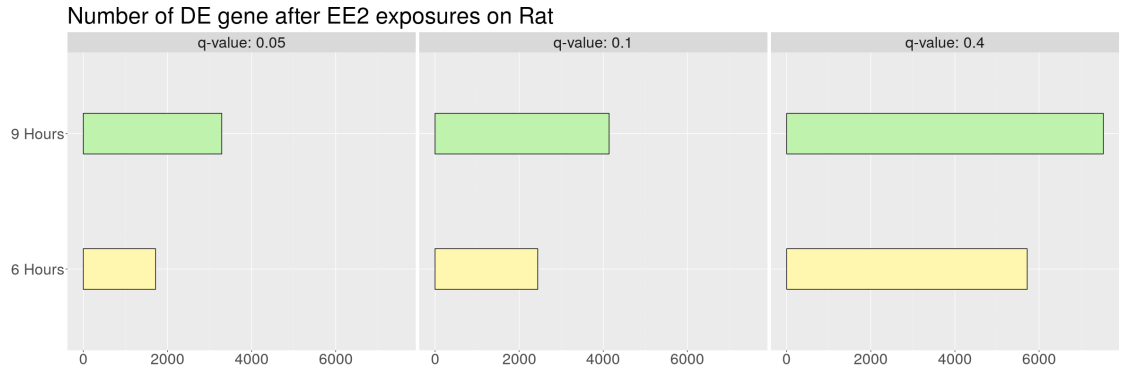

Figure 4.12: Graphic of DE genes after EE2 exposure in rats, filtered by q-value. Compared to AFB1 exposure, EE2 showed a higher amount of DE genes at the more stringent q-value cut-off.

	DE genes		
q-value	0.05	0.1	0.4
6 Hours	1,717	2,442	5,716
9 Hours	3,291	4,138	7,527

Table 4.7: Number of Differentially Expressed genes after 6 and 9 hours of exposure to EE2 (3 mg/kg per rat). Lower numbers of DE genes resulted when compared to the *in vitro* exposures. Despite the higher number of DE genes than the AFB1 *in vivo* exposure, to maintain the same significance, a q-value threshold of 0.4 was selected.

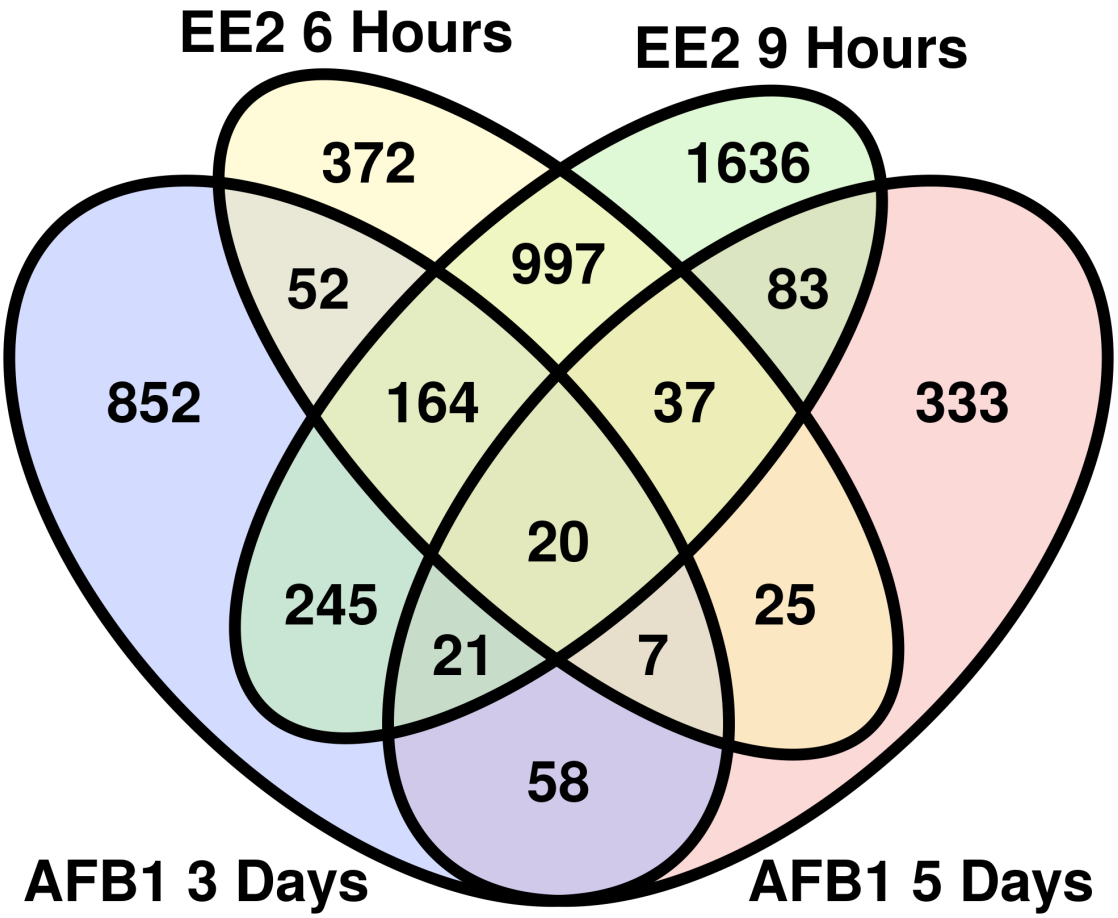

Figure 4.13: Venn Diagram showing overlapping genes among AFB1 0.3 mg/kg bw×day exposure after 3- and 5- day and EE2 3 mg/kg exposure after 6 and 9 hours. Similarly, to the *in vitro* experiment expressions the compounds showed a partial overlap. Interestingly a comparison within the same compound also showed a partial overlap, AFB1 shares only 106 genes of the 1,898 DE genes from both the time steps. Similarly, EE2 expression at 6 and 9 hours shares 1,212 genes from a total of 3,653 DE genes. EE2 expression patterns at 6 hours of exposure had a unique signature of 372 genes, which is significantly lower than the 9 hour sample with 1,636 unique DE genes.

	EE2 6h	EE2 9h
AFB1 3d	243	450
AFB1 5d	89	161

Table 4.8: Overlapping DE genes among EE2 and AFB1 *in vivo* exposures. The AFB1 3-day exposure shares more DE genes with EE2 exposures than with the AFB1 5-day exposure.

4.2.4 GO analysis revealed estrogenic and cancerogenic markers were shared by the *in vivo* exposures to AFB1 and EE2

In order to compare biological effects among the exposures, similarly to the previous analysis, limma outputs filtered by q-values of 0.4 and no Log2FC were inputted into iPathwayGuide (Advaita Bio). BP analysis resulted in a partial overlap among the AFB1 and EE2 exposures (Figure 4.14). The EE2 exposures showed strong unique signatures of 157 and 202 BP for 6 hours and 9 hours exposures respectively (not shown, Table B.4 in Appendix B). Furthermore, a significant amount of the BP terms were shared by the two exposures (134), of these 102 appear to not be related to the AFB1 exposures (not shown, Table B.5 in Appendix B). Interestingly the AFB1 3 and 5-day exposures showed only 4 BP to be shared among the samples, of these only 2 were not related to EE2 exposure (not shown, Table B.3 in Appendix B). Aflatoxin unique signatures are equal to 146 BP terms for the 3-day exposure and 181 for the 5-day exposure (not shown, Table B.2 in Appendix B). Moreover, as seen with gene level analysis there

are higher overlapping BP terms among the AFB1 3-day and EE2 exposures than with AFB1 at 5-day of exposure, resulting in a total of 37 shared BP terms. Of these BP terms shared by the AFB1 3-day exposure 27 are common to both of the EE2 exposures. Any of the EE2 time points shared 5 additional BP terms with the AFB1 3-day exposure. The AFB1 5-day exposure resulted in 22 BP terms shared with both of the EE2 exposures (Table 4.9).

AFB1 3-day exposure. As noted above the AFB1 3-day exposure and GO annotation terms resulted in more shared with EE2 than with the AFB1 5-day exposure. Of the 27 GO terms present in both the EE2 exposures strong correlations with immune and inflammation deregulation were found (7 BP terms). Interestingly NIK/NK-kappaB signalling (adj.p-value: 0.02581) and tumour necrosis factor-mediated signalling pathway (adj.p-value: 0.00217) appear to be related to cell proliferation, differentiation and apoptosis. Moreover, deregulation of the non-canonical NF-kappaB pathway and NIK expression was found in cancers [78] and appears to be regulated by estrogenic and xenoestrogenic activities [79], [80]. Other interesting findings included 8 BP terms related to cellular proliferation and differentiation, 3 of them involved Wnt family signalling, already known oncogenes [81] whose regulation is implicated ERα [82]. Other several BP terms include protein regulation, including ubiquitination, protein targeting and folding. Of the 5 BP terms exclusively shared by the 3-day AFB1 exposure and 6 hour EE2 exposure, 3 are related to metabolic processes,

one to energy metabolism (positive regulation of cristae formation, adj.p-value: 0.01168) and one on differentiation related processes (postsynaptic actin cytoskeleton organization adj.p-value: 0.0316). BP terms unique the the AFB1 3-day and EE2 9 hours expression are more related to DNA damage (positive regulation of double-strand break repair adj.p-value: 0.03176, response to X-ray adj.p-value: 0.0142) and inflammatory conditions (complement activation, alternative pathway adj.p-value: 0.02408v). The complete list of BP terms shared by EE2 exposures and AFB1 3-day exposure is in Table 4.10.

AFB1 5-day exposure 5-day exposure to AFB1 led to 22 shared BP terms with the EE2 exposures, despite the lower number of terms than the AFB1 3-day exposure some are interesting for their involvement in estrogenic response and cancer development. Indeed, 2 processes concerning telomerase activity were deregulated among AFB1 and both EE2 exposures (positive regulation of telomerase activity adj.p-value: 0.03867 and positive regulation of telomere maintenance via telomerase adj.p-value: 0.00539). Also, the upregulation of transferrin transport (adj.p-value: 0.0478) appear to be linked to oral contraceptive intake [83]. Other processes involving protein regulation and metabolism are present in both and 6 hours expressions, while 9 hour of EE2 exposure shared with AFB1 two more processes involving DNA damage and 4 processes inherent inflammatory and immune deregulations (Table 4.11). One more interesting term appears in the AFB1 5-day unique signature, with an adj.p-value of 0.0112, the BP

cellular response to estrogen stimulus appears after only 5-day AFB1 exposure and none of the EE2 exposures (not shown, Table B.2 in Appendix B). This result appears controversial, it could suggest the involvement of estrogenic pathways in AFB1 exposure, but it appears in none of the EE2 exposures. This BP term could appear in the AFB1 5-day exposure because this is the only one exposure that led to a deregulation of ERα gene (*ESR1*) with a significant q-value (*ESR1* Log2FC: -0.864, q-value: 0.396).

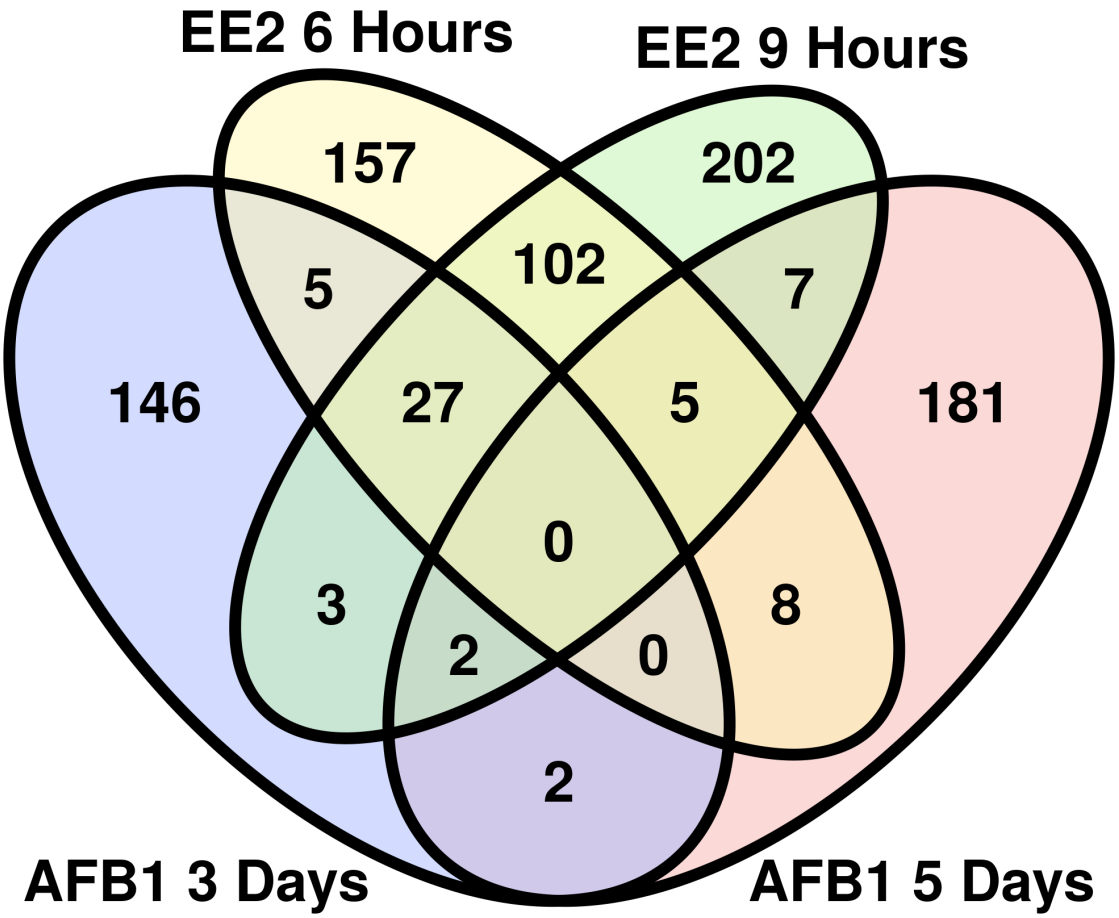

Figure 4.14: Venn diagram showing overlapping BP after *in vivo* exposures. GO analysis resulted in a partial overlap of BP terms among AFB1 and EE2 exposures. Also, as seen with the expression analysis there is higher overlap among the AFB1 3-day and EE2 exposures, resulting in a total of 37 shared BP terms. Of these 27 appear to be in common with both the two exposures of EE2 and any EE2 time ponts shared 5 additional BP terms with the AFB1 3-day exposure. The AFB1 5-day exposure resulted in 22 BP terms shared with EE2 exposures and only 4 with the AFB1 3-day exposure.

	EE2 6h	**EE2 9h**
AFB1 3d	32	32
AFB1 5d	13	14

Table 4.9: Overlapping BP terms among EE2 and AFB1 *in vivo* exposures. Coherently with expression analysis, the AFB1 3-day exposure showed more shared BP terms with EE2 than AFB1 5-day exposures.

GO terms shared by AFB1 3-day and both EE2 exposures	
GO:Biological Process	**adj.p-value**
anaphase-promoting complex-dependent catabolic process	0.0006
antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent	0.00144
chaperone-mediated protein folding	0.00844
DNA replication-dependent nucleosome assembly	0.04657
Fc-epsilon receptor signaling pathway	0.01666
interleukin-1-mediated signaling pathway	0.01099
mitochondrial translational elongation	0.02845
mitochondrial translational termination	0.03453
negative regulation of canonical Wnt signaling pathway	0.02684
negative regulation of G2/M transition of mitotic cell cycle	0.00167
NIK/NF-kappaB signaling	0.02581
nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	3.5E-06
positive regulation of canonical Wnt signaling pathway	0.03642
protein deubiquitination	0.02033
protein polyubiquitination	0.04653
regulation of cellular amino acid metabolic process	5.5E-05
regulation of hematopoietic stem cell differentiation	0.00205
regulation of mRNA stability	0.00725
regulation of transcription from RNA polymerase II promoter in response to hypoxia	2.1E-05
ribosomal small subunit assembly	0.03253
SCF-dependent proteasomal ubiquitin-dependent protein catabolic process	0.00172
SRP-dependent cotranslational protein targeting to membrane	0.00027
stimulatory C-type lectin receptor signaling pathway	0.04943
translational initiation	0.01513
tumor necrosis factor-mediated signaling pathway	0.00217
viral transcription	0.00197
Wnt signaling pathway, planar cell polarity pathway	0.00998
GO terms shared by AFB1 3-day and EE2 6 hours only	
GO:Biological Process	**adj.p-value**
galactose catabolic process	0.0316
glycolytic process	0.03014
positive regulation of cristae formation	0.01168
postsynaptic actin cytoskeleton organization	0.0316
pyrimidine deoxyribonucleoside triphosphate metabolic process	0.03253
GO terms shared by AFB1 3-day and EE2 9 hours only	
GO:Biological Process	**adj.p-value**
complement activation, alternative pathway	0.02408
glyoxylate catabolic process	0.03253
positive regulation of double-strand break repair	0.03176
post-translational protein modification	0.01784
response to X-ray	0.0142

Table 4.10: BP terms shared by AFB1 3-day and both EE2 *in vivo* exposures from iPathway Guide (Advaita) meta-analysis. Correlations with immune, proliferation and apoptosis deregulation were found among the AFB1 3-day exposure and both the EE2 time steps. Additionally, BP terms inherent to DNA damage and inflammatory conditions were found among the AFB1 3-day and EE2 9 hour exposures.

GO terms shared by AFB1 5-day and both EE2 exposures	
GO:Biological Process	**adj.p-value**
cellular response to actinomycin D	0.00581
positive regulation of telomerase activity	0.04309
positive regulation of telomere maintenance via telomerase	0.03867
rRNA processing	0.00539
transferrin transport	0.0478
GO terms shared by AFB1 5-day and EE2 6 hours only	
GO:Biological Process	**adj.p-value**
positive regulation of cellular protein localization	0.04772
positive regulation of fatty acid biosynthetic process	0.03222
positive regulation of proteasomal ubiquitin-dependent protein catabolic process	0.04101
positive regulation of transcription from RNA polymerase II promoter in response to hypoxia	0.02654
regulation of Golgi organization	0.00363
UDP-glucuronic acid transmembrane transport	0.00199
UDP-N-acetylgalactosamine transmembrane transport	0.00199
UDP-N-acetylglucosamine transmembrane transport	0.01127
GO terms shared by AFB1 5-day and EE2 9 hours only	
GO:Biological Process	**adj.p-value**
antigen processing and presentation of exogenous peptide antigen via MHC class II	0.04688
basement membrane assembly	0.04669
cellular response to gamma radiation	0.01777
complement activation, alternative pathway	0.0112
nucleus organization	0.02796
positive regulation of T cell apoptotic process	0.00101
regulation of complement activation	0.00645
response to X-ray	0.02367
valine catabolic process	0.01127

Table 4.11: BP terms shared by the AFB1 5-day and both EE2 *in vivo* exposures from iPathway Guide (Advaita) meta-analysis. Estrogenic and cancerogenic markers were found among the AFB1 5-day and both EE2 exposures. EE2 9 hour exposure shared additional DNA damage and inflammatory BP terms with the AFB1 5-day exposure.

4.3 Protein-Protein Interaction (PPI) networks analysis

4.3.1 *In vitro* exposure

The genes list used in the GO analysis inherent to the AFB1 exposure (filtered by absolute logFC of 0.6 and q-value of 0.05), with the addition of estrogen receptors genes, was inputted into STRING to construct a PPI Network consisting of 1,084 nodes. Of these nodes, 197 were as Downstream Estrogen Receptor Targets (DERT). The centrality analysis and selection of the top 197 central nodes for every centrality measure resulted in a partial overlap of the nodes lying in the most central zone of the network. High closeness centrality means that the node can reach the whole network more quickly than other nodes and the top nodes with high closeness centrality values were in the most central area of the network. Betweenness centrality highlights nodes that control information flow on their neighbours and resulted in a more scattered distribution than the other centrality measures, connecting well branches to the core of the network. Subgraph centrality measure represents the involvement of a node in all subgraphs of the networks and resulted in a distribution of high centrality nodes in a less scattered distribution than Betweenness centrality and showing high overlapping level with closeness centrality (Figure 4.15). Despite the differences of distributions, the top 197 nodes with high closeness central-

ity uncovered 59 DERTs, subgraph centrality uncovered 55 and betweenness centrality 53. The Pearson's chi-square test for independence refuses the Null hypothesis indicating independence among the DERTs found in the entire network and the DERTs recognized by centrality measures (χ^2: 23.422, df: 3, p-value: 3.297E-5) suggesting that the distribution of the estrogen receptor targets is along the most central nodes. Moreover, the post hoc FDR correction and pairwise testing between the groups confirmed the independence of DERTs found by centralities from DERTs in the entire network (Betweenness: 0.012, Closeness: 0.0013, Subgraph: 0.0065 q-value) (Table 4.12). The subnetwork resulting from overlapping nodes consisted of 96 nodes that showed high correlation with DERTs and direct neighbours. Of the 96 nodes, 31 are DERTs and 59 are their direct neighbours, including the Estrogen Receptors α and β, suggesting that estrogen-related proteins have essential roles mediating the expression from AFB1 *in vitro* exposure. The Pearson's chi-square test for independence and FDR correction also showed the independence of the DERTs found in the subnetwork from the entire network, indicating a further suggestion of the non-casual distribution of the DERTs (χ^2: 10.389, df: 1, p-value: 0.001268, q-value: 0.0013). Only 6 proteins were neither DERT nor their neighbours (Figure 4.16).

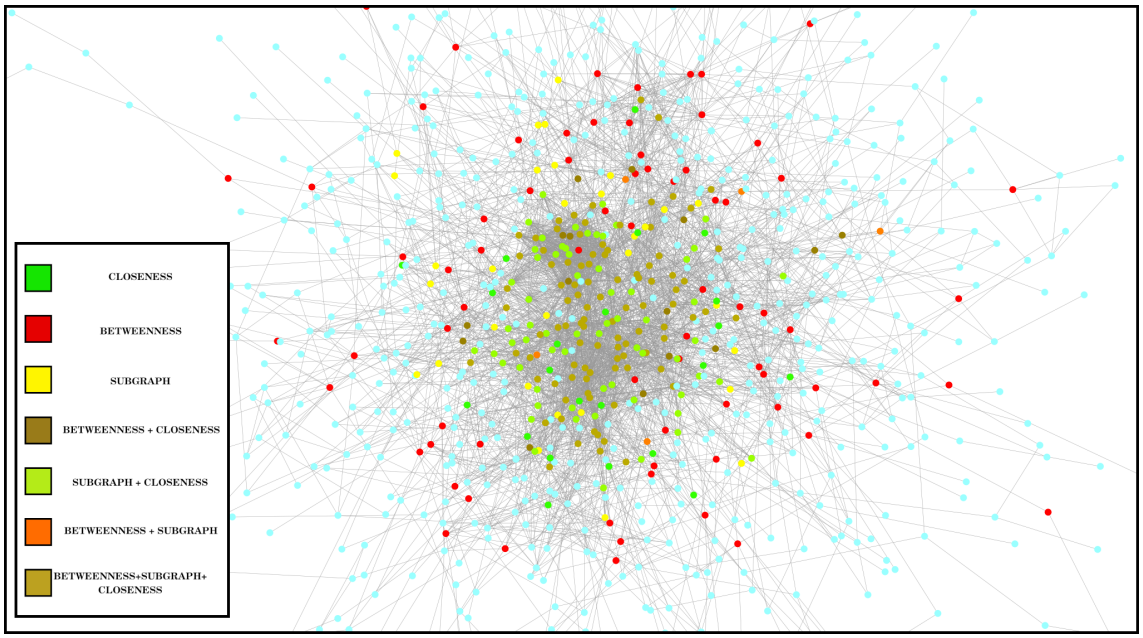

Figure 4.15: Comparison of the high centrality nodes in the PPI network from AFB1 *in vitro* exposure. Betweenness centrality highlights nodes that control information flow on their neighbours and resulted a scattered distribution well connecting branches to the core of the network. Subgraph centrality measure represent the involvement of a node in all subgraphs of the networks and resulted in a less scattered distribution of high centrality nodes than betweenness centrality and showing high overlapping levels with closeness centrality, that resulted in the centrality measure with more central nodes than the others.

Centrality	DERTs	q-value
Betweenness	53	0.0121
Subgraph	55	0.0065
Closeness	59	0.0013

Table 4.12: DERTs recognised by centrality analysis and q-values for AFB1 *in vitro*. All the DERTs recognised by different centrality measures were independent of the amount of DERTs present in the entire network suggesting the effective central distribution of these proteins and a central role on network regulation.

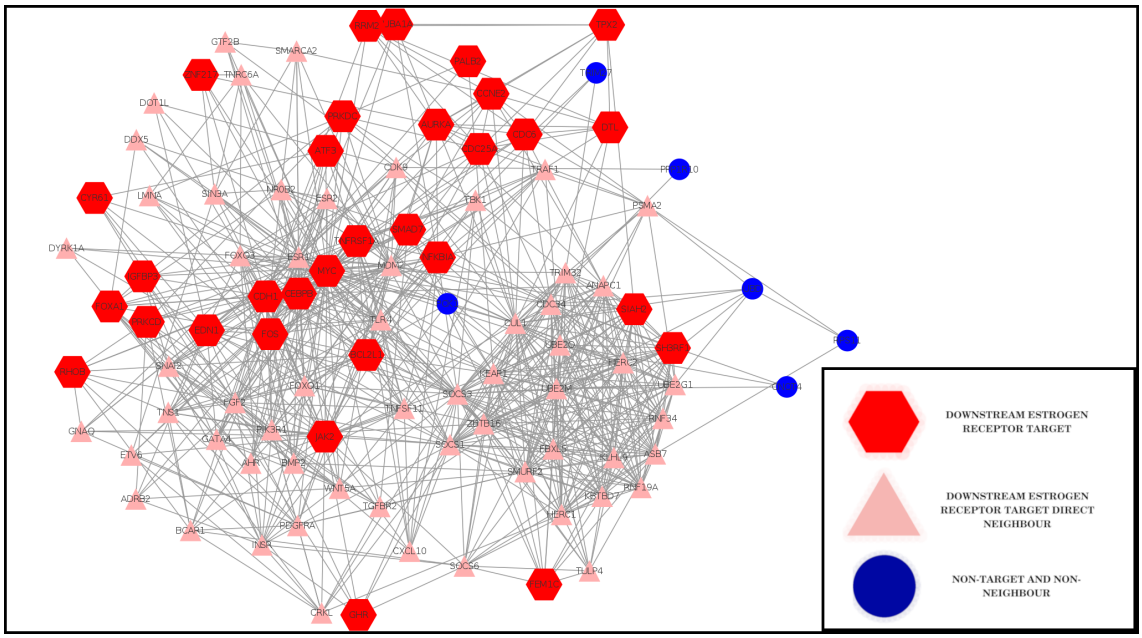

Figure 4.16: Most central nodes and DERTs of the PPI network from AFB1 *in vitro* exposure. The sub network build with overlapping nodes was constituted by 96 nodes, of these 31 are DERTs and 59 direct their neighbours showing strong estrogenic involvement in the core of the expression led from AFB1 *in vitro* exposure.

4.3.2 *In vivo* 3-day exposure

The list of differentially expressed genes inherent to the AFB1 exposure, used in GO analysis (filtered q-value of 0.4), with the addition of the estrogen receptors genes, was inputted into STRING to construct a Network consisting of 1,390 nodes. Of these nodes, 242 resulted as Downstream Estrogen Receptor Targets (DERT). Similarly to the *in vitro* analysis, the distribution of the most central nodes for the different centrality measures demonstrated the properties of these measures. Betweenness centrality result were scattered and able to link singular branches to the core network, subgraph and closeness centralities resulted in more groups in the core

of the network (Figure 4.17). Compared to the *in vitro* network DERTs found by centrality measures were more variable, the betweenness centrality highlighted 56 DERTs, closeness centrality 64 and subgraph centrality 48. The Pearson's chi-square test for independence refused the Null hypothesis indicating independence among the DERTs found in the entire network and the DERTs recognized by centrality measures (χ^2: 13.411, df: 3, p-value: 0.003828). Due to the differences among the centrality measures the post hoc FDR correction and pairwise testing showed that only DERTs found by closeness centrality were statistically significant (betweenness: 0.1242, closeness: 0.0073, subgraph: 0.4612 q-value), suggesting that the DERTs in the network are distributed in those nodes that are closer to all other nodes (Table 4.13). The subnetwork resulting from overlapping nodes consisted of 120 nodes, 27 are DERTs and, excluding SMURF1, all the other nodes are their direct neighbours, including the estrogen receptors α and β, showing high correlation of the core network with DERT and direct neighbours (Figure 4.18). The Pearson's chi-square test for independence and FDR correction did not refuse the Null hypothesis thus not allowing the assertion that the DERTs found in the subnetwork are independent of the DERTs distributed in the entire network (χ^2: 1.6225, df: 1, p-value: 0.2027, q-value: 0.2).

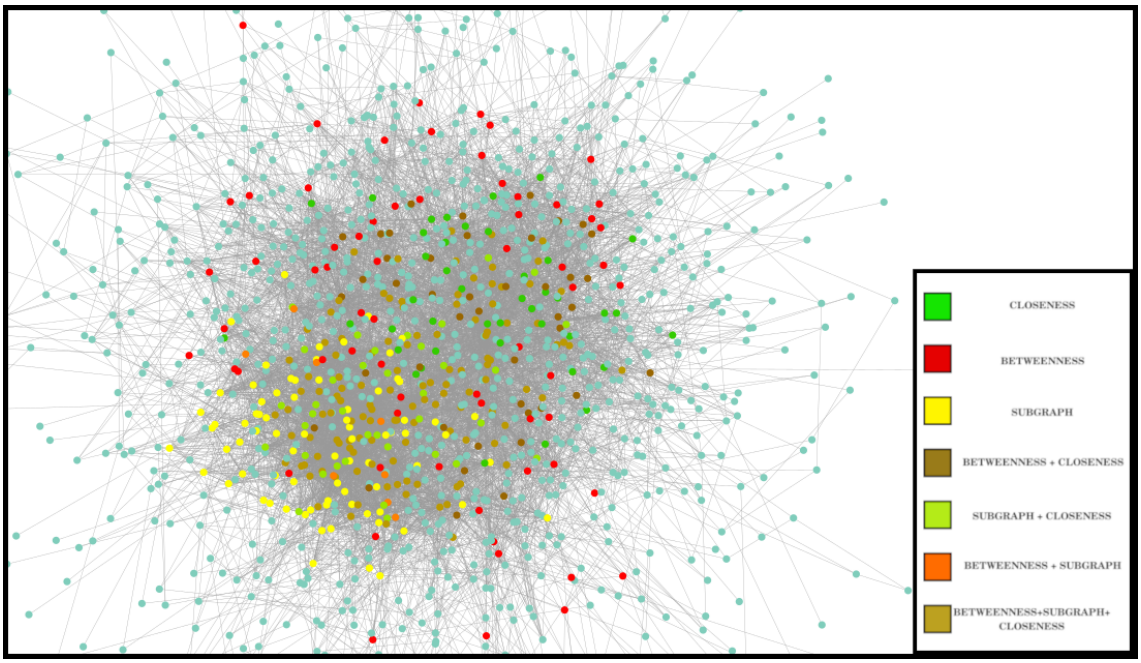

Figure 4.17: Comparison of the high centrality nodes in the PPI network from AFB1 *in vivo* 3-day exposure. Similarly, to *in vitro* analysis the high betweenness centrality nodes showed a dispersed distribution linking peripheral modules to the core network. Subgraph and closeness centralities recognised nodes more distributed in the central area of the network.

Centrality	DERTs	q-value
Betweenness	56	0.1242
Subgraph	64	0.0073
Closeness	48	0.4612

Table 4.13: DERTs recognised by centrality analysis and q-values for AFB1 *in vivo* 3-day exposure. Interestingly, differently to *in vitro* network the amount of DERTs found by centrality measures is more variable. In addition, only high closeness centrality nodes were independent of DERTs distribution in the entire network.

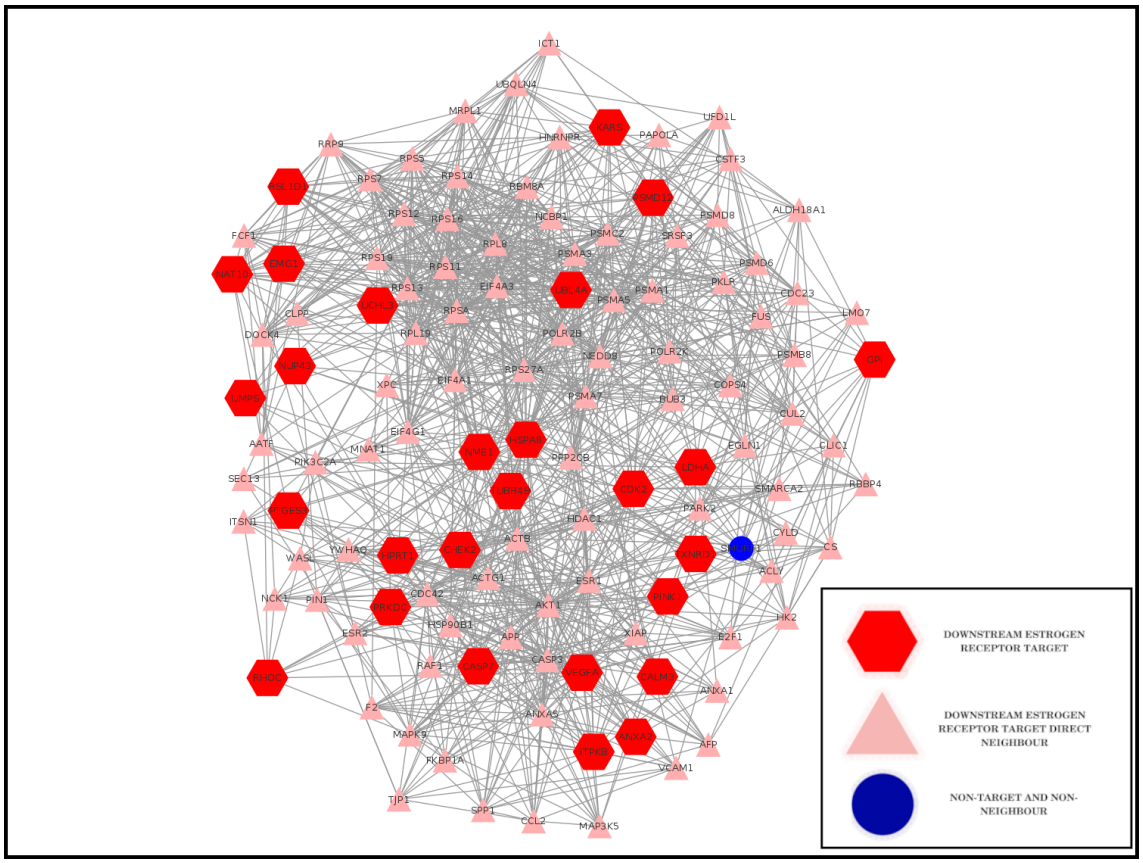

Figure 4.18: Most central nodes and DERTs of the PPI network from AFB1 *in vivo* 3-day exposure. Despite the likely more random distribution of the DERTs, the subnetwork build from most central nodes are highly estrogen correlated. Almost all the nodes are involved with estrogen receptor targets (27 DERTs and 92 direct neighbours).

4.3.3 *In vivo* 5-day exposure

The list of differentially expressed genes inherent to the AFB1 exposure, used in GO analysis (filtered q-value of 0.4), with the addition of the estrogen receptors genes, was inputted in STRING to construct a network consisting of 574 nodes, 101 of which are DERTs. Similarly, to the previous analyses, betweenness centrality highlighted nodes in a scattered distribution, subgraph and closeness centralities showed many overlapping nodes

(Figure 4.19). Low amount of DERTs were found by centrality measure (betweenness: 20, closeness: 22, subgraph: 22) and the Pearson's chi-square test for independence did not refuse the Null hypothesis thus not allowing any assumption regarding the distribution of the DERTs in the network (χ^2: 1.7661, df: 3, p-value: 0.6223). Coherently, the post hoc FDR correction and pairwise testing showed that none of the independence of the DERTs found by the centrality measure were statistically significative (betweenness: 1, closeness: 1, subgraph: 1 q-value) (Table 4.14). The overlapping nodes produced a subnetwork of 43 nodes, of which only 6 are DERTs and 21 their direct neighbours. Nodes in the subnetwork that are not involved in the estrogenic component total 16 and suggest the lower involvement of estrogen receptors pathways in the 5-day exposure to AFB1. Despite the lower correlation with estrogenic components both the estrogen receptors α and β are present in the subnetwork (Figure 4.20). The Pearson's chi-square test for independence and FDR correction did not refuse the Null hypothesis thus not allowing the assertion that the DERTs found in the subnetwork are independent of the DERTs distributed in the entire network (χ^2: 0.15973, df: 1, p-value: 0.6894, q-value: 0.69).

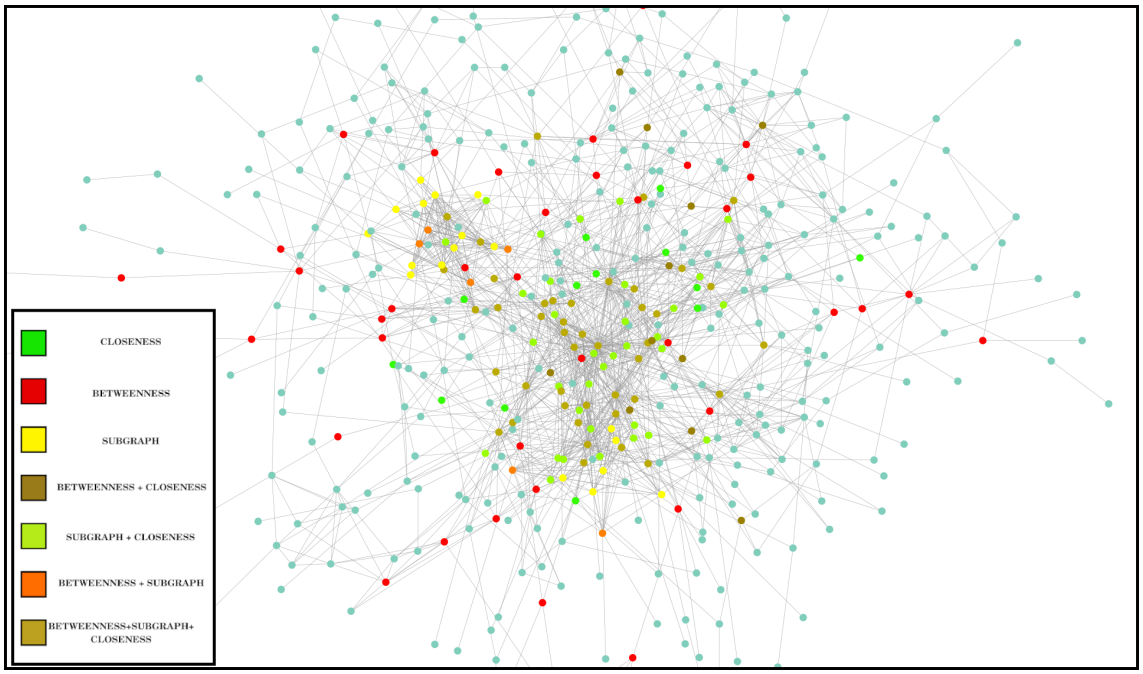

Figure 4.19: Comparison of the high centrality nodes in the PPI network from AFB1 *in vivo* 5-day exposure. The distribution of high centrality nodes resulted in a network structure similar to the previous analysis showing high betweenness centrality nodes linking peripheral modules to the core network. Subgraph and closeness centralities recognised nodes more distributed in the central area of the network and sharing a high amount of nodes.

Centrality	DERTs	q-value
Betweenness	20	1
Subgraph	22	1
Closeness	22	1

Table 4.14: DERTs recognised by centrality analysis and q-values for the AFB1 *in vivo* 5-day exposure. Due to the low number of DERTs recognised by the centrality measures none were independent by DERTs distribution on the network.

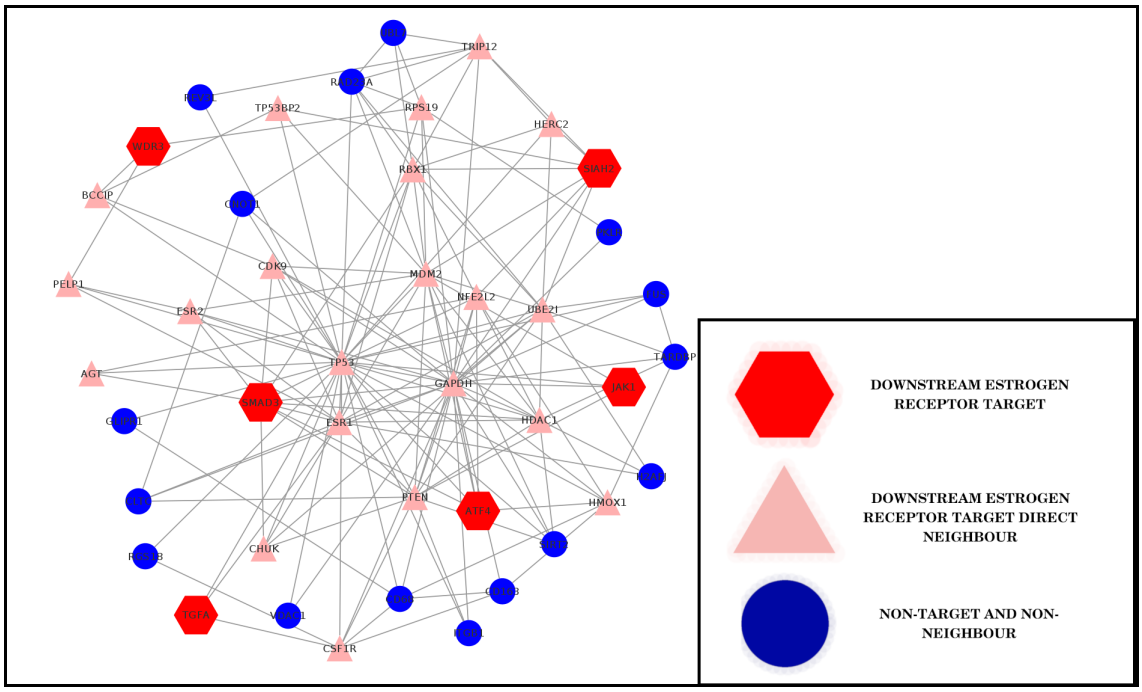

Figure 4.20: Most central nodes and DERTs of the PPI network from AFB1 *in vivo* 5-day exposure. Lower correlation with estrogenic components were found, only 6 proteins were DERTs with 21 direct neighbours. 16 nodes do not share direct contacts with the DERTs. Despite the lower degree of estrogenic correlation both estrogen receptors were found in the subnetwork.

Chapter 5

DISCUSSION

5.1 Docking and Interaction analysis

Different observations can be drawn from the molecular docking simulations and interaction analysis. The free energies of binding and estimated K_D obtained are showing that AFB1 and in particular its metabolite AFQ1 can interact with the estrogen receptors α and β, particularly AFQ1 exhibited greater affinity for ERβ than the known mycoestrogen ZEA (Tables 4.1 and 4.2). Moreover, the constants of dissociation at the micromolar (μM) level, although not typical for high-affinity ligands, are used in chemical gene interaction studies [84] and to highlight the structures that can interact, and even inhibit, specific receptors and transcription factors [85]. This suggests the possible mechanism of AFB1 and its accumulated metabolites, deriving from repeated intakes of contaminated foods, targeted against the ERs and leading to an estrogenic component in the biological activity of AFB1. Interactions formed among the analysed compounds and residues in the binding pocket of ERs showed similarities with

the positive control EE2. AFB1, some of its metabolites and EE2 formed an H-Bond with the same residue (*Arg394*) when binding the ERα and shared several weak interactive residues. Additionally, looking at the interaction between 17-β-Estradiol (EST) and ERα in the crystallised structure on PDB (1ERE)(Figure 5.1), besides the same H-Bond previously highlighted, another specific interaction was shared among EST, AFB1 and all its metabolites, *Phe404* appears to form stacking interactions with the aromatic ring of the ligands (not shown, images in appendix A). Through the superimposition of ERα and β it was possible to find that *Phe356* in ERβ binding pocket is in the same position of ERα *Phe404* suggesting similar roles in the binding ligands (Figure 5.2). This residue appears to make the same stacking interactions with AFB1, all its metabolites and EE2 from the docking simulations (not shown, images in appendix A). Interestingly the different H-Bonds formed among ERs residues and the ligand, besides the several weak interactions shared with EE2, suggest a different effect of the ligand on ERs, similarly to ZEA that despite the high affinity does not form any H-Bond with ERs residues (not shown, images in appendix A). Taken together these results suggest that AFB1 and its metabolites can disrupt ER activity, showing high affinity and, excluding AFQ1, similar binding modes with the positive control EE2. The differences in binding position can also suggest the different effect of aflatoxins on activation of ERs than natural and synthetic activators such as EE2 and EST.

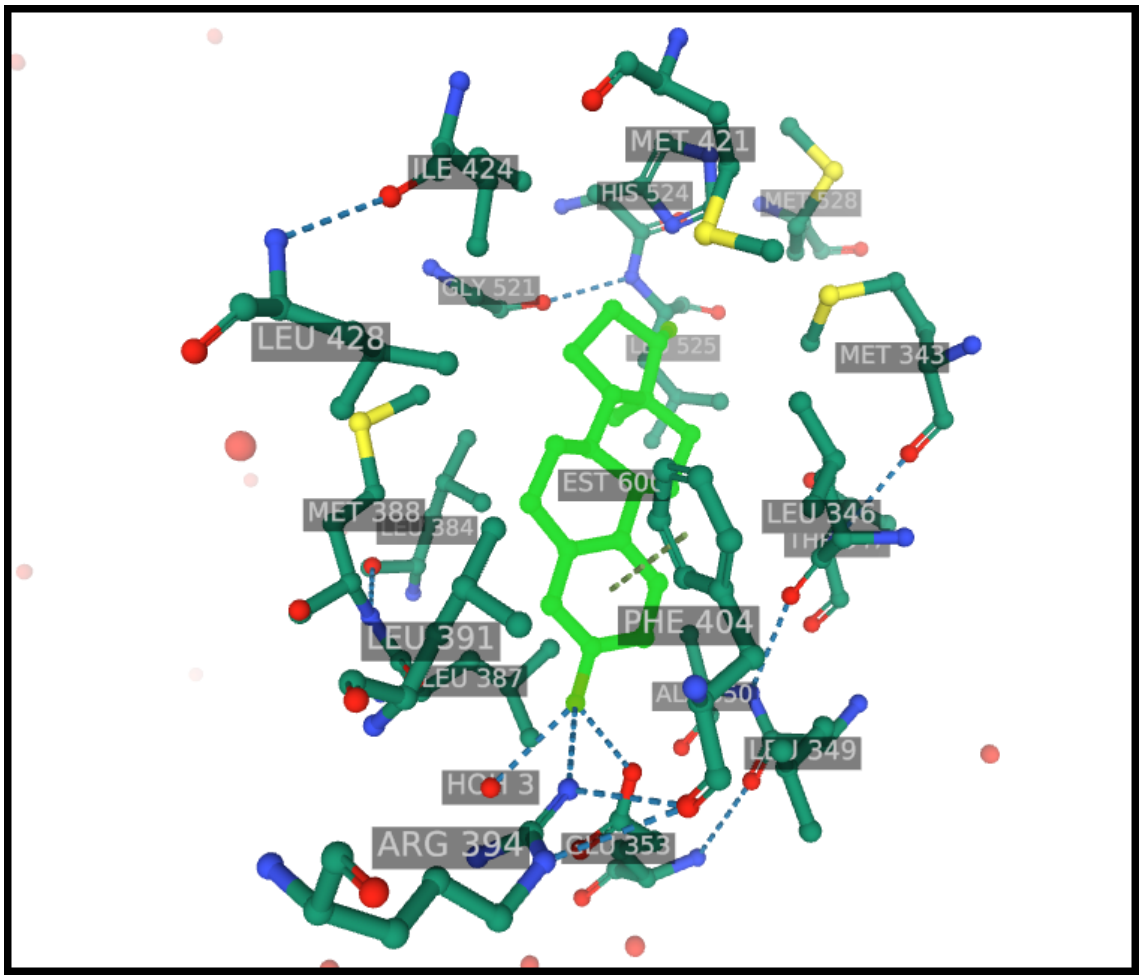

Figure 5.1: Stacking interaction among aromatic ring of EST and *Phe404* of ERα.

5.2 Expression analysis

5.2.1 *in vitro*

The incomplete overlap of the DE genes from the *in vitro* exposure to AFB1 0.8 μM and EE2 3 μM suggests that only a part of the mechanism of action among the substances is shared, this is consistent with previous knowledge about the mechanism of actions of AFB1 that involves geno-

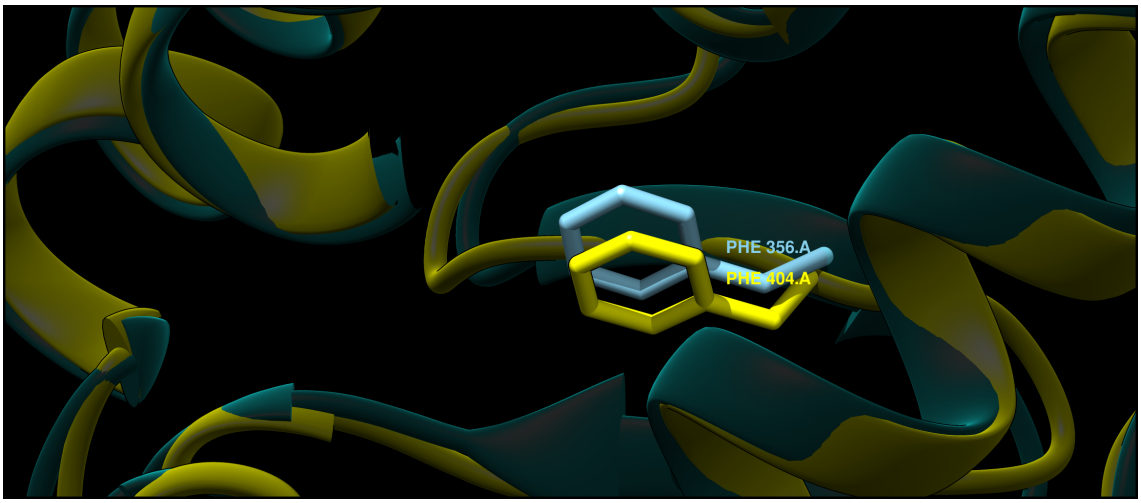

Figure 5.2: Superimposition of ERα and β. *Phe404* of ERα and *Phe356* of ERβ fall in the same position inside the binding pocket, suggesting a similar role in forming interactions with the ligands.

toxicity, immunotoxicity and altered lipid metabolism [3] and EE2 that is known to act mainly through estrogen receptors signalling [86]. Probing deeper by examining GO annotations and observing the singular genes involved in altered BP terms important finding emerge. It is reported that EE2 treatment can induce liver damage by altering choline metabolism [19] as suggested by the alteration of Choline catabolic process (adj.p-value $<$ 0.05 for both AFB1 and EE2 exposures). Moreover, the gene *BHMT* was up-regulated in both EE2 and AFB1 exposures (AFB1: 0.778, EE2: 0.764 log2FC), interestingly the altered expression of *BHMT* was already reported as a biomarker for Hepatocellular Carcinoma (HCC), although in these studies it was down-regulated [87] [88]. Observing the biological processes inherent to apoptosis and the differentially expressed genes involved in these processes the gene *ID1*, involved in the negative regulation of apoptotic signalling, is an inhibitor of DNA binding and was

down-regulated after AFB1 exposure (log2FC: -0.752) and up-regulated after EE2 exposure(log2FC: 0.973). It has been reported that *ID1* responds to ERs signalling and is up-regulated by 17-β-Estradiol (EST) exposure and down-regulated by antagonists of ERs [89]. This result suggests the different effects of AFB1 on the activation state of ERs, effects hypothesized from the docking studies. In a negative regulation of apoptotic signalling (AFB1: 0.04496, EE2: 0.02079 adj.p-value), another interesting deregulated proto-oncogene is *MDM2*, which can target suppressor proteins and appears to be regulated by ERs [90] Similarly, to *ID1*, it shows opposite expression levels after AFB1 and EE2 expression (AFB1: -1.101, EE2: 0.985 log2FC), suggesting a different activation level of the ERs after AFB1 exposure. Additionally, suppressor of cytokine signalling 3 (*SOCS3*), was down-regulated in both exposures (AFB1: -0.837, EE2: -1.027 log2FC). *SOCS3* is regulated by *IL-6* which is responsive to ER signalling, providing a hint on ER involvement in expression mediated by AFB1. Moreover, *SOCS3* was downregulated in different tumours [91], including HCC [92], where liver fibrosis occurs through the activation of TGF-β1 [93]. When negative regulation of transcription by transcription factor localization involved genes are considered, excluding *ID1*, hematopoietically expressed homeobox *HHEX* was similarly down-regulated by both the AFB1 and EE2 exposures (AFB1: -0.813, EE2: -0.609 log2FC). *HHEX* was down-regulated in HHC samples and its over-expression appeared to reduce malignancy in HCC cells

[94]. From an immune response related BP perspective we note two interesting observations. Firstly, *CXCL6* was up-regulated (AFB1: 1.025, EE2: 1.107 log2FC). Over-expression of *CXCL6* enhances proliferation in hepatic cells [95] and is related to pathological conditions in the liver, including HCC [96] and liver fibrosis [97]. Secondly, the other gene to be noted is *F2RL1* that in both exposures was down-regulated (AFB1: -0.871, EE2: -0.812 log2FC), contrary to our situation the over-expression of this gene is related to HCC proliferation [98]. Despite the opposite condition among cell growth and *F2RL1*, the similar alteration of the expression of proliferative signaling such as *CXCL6* and *F2RL1* resulting from both exposures to AFB1 and EE2 suggests that common mechanisms are involved in the hepatotoxic effects of both the substances. The GO annotation analysis and alterations in singular genes suggests a tendency towards oncogenic endpoints from both exposures. The presence of altered BP terms inherent to proliferation, immune response and directly related to ERs activity provide a hint of the estrogenic effects from AFB1 exposures.

5.2.2 *in vivo*

Differential gene expression and overlap among EE2 3 mg/kg 6 and 9 hours treatments and AFB1 0.3 mg/kg bw$\times$day treatments suggest the possibility of a common mechanism of action among the two substances. The higher number of deregulated genes among AFB1 3-day and EE2 treatments and AFB1 5-day and EE2 9 hours treatments than those genes that

were shared by the two AFB1 treatments suggests the effective existence of an estrogenic component of Aflatoxin's mechanism of action. Moreover, the lower number of deregulated genes shared by 5-day of treatment with AFB1 and EE2 treatments gives a hint that the accumulation of cytotoxic effects caused by AFB1 exceed the estrogenic effects found with 3-day exposure to AFB1. In the BP chaperone-mediated protein folding term (adj.p-value: 0.00844) the heat-shock protein *HSPA8* was up-regulated by AFB1 3-day treatment and both the EE2 treatments, only in the AFB1 5-day treatment it was down-regulated (AFB1 3-day: 1.180, AFB1 5-day: -0.396, EE2 6 hours: 0.996, EE2 9 hours: 1.308 log2FC). Interestingly this, and other heat-shock proteins are directly related to ERα activity [99], the different level of expression led by AFB1 5-day treatment strengthens the hypothesis that the longer exposure to AFB1 could result in more cytotoxic effects than ER related activities. Evidence suggests that NF-κB responds to estrogen modulation through ERs activity [100]. Observing deregulated genes involved in the BP term inherent to the non-canonical NF-κB signalling pathway NIK/NF-kappaB signalling (adj.p-value: 0.02581), we noted that the AKT serine/threonine kinase 1 *ATK1* was up-regulated after 3-day AFB1 treatment and in both the EE2 treatments at similar levels (AFB1 3-day: 0.515, EE2 6 hours: 0.470, EE2 9 hours: 0.510 log2FC). *AKT1* is an estrogen activated gene, perhaps through non-nuclear estrogen receptors such as *GPER1* [101] and it is related to the suppression of apoptosis through NF-κB non-canonical signalling [102], indicating that

proliferative effects of estradiol could be also shared by the early exposure response to AFB1. It should be emphasized that the longer exposure to AFB1 involved a down-regulation of *AKT1* (log2FC: -0.414) which did not allow the 5-day exposure to fall within the BP term and suggesting that the accumulation of cytotoxic effects could inhibit the proliferative, and presumably estrogenic component, of AFB1 early exposure. DNA replication-dependent nucleosome assembly (adj.p-value: 0.04657) is another term that is involved in proliferation and found in AFB1 3-day and both EE2 treatments. Here the gene encoding for the chromatin remodelling factor RB binding protein 4 (*RBBP4*) was up-regulated for all of the treatments (AFB1 3-day: 0.832, AFB1 5-day: 0.275, EE2 6 hours: 0.491, EE2 9 hours: 0.343 log2FC). Nevertheless, the lower up-regulation in the 5-day AFB1 exposure did not permit consideration of this longer exposure within the BP term analyses. This protein is reported to interact with ERα mediated gene expression [103]. Its direct interaction with the tumour suppressor gene *RB* and expression were studied in human primary HCC and were found to be up-regulated in this type of tumour [104]. Moreover, this gene is reported to be a NF-κB target. Inhibiting NF-κB down-regulated *RBBP4* suggesting its involvement in the estrogen receptors related pathways and proliferative effects [105]. Looking at other deregulated genes involved in NF-κB signalling, in the BP term positive regulation of double-strand break repair (adj.p-value: 0.03176), the gene *MGMT* [106] was found be differentially expressed by AFB1 and EE2

treatments. Indeed, *MGMT* encodes for methylguanine methyltransferase involved in DNA repair from the activity of alkylating agents and the genotoxic mechanism of action of AFB1 and its metabolites involves guanine alkylation [3]. The differential expression of *MGMT* from AFB1 3-day exposure resulted in a strong down-regulation (log2FC: -2.207) while the longer exposure led to a strong up-regulation of the gene (log2FC: 2.467). EE2 exposures led to a lower but constant up-regulation of the gene (6 hours: 0.553, 9 hours: 0.866). These effects allow two interpretations. On one hand the down-regulation found in the early exposure could suggest that the lower level of genotoxic effects in rats after early AFB1 exposure arose with the accumulation of AFB1 and its principal genotoxic metabolite Epo-AFB1. On the other hand, the strong initial down-regulation suggests that the early exposure to AFB1 resulted in a strong inhibition of DNA repair mechanism that could improve the susceptibility to genotoxic and cancerogenic effects of the mycotoxin. Also, the higher expression of *MGMT* in EE2 treated rats was previously reported *in vitro* as a mechanism in response to estrogen receptors mediated DNA alkylation [107]. A gene showing similar behaviour to *MGMT* was *NFE2L2*, indeed from the AFB1 exposures it was initially down-regulated and subsequently up-regulated (AFB1 3-day: -1.220, AFB1 5-day: 0.628 log2FC). Similarly to *MGMT*, *NFE2L2* was slightly up-regulated by the EE2 treatments (EE2 6 hours: 0.593, EE2 9 hours: 0.767 log2FC). *NFE2L2* encodes a transcription factor that regulates genes containing antioxidant response elements

in response to oxidative stress, its down-regulation in the initial phase of AFB1 treatment suggests a lower level of oxidative stress and DNA damage if compared with later AFB1 exposures or an increased susceptibility to oxidative stress related DNA damage. Estrogen and estrogen receptors related pathways were previously associated with the increase of oxidative stress and genotoxic effects [108] [109]. These effects explain the up-regulation of *NFE2L2* expression. The differential expression of *MGMT* and *NFE2L2* were not directly linked with our hypothesis of estrogenic components in the AFB1 mechanism of action, but hint that the initial exposure to AFB1 could inhibit different and unrelated mechanisms of protection for DNA damage, improving the susceptibility and deleterious effects that the toxin exerts against hepatocytes. Another differentially expressed gene that showed effects similar to what is previously reported in the literature is *TP53*, from which it is already known that the genotoxic effects of AFB1 are particularly focused on codon 249 of *TP53* which could compromise normal expression of this tumour suppressor gene. In our analysis, *TP53* was not deregulated in the early exposure to AFB1 but after 5-day of treatment there was a significant down-regulation of the gene (log2FC: -0.510). On the other hand EE2 slightly up-regulated *TP53* in both treatments (EE2 6 hours: 0.399, EE2 9 hours: 0.378 log2FC). These results suggest that the genotoxic effects of AFB1 against *TP53* gene are mostly like to rise after the longer exposure, leading to a down-regulation of the gene and its cancerogenic effects; Also, the fact that EE2 probably

had no effects on *TP53* expression but the increase in other DNA damages owing to by the rise of oxidative stress led to its slight over-expression provides insights on the genotoxicity of EE2.

5.3 PPI network analysis

Network analysis from *in vitro* 24 hour exposure to 0.8 μM AFB1 highlighted 96 core nodes, of them the 31 where DERTs (32.3%) and 59 are direct neighbours (61.5%) including the ER α and β. The 93.8% of the core and most central nodes are related to the estrogenic pathways suggesting their high involvement in the PPI network and related regulation. The subnetwork is organized in two possible clusters, the bigger one contains most of the DERTs including ERα and β (in the figures reported as ESR1 and ESR2)and some interesting proteins including the proto-oncogene MYC whose expression is enhanced by estrogen [110] and is recruited by ERα in estrogen responsive promoters [111] indicating direct interactions between the receptor and MYC and its involvement in estrogen related proliferation. Other evidence of the involvement of MYC in liver diseases are provided by its interactions in NF-κB and CXCL6 pathways in liver fibrosis [97]. Overexpression of MYC led to an increased HCC progression in zebrafish [112]. Finally, MYC is recognized to have a central role in chronic liver diseases including fibrosis, cirrhosis, and HCC [113]. Another interesting node is the component of the transcription factor complex AP-1 FOS, that is implied in proliferation regulation. AP-1 FOS expres-

sion is regulated by extranuclear estrogen signalling [114] and plays important roles in many tumours. Moreover, in a study that analysed 422 HCC samples, FOS was found to be a possible therapeutic target for its important role in tumorigenesis [115]. Despite the fact that GPER1 was not present in the core network, the presence of FOS as key node suggests a possible disrupting effect of AFB1 via estrogen receptors located in the membrane. Fewer studies were carried on GPER1, resulting in less information about its interactions and a lower score obtained by the STRING scoring function, perhaps this has led to a lower degree level (edges) and the exclusion of the node from centrality studies. The cell cycle-regulated kinase AURKA, found during the mitosis and involved in stabilization of spindle during chromosome segregation, was another key node. AURKA expression is estrogen-induced [116]; It is overexpressed in many tumours and was found in HCC provide radiation resistance by activating NF-κB [117]. The activation of NF-κ pathway has already been noted by different our results in differential gene expression and gene ontology analysis. Also, still inherent to NF-κB pathways JAK2 was found, this tyrosine kinase is a regulator of ERα and modulates the expression of *SOCS3*, that in the core network is a key node that links the two clusters [118]. Moreover, JAK2 responds to other xenoestrogens such as Bisphenol F a substitute for the endocrine disruptor Bisphenol A, thereby providing another link for ER disruption in response to AFB1 exposure [118]. The second cluster linked to the first one mainly through the proteins SOCS1-3-6 contains

only 3 DERTS. Four of the six nodes were not directly linked to DERTs.

In vivo AFB1 0.3 mg/kg bw$\times$day treatment for 3-day led to a core network of 120 nodes, 22.5% of these are DERTs (27 nodes) and 76.7% are direct neighbours (92 nodes). 99.2% of the core network is related to estrogenic targets, and of these nodes some key proteins highlighted in the expression analysis were revealed to play a crucial role in regulatory process, these proteins include AKT1, RBBP4 and HSPA8. Moreover, many nodes inherent to translation such as the elongation factors EIF4A1, EIF4G1 and many ribosomal proteins including RPS5, RPS7, RPS15 suggest a strong impairment of protein synthesis, already known as an AFB1 exposure effect [3]. Also, numerous nodes inherent to proteasomal subunits including PSMA1, PSMA5, PSMC7, PSMD6 suggest two things, the first is deregulation of protein stability, in line with previous knowledge on AFB1 effects, and the second is the possible deregulation of cell cycle due to the involvement of the proteasome in cell cycle regulation [119]. In support to the deregulation of cell cycle other key factors are found as core nodes, the cell division cycle proteins CDC23 and CDC42 that directly interact with ERβ [120], the cyclin dependent kinase 2 CDK2 that was shown to be down-regulated by long-term estradiol treatments and to prevents liver regeneration [121]. Finally, calmondulin3 CALM3, a regulator of the cell cycle and cytokinesis whose promoter contains estrogen responsive elements and can modulate ERα activity, stabilizing the receptor dimers and increasing its activity [122] [123]. The PPI network from

the 5-day AFB1 treatment resulted in a subnetwork smaller and less related to estrogen target, of the 43 nodes only 6 were DERTs (14%) and 21 their direct neighbours (48.8%) suggesting a significant lower involvement of the estrogenic components in the network and expression regulation. Observing the nodes with higher degree level uncovered TP53, suggesting a high involvement of cytotoxic and genotoxic effects that led to apoptotic processes. A further indication of the accumulation of genotoxic effects is the presence of NFE2L2, whose expression was up-regulated after 5-day of treatment and suggests an increased level of oxidative stress due to accumulation of AFB1 metabolites such as Epo-AFB1. Another interesting node is the proto-oncogene MDM2, found in the *in vitro* expression analysis, after 5-day of exposure its expression was increased in the rat liver (log2FC: 2.094). Moreover, the presence of CDK9 and the transforming growth factor alpha TGFA, that is a DERT highlights the strong tumorigenic component of the core network, suggesting that prolonged exposure to AFB1 and metabolites, and the accumulation of genotoxic effects resulted in carcinogenic development by the direct DNA damage exerted by the toxins.

Chapter 6

CONCLUSION

The integration of different system analyses revealed the possibility of interaction among Aflatoxins and Estrogen Receptors, and with better affinities for ERβ than Zearalenone. A certain overlap of the differentially expressed genes and biological processes among AFB1 and EE2 exposures both *in vitro* and *in vivo*, highlighted key processes and expression of genes involved in liver diseases and carcinogenesis, the principal effect of aflatoxin on the liver. Finally, the strong estrogenic components in the core of PPI networks constructed using gene expression information, features found mainly in the *in vitro* and early *in vivo* exposures, provides an additional hint of the estrogenic component in the mechanisms of action of aflatoxin. All these results indicate that Aflatoxin B1 has the potential to be considered an endocrine disruptor and a xenoestrogen, that can alter the activation state of nuclear and membrane estrogen receptors in hepatocytes during the early exposure response. These results are strengthened by the ambiguous effect of estradiol and EE2 on the liver, indeed a physiological exposure to estradiol appears to protect against the development of

HCC[45] but estrogen deregulation and exposure to xenoestrogens appear to increase the chance of developing HCC [20] [46]. Further studies are needed to elucidate the involvement of ERs in Aflatoxin exposures, studies of *in vitro* binding assays could confirm if ERs effectively bind AFB1. *In vivo* exposures with lower concentrations of both the compounds could give a better idea of the effects on hepatocytes at environmental and medically relevant concentrations. Overexpression and ablation of key genes found in expression analysis could suggest the outcomes and may confirm the involvement in HCC development. In conclusion, this integrative systems approach provide insights into the estrogenic component of aflatoxin's mechanism of action from different sources, but further *in vitro* and *in vivo* studies will be needed to further elucidate and eventually confirm our findings.

Chapter 7

ACKNOWLEDGMENTS

First and foremost I am extremely grateful to my supervisors, Prof. Oliana Carnevali from UNIVPM and Prof. Gary Hardiman from QUB for the wonderful opportunity they gaven me and for their insightful comments and suggestions. This experience abroad gave me the opportunity to meet beautiful people and to mature. I would also like to thank Dr. Willian Abraham da Silveira for the infinite support and the thousand knowledge that he transmitted to me. I would like to thank all the members in the Hardiman Group. It is their kind help and support that have made my study and life in the Belfast a wonderful time. Thank you James, Christina, Tom, Clara and everyone else. I would like to express my gratitude to my parents who have supported me for years by allowing me to chase the biggest of my dreams. Finally, I would like to thank my fiancee Eleonora for her infinite love and for being able to reach me despite all the difficulties.

Bibliography

[1] B. R. Rushing and M. I. Selim, "Aflatoxin b1: A review on metabolism, toxicity, occurrence in food, occupational exposure, and detoxification methods," *Food and Chemical Toxicology*, vol. 124, pp. 81–100, 2019, ISSN: 0278-6915. DOI: https://doi.org/10.1016/j.fct.2018.11.047. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0278691518308494.

[2] E. P. on Contaminants in the Food Chain (CONTAM), D. Schrenk, M. Bignami, L. Bodin, J. K. Chipman, J. del Mazo, B. Grasl-Kraupp, C. Hogstrand, L. (Hoogenboom, J.-C. Leblanc, C. S. Nebbia, E. Nielsen, E. Ntzani, A. Petersen, S. Sand, T. Schwerdtle, C. Vleminckx, D. Marko, I. P. Oswald, A. Piersma, M. Routledge, J. Schlatter, K. Baert, P. Gergelova, and H. Wallace, "Risk assessment of aflatoxins in food," *EFSA Journal*, vol. 18, no. 3, e06040, 2020. DOI: 10.2903/j.efsa.2020.6040. eprint: https://efsa.onlinelibrary.wiley.com/doi/pdf/10.2903/j.efsa.2020.6040. [Online]. Available: https://efsa.onlinelibrary.wiley.com/doi/abs/10.2903/j.efsa.2020.6040.

[3] N. Benkerroum, "Chronic and acute toxicities of aflatoxins: Mechanisms of action," *International Journal of Environmental Research and Public Health*, vol. 17, no. 2, p. 423, Jan. 2020, ISSN: 1660-4601. DOI: 10.3390/ijerph17020423. [Online]. Available: http://dx.doi.org/10.3390/ijerph17020423.

[4] S. Marchese, A. Polo, A. Ariano, S. Velotto, S. Costantini, and L. Severino, "Aflatoxin b1 and m1: Biological properties and their involvement in cancer development," *Toxins*, vol. 10, no. 6, p. 214, May 2018, ISSN: 2072-6651. DOI: 10.3390/toxins10060214. [Online]. Available: http://dx.doi.org/10.3390/toxins10060214.

[5] M. Storvik, P. Huuskonen, T. Kyllönen, S. Lehtonen, H. El-Nezami, S. Auriola, and M. Pasanen, "Aflatoxin b1 – a potential endocrine disruptor – up-regulates cyp19a1 in jeg-3 cells," *Toxicology Letters*, vol. 202, no. 3, pp. 161–167, 2011, ISSN: 0378-4274. DOI: https://doi.org/10.1016/j.toxlet.2011.01.028. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378427411000464.

[6] T. Zamir-Nasta, M. Razi, H. Shapour, and H. Malekinejad, "Roles of p21, p53, cyclin d1, CDK-4, estrogen receptor alpha in aflatoxin b1-induced cytotoxicity in testicular tissue of mice," *Environmental Toxicology*, vol. 33, no. 4, pp. 385–395, Dec. 2017. DOI: 10.1002/tox.22524. [Online]. Available: https://doi.org/10.1002/tox.22524.

[7] D. Benford, J.-C. Leblanc, and R. W. Setzer, "Application of the margin of exposure (MoE) approach to substances in food that are genotoxic and carcinogenic," *Food and Chemical Toxicology*, vol. 48, S34–S41, Jan. 2010. DOI: 10.1016/j.fct.2009.10.037. [Online]. Available: https://doi.org/10.1016/j.fct.2009.10.037.

[8] E. V. de Perre, L. Jacxsens, C. Lachat, F. E. Tahan, and B. D. Meulenaer, "Impact of maximum levels in european legislation on exposure of mycotoxins in dried products: Case of aflatoxin b1 and ochratoxin a in nuts and dried fruits," *Food and Chemical Toxicology*, vol. 75, pp. 112–117, Jan. 2015. DOI: 10.1016/j.fct.2014.10.021. [Online]. Available: https://doi.org/10.1016/j.fct.2014.10.021.

[9] K. D. Ruyck, I. Huybrechts, S. Yang, D. Arcella, L. Claeys, S. Abbeddou, W. D. Keyzer, J. D. Vries, M. Ocke, J. Ruprich, M. D. Boevre, and S. D. Saeger, "Mycotoxin exposure assessments in a multi-center european validation study by 24-hour dietary recall and biological fluid sampling," *Environment International*, vol. 137, p. 105 539, Apr. 2020. DOI: 10.1016/j.envint.2020.105539. [Online]. Available: https://doi.org/10.1016/j.envint.2020.105539.

[10] M. Sayiner, P. Golabi, and Z. M. Younossi, "Disease burden of hepatocellular carcinoma: A global perspective," *Digestive Dis-*

eases and Sciences, vol. 64, no. 4, pp. 910–917, Mar. 2019. DOI: 10.1007/s10620-019-05537-2. [Online]. Available: https://doi.org/10.1007/s10620-019-05537-2.

[11] P. Battilani, P. Toscano, H. J. V. der Fels-Klerx, A. Moretti, M. C. Leggieri, C. Brera, A. Rortais, T. Goumperis, and T. Robinson, "Aflatoxin b1 contamination in maize in europe increases due to climate change," *Scientific Reports*, vol. 6, no. 1, Apr. 2016. DOI: 10.1038/srep24328. [Online]. Available: https://doi.org/10.1038/srep24328.

[12] A. Medina, A. Rodriguez, and N. Magan, "Effect of climate change on aspergillus flavus and aflatoxin b1 production," *Frontiers in Microbiology*, vol. 5, p. 348, 2014, ISSN: 1664-302X. DOI: 10.3389/fmicb.2014.00348. [Online]. Available: https://www.frontiersin.org/article/10.3389/fmicb.2014.00348.

[13] C. Monneret, "What is an endocrine disruptor?" *Comptes Rendus Biologies*, vol. 340, no. 9-10, pp. 403–405, Sep. 2017. DOI: 10.1016/j.crvi.2017.07.004. [Online]. Available: https://doi.org/10.1016/j.crvi.2017.07.004.

[14] K. Kowalska, D. E. Habrowska-Górczyńska, and A. W. Piastowska-Ciesielska, "Zearalenone as an endocrine disruptor in humans," *Environmental Toxicology and Pharmacology*, vol. 48, pp. 141–149, 2016, ISSN: 1382-6689. DOI: https://doi.org/10.1016/

j . etap . 2016 . 10 . 015. [Online]. Available: http : / / www . sciencedirect.com/science/article/pii/S138266891630268X.

[15] J. Fink-Gremmels and H. Malekinejad, "Clinical effects and biochemical mechanisms associated with exposure to the mycoestrogen zearalenone," *Animal Feed Science and Technology*, vol. 137, no. 3-4, pp. 326–341, Oct. 2007. DOI: 10.1016/j.anifeedsci.2007.06.008. [Online]. Available: https://doi.org/10.1016/j.anifeedsci.2007.06.008.

[16] M. Adeel, X. Song, Y. Wang, D. Francis, and Y. Yang, "Environmental impact of estrogens on human, animal and plant life: A critical review," *Environment International*, vol. 99, pp. 107–119, Feb. 2017. DOI: 10.1016/j.envint.2016.12.010. [Online]. Available: https://doi.org/10.1016/j.envint.2016.12.010.

[17] Y.-C. Wang, G.-L. Xu, W.-D. Jia, S.-J. Han, W.-H. Ren, W. Wang, W.-B. Liu, C.-H. Zhang, and H. Chen, "Estrogen suppresses metastasis in rat hepatocellular carcinoma through decreasing interleukin-6 and hepatocyte growth factor expression," *Inflammation*, vol. 35, no. 1, pp. 143–149, Feb. 2011. DOI: 10.1007/s10753-011-9299-3. [Online]. Available: https://doi.org/10.1007/s10753-011-9299-3.

[18] M. M. Hassan, G. Botrus, R. Abdel-Wahab, R. A. Wolff, D. Li, D. Tweardy, A. T. Phan, E. Hawk, M. Javle, J.-S. Lee, H. A. Torres, A. Rashid, R. Lenzi, H. M. Hassabo, Y. Abaza, A. S. Shalaby, S.

Lacin, J. Morris, Y. Z. Patt, C. I. Amos, S. A. Khaderi, J. A. Goss, P. K. Jalal, and A. O. Kaseb, "Estrogen replacement reduces risk and increases survival times of women with hepatocellular carcinoma," *Clinical Gastroenterology and Hepatology*, vol. 15, no. 11, pp. 1791–1799, Nov. 2017. DOI: 10.1016/j.cgh.2017.05.036. [Online]. Available: https://doi.org/10.1016/j.cgh.2017.05.036.

[19] M. J. Chang, Y. J. Xu, W. X. He, C. L. Zhang, X. P. Li, and D. Liu, "Intestinal injury in the rat model of 17alpha-ethynylestradiol-induced intrahepatic cholestasis," *Journal of Digestive Diseases*, vol. 17, no. 11, pp. 756–763, 2016. DOI: 10.1111/1751-2980.12407. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1751-2980.12407. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/1751-2980.12407.

[20] L. Wan and P. O'Brien, "Molecular mechanism of 17alpha-ethinylestradiol cytotoxicity in isolated rat hepatocytes," *Canadian Journal of Physiology and Pharmacology*, vol. 92, no. 1, pp. 21–26, Jan. 2014. DOI: 10.1139/cjpp-2013-0267. [Online]. Available: https://doi.org/10.1139/cjpp-2013-0267.

[21] B. J. Young, G. C. Lopez, D. S. Cristos, D. C. Crespo, G. M. Somoza, and P. Carriquiriborde, "Intersex and liver alterations induced by long-term sublethal exposure to 17-alpha-ethinylestradiol

in adult male cnesterodon decemmaculatus (pisces: Poeciliidae),” *Environmental Toxicology and Chemistry*, vol. 36, no. 7, pp. 1738–1745, Aug. 2016. DOI: 10.1002/etc.3547. [Online]. Available: https://doi.org/10.1002/etc.3547.

[22] M. T. Hultman, Y. Song, and K. E. Tollefsen, “17-alpha-ethinylestradiol (EE2) effect on global gene expression in primary rainbow trout (oncorhynchus mykiss) hepatocytes,” *Aquatic Toxicology*, vol. 169, pp. 90–104, Dec. 2015. DOI: 10.1016/j.aquatox.2015.10.004. [Online]. Available: https://doi.org/10.1016/j.aquatox.2015.10.004.

[23] S. Chaturantabut, A. Shwartz, K. J. Evason, A. G. Cox, K. Labella, A. G. Schepers, S. Yang, M. Acuna, Y. Houvras, L. Mancio-Silva, S. Romano, D. A. Gorelick, D. E. Cohen, L. I. Zon, S. N. Bhatia, T. E. North, and W. Goessling, “Estrogen activation of g-protein–coupled estrogen receptor 1 regulates phosphoinositide 3-kinase and mTOR signaling to promote liver growth in zebrafish and proliferation of human hepatocytes,” *Gastroenterology*, vol. 156, no. 6, 1788–1804.e13, May 2019. DOI: 10.1053/j.gastro.2019.01.010. [Online]. Available: https://doi.org/10.1053/j.gastro.2019.01.010.

[24] T. M. Menchaca, C. Juarez-Portilla, and R. C. Zepeda, “Past, present, and future of molecular docking,” in *Drug Discovery and Development - New Advances*, IntechOpen, Mar. 2020. DOI: 10.5772/

intechopen.90921. [Online]. Available: https://doi.org/10.5772/intechopen.90921.

[25] T. J. A. Ewing and I. D. Kuntz, "Critical evaluation of search algorithms for automated molecular docking and database screening," *Journal of Computational Chemistry*, vol. 18, no. 9, pp. 1175–1189, Jul. 1997. DOI: 10.1002/(sici)1096-987x(19970715)18:9<1175::aid-jcc6>3.0.co;2-o. [Online]. Available: https://doi.org/10.1002/(sici)1096-987x(19970715)18:9%3C1175::aid-jcc6%3E3.0.co;2-o.

[26] R. L. DesJarlais, G. L. Seibel, I. D. Kuntz, P. S. Furth, J. C. Alvarez, P. R. Ortiz de Montellano, D. L. DeCamp, L. M. Babé, and C. S. Craik, "Structure-based design of nonpeptide inhibitors specific for the human immunodeficiency virus 1 protease," *Proceedings of the National Academy of Sciences*, vol. 87, no. 17, pp. 6644–6648, 1990, ISSN: 0027-8424. eprint: https://www.pnas.org/content/87/17/6644.full.pdf. [Online]. Available: https://www.pnas.org/content/87/17/6644.

[27] C. Venkatachalam, X. Jiang, T. Oldfield, and M. Waldman, "LigandFit: A novel method for the shape-directed rapid docking of ligands to protein active sites," *Journal of Molecular Graphics and Modelling*, vol. 21, no. 4, pp. 289–307, Jan. 2003. DOI: 10.1016/s1093-3263(02)00164-x. [Online]. Available: https://doi.org/10.1016/s1093-3263(02)00164-x.

[28] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, and P. S. Shenkin, "Glide: a new approach for rapid, accurate docking and scoring. 1. method and assessment of docking accuracy," *Journal of Medicinal Chemistry*, vol. 47, no. 7, pp. 1739–1749, Mar. 2004. DOI: 10.1021/jm0306430. [Online]. Available: https://doi.org/10.1021/jm0306430.

[29] W. Welch, J. Ruppert, and A. N. Jain, "Hammerhead: Fast, fully automated docking of flexible ligands to protein binding sites," *Chemistry & biology*, vol. 3, no. 6, pp. 449–462, 1996.

[30] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, and A. J. Olson, "Autodock4 and autodocktools4: Automated docking with selective receptor flexibility," eng, *Journal of computational chemistry*, vol. 30, no. 16, pp. 2785–2791, Dec. 2009, PMC2760638[pmcid], ISSN: 1096-987X. DOI: 10.1002/jcc.21256. [Online]. Available: https://doi.org/10.1002/jcc.21256.

[31] D. S. Goodsell and A. J. Olson, "Automated docking of substrates to proteins by simulated annealing," *Proteins: Structure, Function, and Genetics*, vol. 8, no. 3, pp. 195–202, 1990. DOI: 10.1002/prot.340080302. [Online]. Available: https://doi.org/10.1002/prot.340080302.

[32] O. Trott and A. J. Olson, "Autodock vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading," eng, *Journal of computational chemistry*, vol. 31, no. 2, pp. 455–461, Jan. 2010, PMC3041641[pmcid], ISSN: 1096-987X. DOI: 10.1002/jcc.21334. [Online]. Available: https://doi.org/10.1002/jcc.21334.

[33] R. Wang, L. Lai, and S. Wang, "Further development and validation of empirical scoring functions for structure-based binding affinity prediction," *Journal of Computer-Aided Molecular Design*, vol. 16, no. 1, pp. 11–26, Jan. 2002, ISSN: 1573-4951. DOI: 10.1023/A:1016357811882. [Online]. Available: https://doi.org/10.1023/A:1016357811882.

[34] M. Grunstein and D. S. Hogness, "Colony hybridization: A method for the isolation of cloned DNAs that contain a specific gene.," *Proceedings of the National Academy of Sciences*, vol. 72, no. 10, pp. 3961–3965, Oct. 1975. DOI: 10.1073/pnas.72.10.3961. [Online]. Available: https://doi.org/10.1073/pnas.72.10.3961.

[35] R. Bumgarner, "Overview of dna microarrays: Types, applications, and their future," *Current Protocols in Molecular Biology*, vol. 101, no. 1, pp. 22.1.1–22.1.11, 2013. DOI: 10.1002/0471142727.mb2201s101. eprint: https://currentprotocols.onlinelibrary.wiley.com/doi/pdf/10.1002/0471142727.mb2201s101.

[Online]. Available: https://currentprotocols.onlinelibrary.wiley.com/doi/abs/10.1002/0471142727.mb2201s101.

[36] S. C. Sealfon and T. T. Chu, "RNA and DNA microarrays," in *Methods in Molecular Biology*, Humana Press, Sep. 2010, pp. 3–34. DOI: 10.1007/978-1-59745-551-0_1. [Online]. Available: https://doi.org/10.1007/978-1-59745-551-0_1.

[37] A. J. Trachtenberg, J.-H. Robert, A. E. Abdalla, A. Fraser, S. Y. He, J. N. Lacy, C. Rivas-Morello, A. Truong, G. Hardiman, L. Ohno-Machado, F. Liu, E. Hovig, and W. P. Kuo, "A primer on the current state of microarray technologies," in *Next Generation Microarray Bioinformatics*, Humana Press, Nov. 2011, pp. 3–17. DOI: 10.1007/978-1-61779-400-1_1. [Online]. Available: https://doi.org/10.1007/978-1-61779-400-1_1.

[38] C. R. Bhasker and G. Hardiman, "Advances in pharmacogenomics technologies," *Pharmacogenomics*, vol. 11, no. 4, pp. 481–485, Apr. 2010. DOI: 10.2217/pgs.10.10. [Online]. Available: https://doi.org/10.2217/pgs.10.10.

[39] G. Hardiman, "Applications of microarrays and biochips in pharmacogenomics," in *Methods in Molecular Biology™*, Humana Press, 2008, pp. 21–30. DOI: 10.1007/978-1-59745-205-2_2. [Online]. Available: https://doi.org/10.1007/978-1-59745-205-2_2.

[40] ——, "Microarray technology–advances, applications, future prospects," *Pharmacogenomics*, vol. 8, no. 12, pp. 1639–1642, Dec. 2007. DOI: 10.2217/14622416.8.12.1639. [Online]. Available: https://doi.org/10.2217/14622416.8.12.1639.

[41] I. Wick and G. Hardiman, "Biochip platforms as functional genomics tools for drug discovery," *Current opinion in drug discovery & development*, vol. 8, pp. 347–54, Jun. 2005.

[42] G. Hardiman, "Microarray platforms–comparisons and contrasts," *Pharmacogenomics*, vol. 5, no. 5, pp. 487–502, Jul. 2004. DOI: 10.1517/14622416.5.5.487. [Online]. Available: https://doi.org/10.1517/14622416.5.5.487.

[43] N. Palanisamy, "Introduction to microarray technology," in *Predictive Biomarkers in Oncology*, Springer International Publishing, Dec. 2018, pp. 75–85. DOI: 10.1007/978-3-319-95228-4_6. [Online]. Available: https://doi.org/10.1007/978-3-319-95228-4_6.

[44] P. Braun and A.-C. Gingras, "History of protein-protein interactions: From egg-white to complex networks," *PROTEOMICS*, vol. 12, no. 10, pp. 1478–1498, May 2012. DOI: 10.1002/pmic.201100563. [Online]. Available: https://doi.org/10.1002/pmic.201100563.

[45] B. T. Palmisano, L. Zhu, and J. M. Stafford, "Role of estrogens in the regulation of liver lipid metabolism," in *Sex and Gender*

Factors Affecting Metabolic Homeostasis, Diabetes and Obesity, F. Mauvais-Jarvis, Ed. Cham: Springer International Publishing, 2017, pp. 227–256, ISBN: 978-3-319-70178-3. DOI: 10.1007/978-3-319-70178-3_12. [Online]. Available: https://doi.org/10.1007/978-3-319-70178-3_12.

[46] E. G. Notch, D. M. Miniutti, and G. D. Mayer, "17alpha-ethinylestradiol decreases expression of multiple hepatic nucleotide excision repair genes in zebrafish (danio rerio)," *Aquatic Toxicology*, vol. 84, no. 3, pp. 301–309, Oct. 2007. DOI: 10.1016/j.aquatox.2007.06.006. [Online]. Available: https://doi.org/10.1016/j.aquatox.2007.06.006.

[47] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang, and E. E. Bolton, "PubChem 2019 update: Improved access to chemical data," *Nucleic Acids Res.*, vol. 47, no. D1, pp. D1102–D1109, Jan. 2019.

[48] N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, and G. R. Hutchison, "Open babel: An open chemical toolbox," *Journal of Cheminformatics*, vol. 3, no. 1, p. 33, Oct. 2011.

[49] A. M. Brzozowski, A. C. W. Pike, Z. Dauter, R. E. Hubbard, T. Bonn, O. Engström, L. Öhman, G. L. Greene, J.-Å. Gustafsson, and M. Carlquist, "Molecular basis of agonism and antagonism

in the oestrogen receptor," *Nature*, vol. 389, no. 6652, pp. 753–758, Oct. 1997, ISSN: 1476-4687. DOI: 10.1038/39645. [Online]. Available: https://doi.org/10.1038/39645.

[50] R. E. Mewshaw, S. M. Bowen, H. A. Harris, Z. B. Xu, E. S. Manas, and S. T. Cohn, "ERb ligands. part 5: Synthesis and structure–activity relationships of a series of 4-hydroxyphenyl-aryl-carbaldehyde oxime derivatives," *Bioorganic and Medicinal Chemistry Letters*, vol. 17, no. 4, pp. 902–906, 2007, ISSN: 0960-894X. DOI: https://doi.org/10.1016/j.bmcl.2006.11.066. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0960894X06013540.

[51] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin, "Ucsf chimera—a visualization system for exploratory research and analysis," *Journal of Computational Chemistry*, vol. 25, no. 13, pp. 1605–1612, 2004. DOI: 10.1002/jcc.20084. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.20084. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.20084.

[52] T. F. Vieira and S. F. Sousa, "Comparing autodock and vina in ligand/decoy discrimination for virtual screening," *Applied Sciences*, vol. 9, no. 21, p. 4538, Oct. 2019, ISSN: 2076-3417. DOI: 10.

3390/app9214538. [Online]. Available: http://dx.doi.org/10.3390/app9214538.

[53] N. T. Nguyen, T. H. Nguyen, T. N. H. Pham, N. T. Huy, M. V. Bay, M. Q. Pham, P. C. Nam, V. V. Vu, and S. T. Ngo, "Autodock vina adopts more accurate binding poses but autodock4 forms better binding affinity," *Journal of Chemical Information and Modeling*, vol. 60, no. 1, pp. 204–211, 2020, PMID: 31887035. DOI: 10.1021/acs.jcim.9b00778. eprint: https://doi.org/10.1021/acs.jcim.9b00778. [Online]. Available: https://doi.org/10.1021/acs.jcim.9b00778.

[54] R. A. Laskowski and M. B. Swindells, "Ligplot+ : Multiple ligand-protein interaction diagrams for drug discovery," *Journal of Chemical Information and Modeling*, vol. 51, no. 10, pp. 2778–2786, Oct. 2011. DOI: 10.1021/ci200227u. [Online]. Available: https://doi.org/10.1021/ci200227u.

[55] Y. Igarashi, N. Nakatsu, T. Yamashita, A. Ono, Y. Ohno, T. Urushidani, and H. Yamada, "Open tg-gates: A large-scale toxicogenomics database," eng, *Nucleic acids research*, vol. 43, no. Database issue, pp. D921–D927, Jan. 2015, gku955[PII], ISSN: 1362-4962. DOI: 10.1093/nar/gku955. [Online]. Available: https://doi.org/10.1093/nar/gku955.

[56] D. Gusenleitner, S. S. Auerbach, T. Melia, H. F. Gómez, D. H. Sherr, and S. Monti, "Genomic models of short-term exposure ac-

curately predict long-term chemical carcinogenicity and identify putative mechanisms of action," eng, *PloS one*, vol. 9, no. 7, e102579–e102579, Jul. 2014, PONE-D-14-11756[PII], ISSN: 1932-6203. DOI: 10.1371/journal.pone.0102579. [Online]. Available: https://doi.org/10.1371/journal.pone.0102579.

[57] R Core Team, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, 2017. [Online]. Available: https://www.R-project.org/.

[58] RStudio Team, *Rstudio: Integrated development environment for r*, RStudio, PBC., Boston, MA, 2020. [Online]. Available: http://www.rstudio.com/.

[59] W. Huber, V. J. Carey, R. Gentleman, S. Anders, M. Carlson, B. S. Carvalho, H. C. Bravo, S. Davis, L. Gatto, T. Girke, R. Gottardo, F. Hahne, K. D. Hansen, R. A. Irizarry, M. Lawrence, M. I. Love, J. MacDonald, V. Obenchain, A. K. Oleś, H. Pagès, A. Reyes, P. Shannon, G. K. Smyth, D. Tenenbaum, L. Waldron, and M. Morgan, "Orchestrating high-throughput genomic analysis with bioconductor," *Nature Methods*, vol. 12, no. 2, pp. 115–121, Feb. 2015, ISSN: 1548-7105. DOI: 10.1038/nmeth.3252. [Online]. Available: https://doi.org/10.1038/nmeth.3252.

[60] R. C. Gentleman, V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M.

Maechler, A. J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J. Y. Yang, and J. Zhang, "Bioconductor: Open software development for computational biology and bioinformatics," *Genome Biology*, vol. 5, no. 10, R80, Sep. 2004, ISSN: 1474-760X. DOI: 10.1186/gb-2004-5-10-r80. [Online]. Available: https://doi.org/10.1186/gb-2004-5-10-r80.

[61] L. Gautier, L. Cope, B. M. Bolstad, and R. A. Irizarry, "Affy—analysis of affymetrix genechip data at the probe level," *Bioinformatics*, vol. 20, no. 3, pp. 307–315, 2004, ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btg405.

[62] M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, and G. K. Smyth, "limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, no. 7, e47–e47, Jan. 2015, ISSN: 0305-1048. DOI: 10.1093/nar/gkv007. eprint: https://academic.oup.com/nar/article-pdf/43/7/e47/7207289/gkv007.pdf. [Online]. Available: https://doi.org/10.1093/nar/gkv007.

[63] W. Huber, V. J. Carey, R. Gentleman, S. Anders, M. Carlson, B. S. Carvalho, H. C. Bravo, S. Davis, L. Gatto, T. Girke, R. Gottardo, F. Hahne, K. D. Hansen, R. A. Irizarry, M. Lawrence, M. I. Love, J. MacDonald, V. Obenchain, A. K. Ole's, H. Pag'es, A. Reyes, P. Shannon, G. K. Smyth, D. Tenenbaum, L. Waldron, and M. Morgan, "Orchestrating high-throughput genomic analysis with Bio-

conductor," *Nature Methods*, vol. 12, no. 2, pp. 115–121, 2015. [Online]. Available: http://www.nature.com/nmeth/journal/v12/n2/full/nmeth.3252.html.

[64] U. Raudvere, L. Kolberg, I. Kuzmin, T. Arak, P. Adler, H. Peterson, and J. Vilo, "g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update)," *Nucleic Acids Research*, vol. 47, no. W1, W191–W198, May 2019, ISSN: 0305-1048. DOI: 10.1093/nar/gkz369. eprint: https://academic.oup.com/nar/article-pdf/47/W1/W191/28879887/gkz369.pdf. [Online]. Available: https://doi.org/10.1093/nar/gkz369.

[65] "The gene ontology resource: 20 years and still GOing strong," *Nucleic Acids Research*, vol. 47, no. D1, pp. D330–D338, Nov. 2018. DOI: 10.1093/nar/gky1055. [Online]. Available: https://doi.org/10.1093/nar/gky1055.

[66] S. Ahsan and S. Drăghici, "Identifying significantly impacted pathways and putative mechanisms with ipathwayguide," *Current Protocols in Bioinformatics*, vol. 57, no. 1, pp. 7.15.1–7.15.30, 2017. DOI: 10.1002/cpbi.24. eprint: https://currentprotocols.onlinelibrary.wiley.com/doi/pdf/10.1002/cpbi.24. [Online]. Available: https://currentprotocols.onlinelibrary.wiley.com/doi/abs/10.1002/cpbi.24.

[67] A. Alexa, J. Rahnenführer, and T. Lengauer, "Improved scoring of functional groups from gene expression data by decorrelating GO graph structure," *Bioinformatics*, vol. 22, no. 13, pp. 1600–1607, Apr. 2006, ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btl140. eprint: https://academic.oup.com/bioinformatics/article-pdf/22/13/1600/482171/btl140.pdf. [Online]. Available: https://doi.org/10.1093/bioinformatics/btl140.

[68] D. Szklarczyk, A. L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N. T. Doncheva, J. H. Morris, P. Bork, L. J. Jensen, and C. v. Mering, "String v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets," eng, *Nucleic acids research*, vol. 47, no. D1, pp. D607–D613, Jan. 2019, 5198476[PII], ISSN: 1362-4962. DOI: 10.1093/nar/gky1131. [Online]. Available: https://doi.org/10.1093/nar/gky1131.

[69] P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker, "Cytoscape: A software environment for integrated models of biomolecular interaction networks," eng, *Genome research*, vol. 13, no. 11, pp. 2498–2504, Nov. 2003, 13/11/2498[PII], ISSN: 1088-9051. DOI: 10.1101/gr.1239303. [Online]. Available: https://doi.org/10.1101/gr.1239303.

[70] S. A. Ochsner, D. Abraham, K. Martin, W. Ding, A. McOwiti, W. Kankanamge, Z. Wang, K. Andreano, R. A. Hamilton, Y. Chen, A. Hamilton, M. L. Gantner, M. Dehart, S. Qu, S. G. Hilsenbeck, L. B. Becnel, D. Bridges, A. Ma'ayan, J. M. Huss, F. Stossi, C. E. Foulds, A. Kralli, D. P. McDonnell, and N. J. McKenna, "The signaling pathways project, an integrated 'omics knowledgebase for mammalian cellular signaling pathways," *Scientific Data*, vol. 6, no. 1, p. 252, Oct. 2019, ISSN: 2052-4463. DOI: 10.1038/s41597-019-0193-4. [Online]. Available: https://doi.org/10.1038/s41597-019-0193-4.

[71] X. Lei, S. Wang, and F. Wu, "Identification of essential proteins based on improved hits algorithm," eng, *Genes*, vol. 10, no. 2, p. 177, Feb. 2019, genes10020177[PII], ISSN: 2073-4425. DOI: 10.3390/genes10020177. [Online]. Available: https://doi.org/10.3390/genes10020177.

[72] Y. Tang, M. Li, J. Wang, Y. Pan, and F.-X. Wu, "Cytonca: A cytoscape plugin for centrality analysis and evaluation of protein interaction networks," *Biosystems*, vol. 127, pp. 67–72, 2015, ISSN: 0303-2647. DOI: https://doi.org/10.1016/j.biosystems.2014.11.005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0303264714001944.

[73] A. F. Pierucci, *Rmngb: Miscellaneous collection of functions for medical data analysis*, R package version 0.6-1, 2014. [Online]. Available: `https://CRAN.R-project.org/package=rmngb`.

[74] C. Ismay and A. Kim, *Statistical Inference via Data Science: A ModernDive into R and the Tidyverse*. Dec. 2019, ISBN: 9781000763461. DOI: `10.1201/9780367409913`.

[75] The GIMP Development Team, *Gimp*, version 2.10.12, Jun. 12, 2019. [Online]. Available: `https://www.gimp.org`.

[76] C. A. Heinrich, M. R. Lail-Trecker, J. J. Peluso, and B. A. White, "Negative regulation of n-cadherin-mediated cell-cell adhesion by the estrogen receptor signaling pathway in rat pituitary gh3 cells," *Endocrine*, vol. 10, no. 1, pp. 67–76, Feb. 1999, ISSN: 1559-0100. DOI: `10.1385/ENDO:10:1:67`. [Online]. Available: `https://doi.org/10.1385/ENDO:10:1:67`.

[77] N. A. Gloushankova, "Changes in regulation of cell—cell adhesion during tumor transformation," *Biochemistry (Moscow)*, vol. 73, no. 7, pp. 742–750, Jul. 2008, ISSN: 1608-3040. DOI: `10.1134/S000629790807002X`. [Online]. Available: `https://doi.org/10.1134/S000629790807002X`.

[78] Y. N. Demchenko, O. K. Glebov, A. Zingone, J. J. Keats, P. L. Bergsagel, and W. M. Kuehl, "Classical and/or alternative nf-kappab pathway activation in multiple myeloma," eng, *Blood*, vol. 115, no. 17, pp. 3541–3552, Apr. 2010, blood-2009-09-243535[PII],

ISSN: 1528-0020. DOI: 10.1182/blood-2009-09-243535. [Online]. Available: https://doi.org/10.1182/blood-2009-09-243535.

[79] R. A. Ansari and J. Gandy, "Determining the transrepression activity of xenoestrogen on nuclear factor-kappab in cos-1 cells by estrogen receptor-a," *International Journal of Toxicology*, vol. 26, no. 5, pp. 441–449, 2007, PMID: 17963131. DOI: 10.1080/10915810701620317. eprint: https://doi.org/10.1080/10915810701620317. [Online]. Available: https://doi.org/10.1080/10915810701620317.

[80] D. K. Biswas, S. Singh, Q. Shi, A. B. Pardee, and J. D. Iglehart, "Crossroads of estrogen receptor and nf-kappab signaling," *Science Signaling*, vol. 2005, no. 288, pe27–pe27, 2005. DOI: 10.1126/stke.2882005pe27. [Online]. Available: https://stke.sciencemag.org/content/2005/288/pe27.

[81] T. Zhan, N. Rindtorff, and M. Boutros, "Wnt signaling in cancer," *Oncogene*, vol. 36, no. 11, pp. 1461–1473, Mar. 2017, ISSN: 1476-5594. DOI: 10.1038/onc.2016.304. [Online]. Available: https://doi.org/10.1038/onc.2016.304.

[82] S. Liu, W. Fan, X. Gao, K. Huang, C. Ding, G. Ma, L. Yan, and S. Song, "Estrogen receptor alpha regulates the wnt/b-catenin signaling pathway in colon cancer by targeting the nod-like receptors," *Cellular Signalling*, vol. 61, pp. 86–92, 2019, ISSN: 0898-6568. DOI: https://doi.org/10.1016/j.cellsig.2019.05.

009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0898656819301081.

[83] N. Bhagavan and C.-E. Ha, "Chapter 27 - metabolism of iron and heme," in *Essentials of Medical Biochemistry (Second Edition)*, N. Bhagavan and C.-E. Ha, Eds., Second Edition, San Diego: Academic Press, 2015, pp. 511–529, ISBN: 978-0-12-416687-5. DOI: https://doi.org/10.1016/B978-0-12-416687-5.00027-0. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780124166875000270.

[84] M. Reddy, K. Bachhawat-Sikder, and T. Kodadek, "Transformation of low-affinity lead compounds into high-affinity protein capture agents," *Chemistry & Biology*, vol. 11, no. 8, pp. 1127–1137, Aug. 2004. DOI: 10.1016/j.chembiol.2004.05.013. [Online]. Available: https://doi.org/10.1016/j.chembiol.2004.05.013.

[85] A. N. Koehler, A. F. Shamji, and S. L. Schreiber, "Discovery of an inhibitor of a transcription factor using small molecule microarrays and diversity-oriented synthesis," *Journal of the American Chemical Society*, vol. 125, no. 28, pp. 8420–8421, Jul. 2003. DOI: 10.1021/ja0352698. [Online]. Available: https://doi.org/10.1021/ja0352698.

[86] N. Fuentes and P. Silveyra, "Chapter three - estrogen receptor signaling mechanisms," in *Intracellular Signalling Proteins*, ser. Ad-

vances in Protein Chemistry and Structural Biology, R. Donev, Ed., vol. 116, Academic Press, 2019, pp. 135–170. DOI: https://doi.org/10.1016/bs.apcsb.2019.01.001. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1876162319300112.

[87] H. Yan, Z. Li, Q. Shen, Q. Wang, J. Tian, Q. Jiang, and L. Gao, "Aberrant expression of cell cycle and material metabolism related genes contributes to hepatocellular carcinoma occurrence," *Pathology - Research and Practice*, vol. 213, no. 4, pp. 316–321, 2017, ISSN: 0344-0338. DOI: https://doi.org/10.1016/j.prp.2017.01.019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0344033817300729.

[88] Y. Cao, W. Ding, J. Zhang, Q. Gao, H. Yang, W. Cao, Z. Wang, L. Fang, and R. Du, "Significant down-regulation of urea cycle generates clinically relevant proteomic signature in hepatocellular carcinoma patients with macrovascular invasion," *Journal of Proteome Research*, vol. 18, no. 5, pp. 2032–2044, May 2019, ISSN: 1535-3893. DOI: 10.1021/acs.jproteome.8b00921. [Online]. Available: https://doi.org/10.1021/acs.jproteome.8b00921.

[89] S.-H. Hong, H.-Y. Nah, and C.-H. Kim, "Estrogen-regulated expression and distribution of id-1 in the mouse uterus," *Placenta*, vol. 31, no. 3, pp. 240–244, 2010, ISSN: 0143-4004. DOI: https://doi.org/10.1016/j.placenta.2009.11.015. [On-

line]. Available: http://www.sciencedirect.com/science/article/pii/S0143400409003786.

[90] T. Schlott, W. Thasler, C. Gorzel, S. Pahernik, U. Brinck, H. Eiffert, and M. Droese, "Detection of MDM2 alterations in cultured human hepatocytes treated with 17beta-estradiol or 17alpha-ethinylestradiol," eng, *Anticancer Research*, vol. 22, no. 3, pp. 1545–1551, Jun. 2002, ISSN: 0250-7005.

[91] L. Huang, B. Hu, J. Ni, J. Wu, W. Jiang, C. Chen, L. Yang, Y. Zeng, R. Wan, G. Hu, and X. Wang, "Transcriptional repression of socs3 mediated by il-6/stat3 signaling via dnmt1 promotes pancreatic cancer growth and metastasis," *Journal of Experimental & Clinical Cancer Research*, vol. 35, no. 1, p. 27, Feb. 2016, ISSN: 1756-9966. DOI: 10.1186/s13046-016-0301-7. [Online]. Available: https://doi.org/10.1186/s13046-016-0301-7.

[92] W.-Y. Wu, H. Kim, C.-L. Zhang, X.-L. Meng, and Z.-S. Wu, "Loss of suppressors of cytokine signaling 3 promotes aggressiveness in hepatocellular carcinoma," *Journal of Investigative Surgery*, vol. 27, no. 4, pp. 197–204, 2014. DOI: 10.3109/08941939.2013.873098. eprint: https://doi.org/10.3109/08941939.2013.873098. [Online]. Available: https://doi.org/10.3109/08941939.2013.873098.

[93] H. Ogata, T. Chinen, T. Yoshida, I. Kinjyo, G. Takaesu, H. Shiraishi, M. Iida, T. Kobayashi, and A. Yoshimura, "Loss of socs3 in

the liver promotes fibrosis by enhancing stat3-mediated tgf-beta1 production," *Oncogene*, vol. 25, no. 17, pp. 2520–2530, Apr. 2006, ISSN: 1476-5594. DOI: 10.1038/sj.onc.1209281. [Online]. Available: https://doi.org/10.1038/sj.onc.1209281.

[94] J. Su, P. You, J.-P. Zhao, S.-L. Zhang, S.-H. Song, Z.-R. Fu, L.-W. Ye, X.-Y. Zi, D.-F. Xie, M.-H. Zhu, and Y.-P. Hu, "A potential role for the homeoprotein hhex in hepatocellular carcinoma progression," *Medical Oncology*, vol. 29, no. 2, pp. 1059–1067, Jun. 2012, ISSN: 1559-131X. DOI: 10.1007/s12032-011-9989-6. [Online]. Available: https://doi.org/10.1007/s12032-011-9989-6.

[95] Z. Li, Q. Zhang, Q. Zhang, M. Xu, Y. Qu, X. Cai, and L. Lu, "CXCL6 promotes human hepatocyte proliferation through the CXCR1–NFkap pathway and inhibits collagen i secretion by hepatic stellate cells," *Biochemistry and Cell Biology*, vol. 94, no. 3, pp. 229–235, Jun. 2016. DOI: 10.1139/bcb-2015-0136. [Online]. Available: https://doi.org/10.1139/bcb-2015-0136.

[96] H. Tian, P. Huang, Z. Zhao, W. Tang, and J. Xia, "HIF-1a plays a role in the chemotactic migration of hepatocarcinoma cells through the modulation of CXCL6 expression," *Cellular Physiology and Biochemistry*, vol. 34, no. 5, pp. 1536–1546, 2014. DOI: 10.1159/000366357. [Online]. Available: https://doi.org/10.1159/000366357.

[97] X. Cai, Z. Li, Q. Zhang, Y. Qu, M. Xu, X. Wan, and L. Lu, "CXCL6-EGFR-induced kupffer cells secrete TGF-beta1 promoting hepatic stellate cell activation via the SMAD2/BRD4/c-MYC/EZH2 pathway in liver fibrosis," *Journal of Cellular and Molecular Medicine*, vol. 22, no. 10, pp. 5050–5061, Aug. 2018. DOI: 10.1111/jcmm.13787. [Online]. Available: https://doi.org/10.1111/jcmm.13787.

[98] F. Mußbach, H. Ungefroren, B. Günther, K. Katenkamp, P. Henklein, M. Westermann, U. Settmacher, L. Lenk, S. Sebens, J. P. Müller, F.-D. Böhmer, and R. Kaufmann, "Proteinase-activated receptor 2 (PAR2) in hepatic stellate cells – evidence for a role in hepatocellular carcinoma growth in vivo," *Molecular Cancer*, vol. 15, no. 1, Jul. 2016. DOI: 10.1186/s12943-016-0538-y. [Online]. Available: https://doi.org/10.1186/s12943-016-0538-y.

[99] A. E. Dhamad, Z. Zhou, J. Zhou, and Y. Du, "Systematic proteomic identification of the heat shock proteins (hsp) that interact with estrogen receptor alpha (ERalpha) and biochemical characterization of the ERalpha-hsp70 interaction," *PLOS ONE*, vol. 11, no. 8, D. Picard, Ed., e0160312, Aug. 2016. DOI: 10.1371/journal.pone.0160312. [Online]. Available: https://doi.org/10.1371/journal.pone.0160312.

[100] D. Xing, S. Oparil, H. Yu, K. Gong, W. Feng, J. Black, Y.-F. Chen, and S. Nozell, "Estrogen modulates NFkappab signaling by en-

hancing ikappa balpha levels and blocking p65 binding at the promoters of inflammatory genes via estrogen receptor-beta," *PLoS ONE*, vol. 7, no. 6, S. Kovats, Ed., e36890, Jun. 2012. DOI: 10.1371/journal.pone.0036890. [Online]. Available: https://doi.org/10.1371/journal.pone.0036890.

[101] N. H. Rogers, C. A. Witczak, M. F. Hirshman, L. J. Goodyear, and A. S. Greenberg, "Estradiol stimulates akt, AMP-activated protein kinase (AMPK) and TBC1d1/4, but not glucose uptake in rat soleus," *Biochemical and Biophysical Research Communications*, vol. 382, no. 4, pp. 646–650, May 2009. DOI: 10.1016/j.bbrc.2009.02.154. [Online]. Available: https://doi.org/10.1016/j.bbrc.2009.02.154.

[102] L. V. Madrid, C.-Y. Wang, D. C. Guttridge, A. J. G. Schottelius, A. S. Baldwin, and M. W. Mayo, "Akt suppresses apoptosis by stimulating the transactivation potential of the RelA/p65 subunit of NF-kappab," *Molecular and Cellular Biology*, vol. 20, no. 5, pp. 1626–1638, Mar. 2000. DOI: 10.1128/mcb.20.5.1626-1638.2000. [Online]. Available: https://doi.org/10.1128/mcb.20.5.1626-1638.2000.

[103] A. L. Creekmore, K. A. Walt, J. R. Schultz-Norton, Y. S. Ziegler, I. X. McLeod, J. R. Yates, and A. M. Nardulli, "The role of retinoblastoma-associated proteins 46 and 48 in estrogen receptor alpha mediated gene expression," *Molecular and Cellular Endocrinology*, vol. 291,

no. 1-2, pp. 79–86, Sep. 2008. DOI: 10.1016/j.mce.2008.05.016. [Online]. Available: https://doi.org/10.1016/j.mce.2008.05.016.

[104] H. Song, "Genes encoding pir51, beclin 1, RbAp48 and aldolase b are up or down-regulated in human primary hepatocellular carcinoma," *World Journal of Gastroenterology*, vol. 10, no. 4, p. 509, 2004. DOI: 10.3748/wjg.v10.i4.509. [Online]. Available: https://doi.org/10.3748/wjg.v10.i4.509.

[105] F. Pacifico, M. Paolillo, G. Chiappetta, E. Crescenzi, S. Arena, A. Scaloni, M. Monaco, C. Vascotto, G. Tell, S. Formisano, and A. Leonardi, "RbAp48 is a target of nuclear factor-kappab activity in thyroid cancer," *The Journal of Clinical Endocrinology & Metabolism*, vol. 92, no. 4, pp. 1458–1466, Apr. 2007. DOI: 10.1210/jc.2006-2199. [Online]. Available: https://doi.org/10.1210/jc.2006-2199.

[106] Y. Guo, T. Song, J. Liu, H. Li, Z. Tian, and C. Xu, "Disruption of NF-&kappab signaling by fluoxetine attenuates MGMT expression in glioma cells," *OncoTargets and Therapy*, p. 2199, Aug. 2015. DOI: 10.2147/ott.s85948. [Online]. Available: https://doi.org/10.2147/ott.s85948.

[107] A. K. C. Teo, H. K. Oh, R. B. Ali, and B. F. L. Li, "The modified human DNA repair EnzymeO6-methylguanine-DNA methyltransferase is a negative regulator of estrogen receptor-mediated

transcription upon alkylation DNA damage," *Molecular and Cellular Biology*, vol. 21, no. 20, pp. 7105–7114, Oct. 2001. DOI: 10.1128/mcb.21.20.7105-7114.2001. [Online]. Available: https://doi.org/10.1128/mcb.21.20.7105-7114.2001.

[108] S. H. Lam, S. G. Lee, C. Y. Lin, J. S. Thomsen, P. Y. Fu, K. R. Murthy, H. Li, K. R. Govindarajan, L. C. Nick, G. Bourque, Z. Gong, T. Lufkin, E. T. Liu, and S. Mathavan, "Molecular conservation of estrogen-response associated with cell cycle regulation, hormonal carcinogenesis and cancer in zebrafish and human cancer cell lines," *BMC Medical Genomics*, vol. 4, no. 1, May 2011. DOI: 10.1186/1755-8794-4-41. [Online]. Available: https://doi.org/10.1186/1755-8794-4-41.

[109] D. Roy, Q. Cai, Q. Felty, and S. Narayan, "Estrogen-induced generation of reactive oxygen and nitrogen species, gene damage, and estrogen-dependent cancers," *Journal of Toxicology and Environmental Health, Part B*, vol. 10, no. 4, pp. 235–257, May 2007. DOI: 10.1080/15287390600974924. [Online]. Available: https://doi.org/10.1080/15287390600974924.

[110] C. Wang, J. A. Mayer, A. Mazumdar, K. Fertuck, H. Kim, M. Brown, and P. H. Brown, "Estrogen induces c-myc gene expression via an upstream enhancer activated by the estrogen receptor and the AP-1 transcription factor," *Molecular Endocrinology*, vol. 25, no. 9, pp. 1527–1538, Sep. 2011. DOI: 10.1210/me.2011-1037.

[Online]. Available: https://doi.org/10.1210/me.2011-1037.

[111] A. S. Cheng, V. X. Jin, M. Fan, L. T. Smith, S. Liyanarachchi, P. S. Yan, Y.-W. Leu, M. W. Chan, C. Plass, K. P. Nephew, R. V. Davuluri, and T. H.-M. Huang, "Combinatorial analysis of transcription factor partners reveals recruitment of c-MYC to estrogen receptor-alpha responsive promoters," *Molecular Cell*, vol. 21, no. 3, pp. 393–404, Feb. 2006. DOI: 10.1016/j.molcel.2005.12.016. [Online]. Available: https://doi.org/10.1016/j.molcel.2005.12.016.

[112] H. Li, J.-W. Lu, X. Huo, Y. Li, Z. Li, and Z. Gong, "Effects of sex hormones on liver tumor progression and regression in myc/xmrk double oncogene transgenic zebrafish," *General and Comparative Endocrinology*, vol. 277, pp. 112–121, Jun. 2019. DOI: 10.1016/j.ygcen.2019.03.018. [Online]. Available: https://doi.org/10.1016/j.ygcen.2019.03.018.

[113] T. Liu, Y. Zhou, K. S. Ko, and H. Yang, "Interactions between myc and mediators of inflammation in chronic liver diseases," *Mediators of Inflammation*, vol. 2015, pp. 1–12, 2015. DOI: 10.1155/2015/276850. [Online]. Available: https://doi.org/10.1155/2015/276850.

[114] M. Maggiolini, A. Vivacqua, G. Fasanella, A. G. Recchia, D. Sisci, V. Pezzi, D. Montanaro, A. M. Musti, D. Picard, and S. Andò, "The

g protein-coupled receptor GPR30 mediates c-fosUp-regulation by 17beta-estradiol and phytoestrogens in breast cancer cells," *Journal of Biological Chemistry*, vol. 279, no. 26, pp. 27 008–27 016, Apr. 2004. DOI: 10.1074/jbc.m403588200. [Online]. Available: https://doi.org/10.1074/jbc.m403588200.

[115] J.-W. Hu, G.-Y. Ding, P.-Y. Fu, W.-G. Tang, Q.-M. Sun, X.-D. Zhu, Y.-H. Shen, J. Zhou, J. Fan, H.-C. Sun, and C. Huang, "Identification of FOS as a candidate risk gene for liver cancer by integrated bioinformatic analysis," *BioMed Research International*, vol. 2020, pp. 1–10, Mar. 2020. DOI: 10.1155/2020/6784138. [Online]. Available: https://doi.org/10.1155/2020/6784138.

[116] S. Jiang, H. Katayama, J. Wang, S. A. Li, Y. Hong, L. Radvanyi, J. J. Li, and S. Sen, "Estrogen-induced aurora kinase-a (AURKA) gene expression is activated by GATA-3 in estrogen receptor-positive breast cancer cells," *Hormones and Cancer*, vol. 1, no. 1, pp. 11–20, Feb. 2010. DOI: 10.1007/s12672-010-0006-x. [Online]. Available: https://doi.org/10.1007/s12672-010-0006-x.

[117] Z.-T. Shen, Y. Chen, G.-C. Huang, X.-X. Zhu, R. Wang, and L.-B. Chen, "Aurora-a confers radioresistance in human hepatocellular carcinoma by activating NF-kappab signaling pathway," *BMC Cancer*, vol. 19, no. 1, Nov. 2019. DOI: 10.1186/s12885-019-6312-y. [Online]. Available: https://doi.org/10.1186/s12885-019-6312-y.

[118] M. Shi, Z. Lin, L. Ye, X. Chen, W. Zhang, Z. Zhang, F. Luo, Y. Liu, and M. Shi, "Estrogen receptor-regulated SOCS3 modulation via JAK2/STAT3 pathway is involved in BPF-induced m1 polarization of macrophages," *Toxicology*, vol. 433-434, p. 152 404, Mar. 2020. DOI: 10.1016/j.tox.2020.152404. [Online]. Available: https://doi.org/10.1016/j.tox.2020.152404.

[119] J. Adams, "The proteasome in cell-cycle regulation," in *Proteasome Inhibitors in Cancer Therapy*, J. Adams, Ed. Totowa, NJ: Humana Press, 2004, pp. 77–84, ISBN: 978-1-59259-794-9. DOI: 10.1007/978-1-59259-794-9_6. [Online]. Available: https://doi.org/10.1007/978-1-59259-794-9_6.

[120] G. Giurato, G. Nassa, A. Salvati, E. Alexandrova, F. Rizzo, T. A. Nyman, A. Weisz, and R. Tarallo, "Quantitative mapping of RNA-mediated nuclear estrogen receptor beta interactome in human breast cancer cells," *Scientific Data*, vol. 5, no. 1, Mar. 2018. DOI: 10.1038/sdata.2018.31. [Online]. Available: https://doi.org/10.1038/sdata.2018.31.

[121] J. S. Foster and J. Wimalasena, "Estrogen regulates activity of cyclyn-dependent kinases and retinoblastoma protein phosphorilation in beast cancer cells.," *Molecular Endocrinology*, vol. 10, no. 5, pp. 488–498, May 1996. DOI: 10.1210/mend.10.5.8732680. [Online]. Available: https://doi.org/10.1210/mend.10.5.8732680.

[122] J. M. G. Pedrero, B. del Rio, C. Martinez-Campa, M. Muramatsu, P. S. Lazo, and S. Ramos, "Calmodulin is a selective modulator of estrogen receptors," *Molecular Endocrinology*, vol. 16, no. 5, pp. 947–960, May 2002. DOI: 10.1210/mend.16.5.0830. [Online]. Available: https://doi.org/10.1210/mend.16.5.0830.

[123] Z. Li, Y. Zhang, A. C. Hedman, J. B. Ames, and D. B. Sacks, "Calmodulin lobes facilitate dimerization and activation of estrogen receptor-alpha," *Journal of Biological Chemistry*, vol. 292, no. 11, pp. 4614–4622, Feb. 2017. DOI: 10.1074/jbc.m116.754804. [Online]. Available: https://doi.org/10.1074/jbc.m116.754804.

Appendices

Chapter A

Interactions

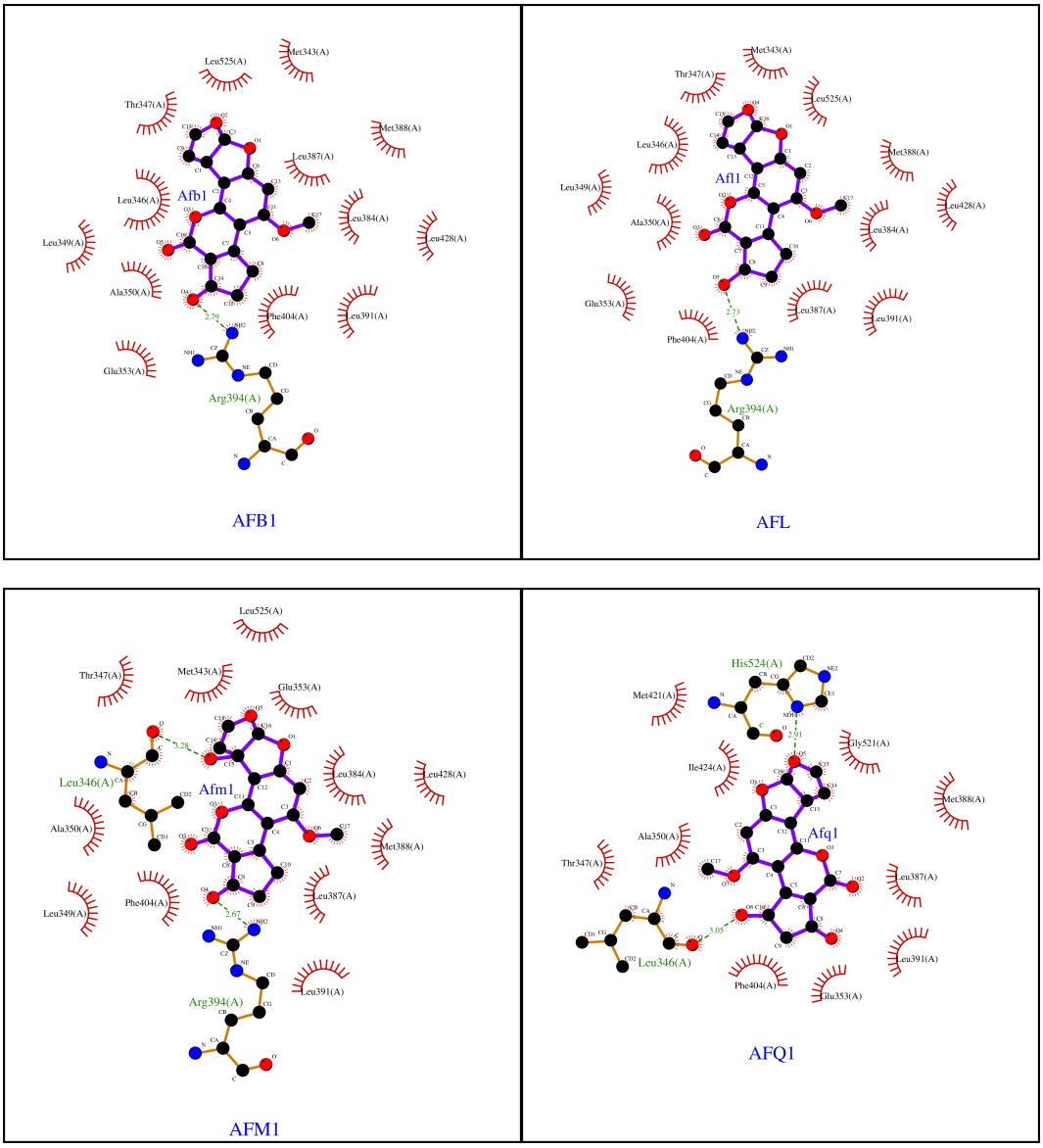

Figure A.1: Interactions occurring among every ligand and ERα residues.

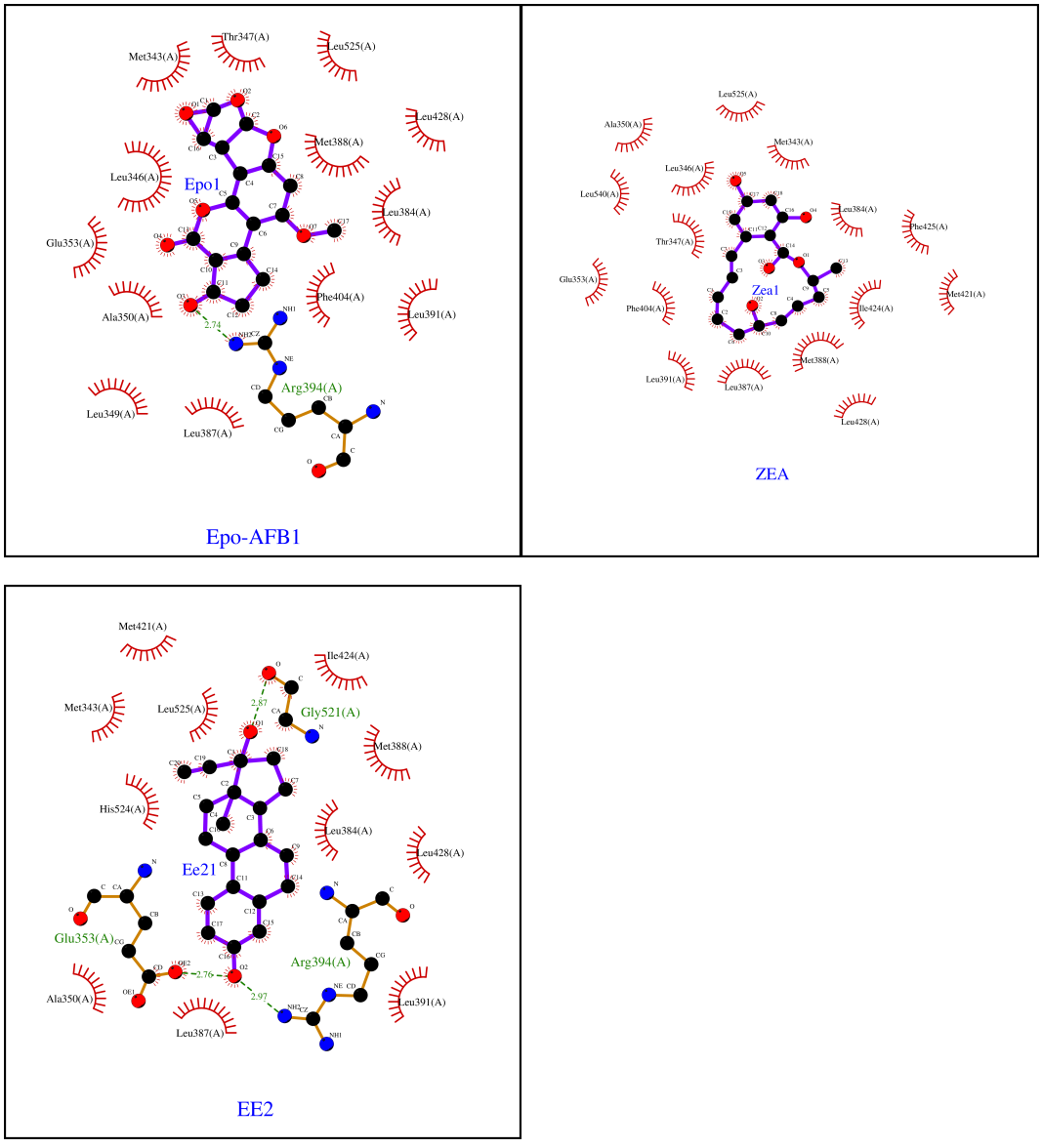

Figure A.2: Interactions occurring among every ligand and ERα residues.

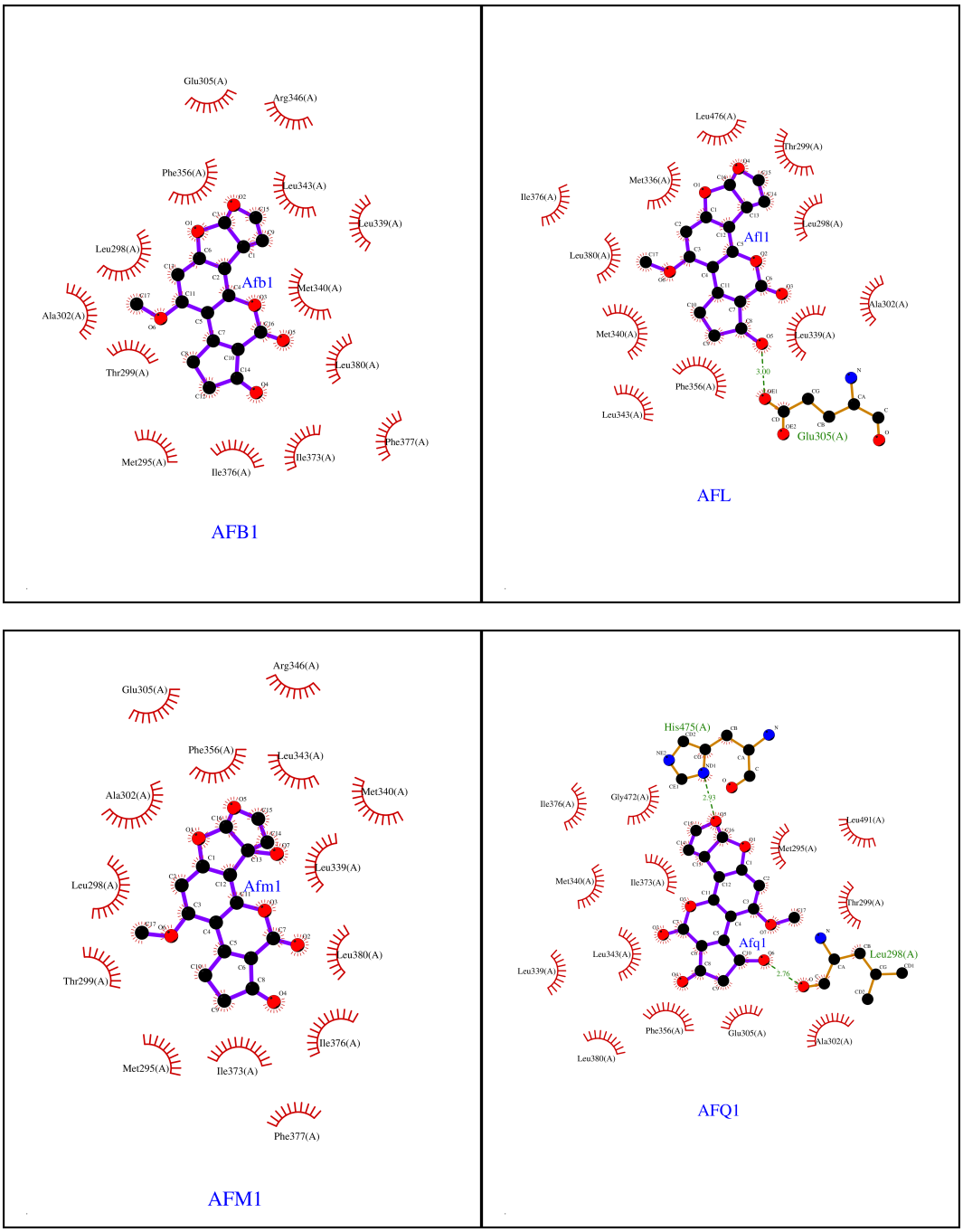

Figure A.3: Interactions occurring among every ligand and ERβ residues.

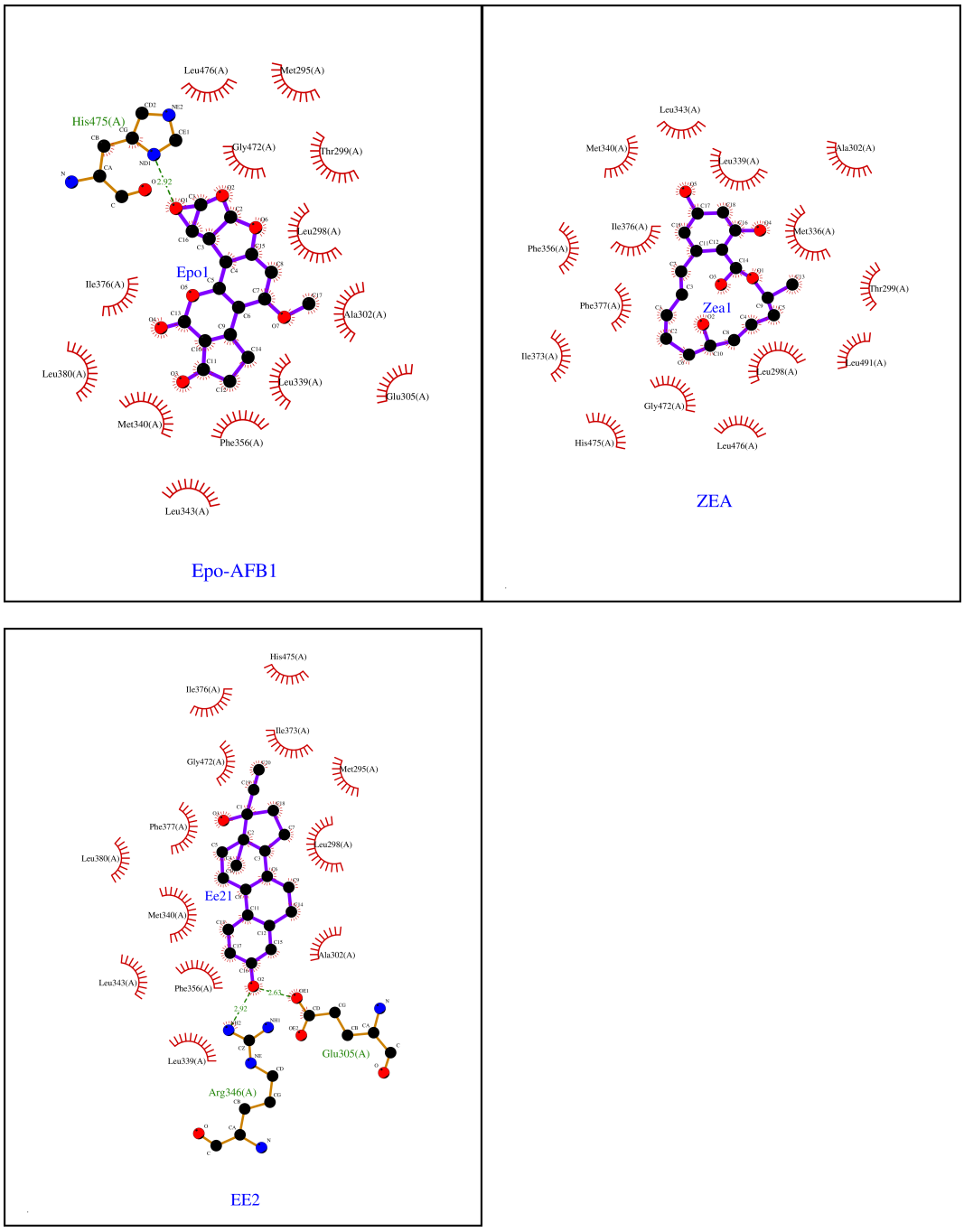

Figure A.4: Interactions occurring among every ligand and ERβ residues.

Chapter B

BP terms

GO unique signature for AFB1 *in vitro* exposure from Advaita	
GO:Biological Process	**Adj.p-value**
activation of MAPK activity	0.00288
activation of protein kinase B activity	0.04078
adrenal gland development	0.01457
amino acid homeostasis	0.01113
apoptotic chromosome condensation	0.03411
asymmetric neuroblast division	0.04908
autocrine signaling	0.04908
B cell proliferation involved in immune response	0.02134
BMP signaling pathway	0.0272
bone trabecula morphogenesis	0.04299
branching morphogenesis of an epithelial tube	0.03274
cell differentiation involved in metanephros development	0.02036
cellular response to amino acid starvation	0.04873
cellular response to biotic stimulus	0.04716
cellular response to cold	0.0152
cellular response to cytokine stimulus	0.02634

cellular response to dsDNA 0.00387
cellular response to glucocorticoid stimulus 0.04488
cellular response to leukemia inhibitory factor 0.00066
cellular response to mineralocorticoid stimulus 0.00696
cellular response to transforming growth factor beta stimulus 0.02711
chondroblast differentiation 0.03411
chylomicron assembly 0.02073
coronary vasculature development 0.02794
dendritic spine maintenance 0.00516
detection of lipopolysaccharide 0.02134
developmental induction 0.04595
DNA damage response, signal transduction by p53 class mediator 0.04854
embryonic heart tube anterior/posterior pattern specification 0.01113
endothelial cell activation 0.02073
epidermal growth factor receptor signaling pathway 0.04317
epithelial cell maturation involved in prostate gland development 0.01113
epithelial cell proliferation involved in renal tubule morphogenesis 0.02134
epithelial to mesenchymal transition involved in cardiac fibroblast development 0.00387
female gonad development 0.02935
fibroblast migration 0.0321
glucosylceramide catabolic process 0.01113
glycogen biosynthetic process 0.03986
growth 0.0473
growth plate cartilage chondrocyte development 0.02134
histone acetylation 0.04243
in utero embryonic development 0.02668
inactivation of MAPK activity 0.00231

insulin metabolic process 0.03411
intestinal epithelial structure maintenance 0.04908
intracellular receptor signaling pathway 0.01517
intrinsic apoptotic signaling pathway in response to oxidative stress 0.03986
lipid hydroxylation 0.03411
lipoprotein catabolic process 0.03463
maintenance of centrosome location 0.04908
male germ-line stem cell asymmetric division 0.01113
mesenchymal cell differentiation 0.04677
metanephric cap mesenchymal cell proliferation involved in
metanephros development 0.00387
middle ear morphogenesis 0.03247
mitochondrial transcription 0.02721
mitotic cell cycle arrest 0.04299
mitotic cell cycle process 0.04565
morphogenesis of a branching epithelium 0.01248
morphogenesis of an epithelial sheet 0.03747
negative regulation of cardiac muscle hypertrophy in response to stress 0.04908
negative regulation of cell migration 0.02797
negative regulation of cell-cell adhesion 0.02123
negative regulation of endoplasmic reticulum stress-induced eIF2
alpha phosphorylation 0.02134
negative regulation of endothelial cell-matrix adhesion via fibronectin 0.00387
negative regulation of epithelial cell migration 0.00471
negative regulation of ERK1 and ERK2 cascade 0.00587
negative regulation of extracellular matrix constituent secretion 0.02134
negative regulation of fat cell differentiation 0.04873
negative regulation of intrinsic apoptotic signaling pathway 0.04637

negative regulation of keratinocyte proliferation 0.00677
negative regulation of microtubule depolymerization 0.0382
negative regulation of multicellular organism growth 0.00346
negative regulation of MyD88-independent toll-like receptor signaling pathway 0.04908
negative regulation of myeloid cell differentiation 0.03055
negative regulation of netrin-activated signaling pathway 0.02134
negative regulation of neurotrophin TRK receptor signaling pathway 0.02134
negative regulation of oxidative stress-induced cell death 0.02021
negative regulation of pathway-restricted SMAD protein phosphorylation 0.00346
negative regulation of peptidyl-serine phosphorylation 0.01215
negative regulation of peptidyl-threonine phosphorylation 0.00305
negative regulation of PERK-mediated unfolded protein response 0.03411
negative regulation of protein kinase activity 0.03037
negative regulation of response to DNA damage stimulus 0.04259
negative regulation of signal transduction by p53 class mediator 0.04078
negative regulation of stress-activated MAPK cascade 0.0069
negative regulation of T-helper 2 cell differentiation 0.02134
negative regulation of transcription by RNA polymerase II 6.9E-05
negative regulation of Wnt signaling pathway involved in heart development 0.02134
neuron death 0.04521
nodal signaling pathway 0.04299
nucleic acid metabolic process 0.02275
phosphatidylethanolamine metabolic process 0.04446
placenta blood vessel development 0.00937
positive regulation of apoptotic process involved in morphogenesis 0.03411

positive regulation of cap-independent translational initiation 0.01113
positive regulation of cardioblast differentiation 0.04908
positive regulation of cardiocyte differentiation 0.00808
positive regulation of cartilage development 0.0079
positive regulation of ceramide biosynthetic process 0.0152
positive regulation of chemokine production 0.04488
positive regulation of DNA damage response, signal transduction by
p53 class mediator 0.00677
positive regulation of ERK1 and ERK2 cascade via
TNFSF11-mediated signaling 0.00387
positive regulation of fever generation by positive regulation of
prostaglandin secretion 0.00387
positive regulation of growth hormone receptor signaling pathway 0.01113
positive regulation of histone H3-K27 methylation 0.00417
positive regulation of histone H3-K9 trimethylation 0.02134
positive regulation of interferon-alpha production 0.03247
positive regulation of interleukin-1 beta secretion 0.00547
positive regulation of JUN kinase activity 0.00901
positive regulation of lipoprotein particle clearance 0.03411
positive regulation of macrophage cytokine production 0.0152
positive regulation of MAPKKK cascade by fibroblast growth factor
receptor signaling pathway 0.03411
positive regulation of NK T cell differentiation 0.02134
positive regulation of ossification 0.00892
positive regulation of oxidative stress-induced neuron death 0.03411
positive regulation of prostaglandin-endoperoxide synthase activity 0.01113
positive regulation of proteasomal ubiquitin-dependent protein
catabolic process 0.00587

positive regulation of protein kinase A signaling 0.02073
positive regulation of protein kinase C activity 0.03411
positive regulation of pseudopodium assembly 0.02073
positive regulation of reactive oxygen species metabolic process 0.03145
positive regulation of RIG-I signaling pathway 0.00696
positive regulation of T cell mediated immunity 0.02861
positive regulation of T-helper 1 cell differentiation 0.00417
positive regulation of transcription by RNA polymerase II 1.4E-06
positive regulation of transcription, DNA-templated 0.00067
positive regulation of trophoblast cell migration 0.04908
positive regulation of tumor necrosis factor production 0.03957
positive regulation of type I interferon-mediated signaling pathway 0.00346
positive regulation of vascular smooth muscle cell proliferation 0.00302
positive regulation of Wnt signaling pathway 0.02958
positive regulation of xenophagy 0.04908
post-translational protein modification 0.00031
production of miRNAs involved in gene silencing by miRNA 0.04873
programmed necrotic cell death 0.02861
proteasomal ubiquitin-independent protein catabolic process 0.02134
protein acylation 0.04752
protein deubiquitination 0.04464
protein K11-linked deubiquitination 0.0152
protein K29-linked deubiquitination 0.01113
protein K33-linked deubiquitination 0.00387
protein localization to endosome 0.03247
protein ubiquitination 4.5E-06
regulation of acetyl-CoA biosynthetic process from pyruvate 0.01062
regulation of alpha-beta T cell proliferation 0.02374

regulation of BMP signaling pathway 0.01742
regulation of bone mineralization 0.02242
regulation of canonical Wnt signaling pathway 0.00864
regulation of CD4-positive, alpha-beta T cell activation 0.03366
regulation of endoplasmic reticulum stress-induced intrinsic apoptotic signaling pathway 0.03598
regulation of endothelial cell differentiation 0.04078
regulation of extrinsic apoptotic signaling pathway 0.01731
regulation of glycogen catabolic process 0.04908
regulation of interferon-gamma-mediated signaling pathway 0.01457
regulation of MDA-5 signaling pathway 0.04908
regulation of mesenchymal stem cell differentiation 0.04908
regulation of nitric-oxide synthase biosynthetic process 0.02263
regulation of oocyte maturation 0.04908
regulation of osteoclast differentiation 0.0345
regulation of phosphatidylinositol 3-kinase activity 0.00335
regulation of phosphoprotein phosphatase activity 0.01775
regulation of protein binding 0.04338
regulation of protein catabolic process at postsynapse, modulating synaptic transmission 0.02134
regulation of secondary metabolic process 0.04299
regulation of systemic arterial blood pressure by endothelin 0.04908
regulation of T cell cytokine production 0.01102
regulation of T cell migration 0.03986
regulation of transcription by RNA polymerase II 4.3E-17
regulation of transcription, DNA-templated 0.00124
regulation of vasculature development 0.04035
response to 2,3,7,8-tetrachlorodibenzodioxine 0.02134

response to dexamethasone	0.02861
response to dsRNA	0.02538
response to methionine	0.01113
response to UV	0.00367
response to water-immersion restraint stress	0.02134
RNA destabilization	0.03598
roof of mouth development	0.0062
secondary alcohol metabolic process	0.04602
signal transduction involved in mitotic G1 DNA damage checkpoint	0.02179
spindle assembly involved in meiosis	0.03411
thymine catabolic process	0.01113
toll-like receptor 4 signaling pathway	0.01286
ureteric bud development	0.00761
vascular endothelial cell response to fluid shear stress	0.00387
vascular smooth muscle cell differentiation	0.02374
vein smooth muscle contraction	0.02134
ventricular septum development	0.0401
ventricular septum morphogenesis	0.01597
vitellogenesis	0.02134

GO unique signature for EE2 *in vitro* exposure from Advaita

GO:Biological Process	**Adj.p-value**
4-hydroxyproline catabolic process	0.00167
9-cis-retinoic acid biosynthetic process	0.03139
activation of cysteine-type endopeptidase activity	0.00285
activation of cysteine-type endopeptidase activity involved in apoptotic process	0.00321
actomyosin contractile ring assembly	0.03139

aging 0.01915
alditol biosynthetic process 0.0497
alkaloid metabolic process 0.01224
alpha-amino acid catabolic process 0.00942
anaphase-promoting complex-dependent catabolic process 0.00081
animal organ regeneration 0.0125
apoptotic cell clearance 0.04676
apoptotic process involved in embryonic digit morphogenesis 0.0058
arsonoacetate metabolic process 0.0058
ascending aorta morphogenesis 0.00167
astrocyte cell migration 0.00741
attachment of mitotic spindle microtubules to kinetochore 0.03531
base-excision repair 0.00693
blood coagulation, intrinsic pathway 0.00727
blood vessel endothelial cell differentiation 0.00717
blood vessel remodeling 0.01106
carboxylic acid catabolic process 0.01013
cardiac muscle hypertrophy in response to stress 0.02253
carnitine shuttle 0.04587
cell cycle 0.01288
cell cycle arrest 0.03424
cell cycle G1/S phase transition 0.02662
cell division 7.3E-22
cell migration involved in sprouting angiogenesis 0.02748
cellular carbohydrate biosynthetic process 0.04396
cellular glucuronidation 0.00046
cellular process regulating host cell cycle in response to virus 0.01653
cellular response to camptothecin 0.03139

cellular response to cisplatin 0.0497
cellular response to DNA damage stimulus 0.00249
cellular response to hepatocyte growth factor stimulus 0.02914
cellular response to hydrogen peroxide 0.0065
cellular response to inorganic substance 0.02309
cellular response to interferon-gamma 0.01587
cellular response to ionizing radiation 0.00166
cellular response to leptin stimulus 0.04341
cellular response to thyroxine stimulus 0.0058
cellular response to UV 0.0108
cellular senescence 0.03246
CENP-A containing nucleosome assembly 3.7E-07
centrosome cycle 0.03604
centrosome duplication 0.00677
centrosome localization 0.03206
chromosome condensation 0.03539
chromosome segregation 0.00032
ciliary basal body-plasma membrane docking 2.2E-05
cobalamin biosynthetic process 0.01653
coenzyme A metabolic process 0.0181
complement activation, classical pathway 0.02624
courtship behavior 0.01653
cysteine metabolic process 0.00285
cytokinetic process 0.02725
cytoskeletal anchoring at nuclear membrane 0.0497
dATP catabolic process 0.01653
de novo centriole assembly involved in multi-ciliated epithelial cell differentiation 0.03139

dehydroascorbic acid transport 0.01653
deoxyribonucleoside monophosphate biosynthetic process 0.0313
deoxyribonucleotide biosynthetic process 0.01219
dGTP catabolic process 0.00044
dichotomous subdivision of terminal units involved in salivary gland branching 0.0497
dipeptide transmembrane transport 0.0497
DNA biosynthetic process 0.00796
DNA damage checkpoint 0.0062
DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest 0.00074
DNA damage response, signal transduction by p53 class mediator resulting in transcription of p21 class mediator 0.00547
DNA geometric change 0.01074
DNA ligation 0.03586
DNA protection 0.0497
DNA replication 1.4E-06
DNA replication checkpoint 0.04587
DNA replication initiation 3.3E-05
DNA replication, synthesis of RNA primer 0.00167
DNA strand elongation involved in DNA replication 1.4E-05
dorsal aorta morphogenesis 0.01848
double-strand break repair via alternative nonhomologous end joining 0.0058
double-strand break repair via homologous recombination 9.9E-05
double-strand break repair via nonhomologous end joining 0.04169
drug metabolic process 0.02245
dTTP biosynthetic process 0.01653
dUMP metabolic process 0.01653

endocardial cushion to mesenchymal transition involved in heart valve formation 0.01653
epoxygenase P450 pathway 0.04587
error-free translesion synthesis 0.01203
error-prone translesion synthesis 0.04341
establishment of mitotic spindle localization 0.03102
establishment of mitotic spindle orientation 3.8E-05
establishment of protein-containing complex localization to telomere 0.01653
establishment of RNA localization to telomere 0.01653
establishment or maintenance of microtubule cytoskeleton polarity 0.03139
ethanol oxidation 0.00128
exit from mitosis 0.01
exogenous drug catabolic process 0.00717
extrinsic apoptotic signaling pathway via death domain receptors 0.00716
fatty acid catabolic process 0.02377
fatty acid oxidation 0.02192
female meiosis chromosome segregation 0.00393
female meiosis I 0.00391
female pregnancy 0.04736
fibroblast activation 0.04587
formation of translation preinitiation complex 0.0497
fructose catabolic process to hydroxyacetone phosphate and glyceraldehyde-3-phosphate 0.0497
G1 to G0 transition involved in cell differentiation 0.01653
G1/S transition of mitotic cell cycle 0.00093
G2/M transition of mitotic cell cycle 0.00024
gamma-aminobutyric acid metabolic process 0.0497
GDP-mannose biosynthetic process 0.03139

glutamate metabolic process 0.03721
glutathione biosynthetic process 0.02914
glutathione derivative biosynthetic process 0.01505
glutathione metabolic process 0.01412
glycerol metabolic process 0.03586
glyoxylate catabolic process 0.01653
gonad development 0.01557
histidine metabolic process 0.03531
histone H3-S10 phosphorylation 0.0497
histone H4-K20 methylation 0.03531
histone phosphorylation 0.01075
icosanoid metabolic process 0.04241
inner cell mass cell proliferation 0.01011
interleukin-10 secretion 0.02617
interleukin-17 production 0.03627
interleukin-33-mediated signaling pathway 0.01653
interstrand cross-link repair 0.00206
intrinsic apoptotic signaling pathway by p53 class mediator 0.02702
intrinsic apoptotic signaling pathway in response to DNA damage by
p53 class mediator 0.02978
isopeptide cross-linking 0.01653
ketone body biosynthetic process 0.01848
ketone body catabolic process 0.03139
ketone catabolic process 0.02617
kinetochore assembly 3.1E-05
kinetochore organization 0.03109
lagging strand elongation 0.01653
liver regeneration 0.00912

long-chain fatty acid biosynthetic process 0.00208
lymphocyte homeostasis 0.0369
maintenance of DNA repeat elements 0.0497
male meiosis I 0.00632
mannose metabolic process 0.01653
medium-chain fatty-acyl-CoA metabolic process 0.01653
meiotic chromosome condensation 0.01653
meiotic nuclear division 0.01792
meiotic sister chromatid cohesion 0.0497
metaphase plate congression 0.01763
metaphase/anaphase transition of mitotic cell cycle 0.0161
microtubule depolymerization 0.00337
microtubule polymerization 0.02451
mismatch repair 0.00266
mitochondrial DNA replication 0.00183
mitotic cell cycle 0.03496
mitotic cell cycle checkpoint 0.00168
mitotic centrosome separation 0.04587
mitotic chromosome condensation 0.00053
mitotic chromosome movement towards spindle pole 0.03139
mitotic cytokinesis 1.2E-07
mitotic metaphase plate congression 0.00027
mitotic nuclear envelope disassembly 0.00012
mitotic recombination-dependent replication fork processing 0.0058
mitotic sister chromatid segregation 0.00548
mitotic spindle assembly 1.5E-05
mitotic spindle assembly checkpoint 2.8E-06
mitotic spindle midzone assembly 0.00044

mitotic spindle organization 1.4E-05

modulation of microtubule cytoskeleton involved in cerebral cortex radial glia guided migration 0.0497

monocarboxylic acid catabolic process 0.03511

muscle hyperplasia 0.0497

NAD biosynthesis via nicotinamide riboside salvage pathway 0.02322

negative regulation of adenylate cyclase-activating adrenergic receptor signaling pathway 0.01653

negative regulation of biomineral tissue development 0.03206

negative regulation of cell division 0.04341

negative regulation of cell growth 0.01824

negative regulation of centriole replication 0.01848

negative regulation of ceramide biosynthetic process 0.0497

negative regulation of chemokine-mediated signaling pathway 0.01224

negative regulation of coagulation 0.03746

negative regulation of DNA replication 0.00957

negative regulation of fibrinolysis 0.02617

negative regulation of G0 to G1 transition 0.00585

negative regulation of G1/S transition of mitotic cell cycle 0.00233

negative regulation of helicase activity 0.0497

negative regulation of heterotypic cell-cell adhesion 0.03531

negative regulation of inclusion body assembly 0.03531

negative regulation of interferon-gamma-mediated signaling pathway 0.00741

negative regulation of intestinal cholesterol absorption 0.0058

negative regulation of intestinal phytosterol absorption 0.0058

negative regulation of lamellipodium assembly 0.01653

negative regulation of lymphocyte apoptotic process 0.01703

negative regulation of molecular function 0.02132

negative regulation of muscle adaptation 0.04587
negative regulation of neuron death 0.0394
negative regulation of nuclear cell cycle DNA replication 0.03139
negative regulation of protein binding 0.00937
negative regulation of protein processing 0.04131
negative regulation of protein ubiquitination 0.03246
negative regulation of ruffle assembly 0.0497
negative regulation of smooth muscle cell chemotaxis 0.0497
negative regulation of transcription from RNA polymerase II promoter involved in smooth muscle cell differentiation 0.01653
negative regulation of transcription initiation from RNA polymerase II promoter 0.00393
negative regulation of type I interferon-mediated signaling pathway 0.04587
negative regulation of vascular endothelial cell proliferation 0.01848
negative regulation of vascular endothelial growth factor signaling pathway 0.01848
negative regulation of viral entry into host cell 0.00221
negative regulation of viral genome replication 0.00219
neuron death in response to oxidative stress 0.04379
neurotransmitter catabolic process 0.01796
neutrophil clearance 0.0497
Notch signaling involved in heart development 0.04587
nucleobase-containing small molecule interconversion 8E-05
nucleoside monophosphate phosphorylation 0.01374
nucleotide-excision repair, DNA gap filling 0.00632
nucleotide-excision repair, DNA incision, 5'-to lesion 0.01735
omega-hydroxylase P450 pathway 0.02617
oocyte maturation 0.00019

opsonization 0.0031
ovulation 0.01505
oxidation-reduction process 0.00576
pallium development 0.02041
parturition 0.02914
peptidyl-glutamine modification 0.03139
plasma membrane repair 0.03531
platelet degranulation 0.00652
positive regulation of adipose tissue development 0.0497
positive regulation of angiogenesis 0.03755
positive regulation of apoptotic DNA fragmentation 0.0497
positive regulation of apoptotic process by virus 0.0058
positive regulation of bleb assembly 0.0058
positive regulation of blood vessel diameter 0.0378
positive regulation of cAMP-dependent protein kinase activity 0.0497
positive regulation of catalytic activity 0.02541
positive regulation of cell growth involved in cardiac muscle cell development 0.01224
positive regulation of chromosome segregation 0.00789
positive regulation of cytokinesis 0.00337
positive regulation of DNA biosynthetic process 0.02687
positive regulation of DNA repair 0.04048
positive regulation of DNA-dependent DNA replication 0.04587
positive regulation of DNA-directed DNA polymerase activity 1.4E-05
positive regulation of endoplasmic reticulum stress-induced intrinsic apoptotic signaling pathway 0.03531
positive regulation of establishment of protein localization 0.04511
positive regulation of exit from mitosis 0.0497

positive regulation of extracellular matrix constituent secretion 0.01653
positive regulation of extrinsic apoptotic signaling pathway 0.00942
positive regulation of fibroblast proliferation 0.02735
positive regulation of G2/M transition of mitotic cell cycle 0.00027
positive regulation of histone H3-K14 acetylation 0.03139
positive regulation of isotype switching to IgA isotypes 0.0497
positive regulation of mitochondrial ATP synthesis coupled electron transport 0.0497
positive regulation of mitochondrial depolarization 0.02617
positive regulation of mitotic cell cycle spindle assembly checkpoint 0.01224
positive regulation of mitotic nuclear division 0.01098
positive regulation of phagocytosis 0.00794
positive regulation of platelet-derived growth factor receptor signaling pathway 0.0497
positive regulation of protein localization to cell cortex 0.0497
positive regulation of protein oligomerization 0.00632
positive regulation of protein polyubiquitination 0.04587
positive regulation of release of cytochrome c from mitochondria 0.00087
positive regulation of smooth muscle cell proliferation 0.02501
positive regulation of spindle assembly 0.0497
positive regulation of toll-like receptor 4 signaling pathway 0.01848
positive regulation of transcription from RNA polymerase II promoter involved in myocardial precursor cell differentiation 0.03139
positive regulation of tumor necrosis factor-mediated signaling pathway 0.02617
positive regulation of tyrosine phosphorylation of STAT protein 0.04832
positive regulation of ubiquitin protein ligase activity 0.00128
positive regulation of viral life cycle 0.02748

post-chaperonin tubulin folding pathway 0.03139
post-embryonic animal organ morphogenesis 0.01848
post-translational protein acetylation 0.01653
pri-miRNA transcription by RNA polymerase II 0.02161
proline catabolic process to glutamate 0.0058
protein depalmitoylation 0.01224
protein destabilization 0.00611
protein homooligomerization 0.01647
protein homotetramerization 0.01558
protein insertion into mitochondrial membrane involved in apoptotic signaling pathway 0.02013
protein K27-linked ubiquitination 0.0497
protein K6-linked ubiquitination 0.01848
protein lipidation 0.02387
protein localization to chromatin 0.01428
protein localization to chromosome, centromeric region 0.04932
protein localization to kinetochore 0.0008
protein localization to microtubule cytoskeleton 0.00374
protein refolding 0.02322
protoporphyrinogen IX biosynthetic process 0.00183
pulmonary valve morphogenesis 0.03586
purine nucleobase catabolic process 0.01653
purine nucleoside bisphosphate catabolic process 0.0497
pyrimidine deoxyribonucleoside monophosphate biosynthetic process 0.01653
pyrimidine dimer repair 0.01224
pyrimidine nucleoside salvage 0.04587
quinone metabolic process 0.01793
regulation of attachment of spindle microtubules to kinetochore 0.00128

regulation of B cell apoptotic process 0.03586
regulation of bile acid biosynthetic process 0.02617
regulation of cardiocyte differentiation 0.0333
regulation of cell-cell adhesion mediated by integrin 0.04587
regulation of centriole replication 0.04563
regulation of chemokine secretion 0.02914
regulation of cilium assembly 0.0378
regulation of complement activation 0.00057
regulation of cyclin-dependent protein serine/threonine kinase activity 0.00071
regulation of cytokinesis 0.04141
regulation of DNA damage response, signal transduction by p53 class mediator 0.02742
regulation of epithelial cell proliferation 0.04117
regulation of exit from mitosis 0.04571
regulation of extracellular matrix disassembly 0.02322
regulation of G2/M transition of mitotic cell cycle 3.8E-05
regulation of glutamate metabolic process 0.01653
regulation of histone H3-K9 acetylation 0.04587
regulation of interleukin-8 secretion 0.03206
regulation of lipid transporter activity 0.03139
regulation of mast cell cytokine production 0.0497
regulation of microtubule polymerization or depolymerization 0.00491
regulation of mitochondrial depolarization 0.04564
regulation of mitochondrial outer membrane permeabilization involved in apoptotic signaling pathway 0.03781
regulation of mitotic metaphase/anaphase transition 0.00032
regulation of mitotic nuclear division 0.01082
regulation of mitotic spindle organization 0.00491

regulation of neuron apoptotic process 0.04986
regulation of nitrogen utilization 0.01653
regulation of nuclease activity 0.04341
regulation of pinocytosis 0.04587
regulation of plasminogen activation 0.03531
regulation of protein ADP-ribosylation 0.02617
regulation of response to DNA damage stimulus 0.01154
regulation of retinal ganglion cell axon guidance 0.0497
regulation of signal transduction by p53 class mediator 0.00095
regulation of sister chromatid cohesion 0.01854
regulation of transcription from RNA polymerase II promoter in response to iron 0.0058
regulation of transcription involved in G1/S transition of mitotic cell cycle 2.3E-05
regulation of transferase activity 0.03517
regulation of vasculogenesis 0.0181
renal sodium ion absorption 0.03139
replication fork protection 0.00741
response to activity 0.00293
response to axon injury 0.04207
response to cadmium ion 0.00255
response to cold 0.00378
response to cortisol 0.0497
response to fluoride 0.0497
response to folic acid 0.03531
response to food 0.03627
response to glucocorticoid 0.00959
response to heat 0.01411

response to human chorionic gonadotropin 0.01653
response to hydrostatic pressure 0.0497
response to immobilization stress 0.01428
response to mechanical stimulus 0.01423
response to ozone 0.0497
response to platinum ion 0.0058
response to progesterone 0.01468
response to virus 0.03988
response to vitamin A 0.01203
response to vitamin B1 . . . . . 0.01653
response to wounding 0.03855
response to X-ray 0.01703
retinol metabolic process 0.0063
retrograde vesicle-mediated transport, Golgi to ER 0.00363
ribonucleoside bisphosphate catabolic process 0.0497
salivary gland cavitation 0.0497
serine family amino acid biosynthetic process 0.0181
signal transduction involved in G2 DNA damage checkpoint 0.01374
somatic hypermutation of immunoglobulin genes 0.00717
spindle assembly involved in female meiosis I 0.0058
spleen development 0.03627
spongiotrophoblast differentiation 0.03139
substrate-dependent cerebral cortex tangential migration 0.0497
synaptonemal complex disassembly 0.0058
tangential migration from the subventricular zone to the olfactory bulb 0.0497
telomere maintenance via recombination 0.01011
telomere maintenance via semi-conservative replication 2E-08
terpenoid catabolic process 0.0497

third ventricle development	0.03139
thrombin-activated receptor signaling pathway	0.01011
toxin biosynthetic process	0.0497
toxin metabolic process	0.0357
trabecula formation	0.02703
triglyceride homeostasis	0.03164
trophoblast giant cell differentiation	0.01011
type 2 immune response	0.04131
type I interferon signaling pathway	0.03337
ubiquinone biosynthetic process	0.01011
ubiquitin-dependent SMAD protein catabolic process	0.0497
umbilical cord morphogenesis	0.0058
vascular smooth muscle cell development	0.0031
ventricular cardiac muscle cell development	0.00485
xenobiotic metabolic process	0.00096
zygotic determination of anterior/posterior axis, embryo	0.0497

Table B.1: Unique BP signature from iPathway Guide (Advaita) after *in vitro* exposures.

GO unique signature for AFB1 3 Days *in vivo* exposure from Advaita	
GO:Biological Process	**Adj.p-value**
GO:Biological Process	Adj.p-value
'de novo' NAD biosynthetic process from tryptophan	0.00464
2-oxobutyrate metabolic process	0.01168
actin rod assembly	0.01168
adhesion of symbiont to host	0.04382

aldehyde catabolic process 0.01062
ammonium ion metabolic process 0.02361
B cell differentiation 0.02805
canalicular bile acid transport 0.03253
cap-dependent translational initiation 0.03253
carbohydrate phosphorylation 0.02057
carnitine metabolic process, CoA-linked 0.03253
CDP-choline pathway 0.01962
cellular chloride ion homeostasis 0.03253
cellular modified amino acid metabolic process 0.02242
cellular nitrogen compound catabolic process 0.04628
cellular protein-containing complex localization 0.01273
cellular response to metal ion 0.02684
cellular response to tumor necrosis factor 0.03283
cofactor metabolic process 0.03491
collateral sprouting in absence of injury 0.03253
common myeloid progenitor cell proliferation 0.0316
creatine transmembrane transport 0.01168
deadenylation-dependent decapping of nuclear-transcribed mRNA 0.03253
detection of oxygen 0.01168
diadenosine polyphosphate metabolic process 0.03253
DNA replication checkpoint 0.0316
endoplasmic reticulum membrane organization 0.01962
epithelial cell proliferation involved in renal tubule morphogenesis 0.03253
establishment of blood-nerve barrier 0.03253
establishment of endothelial intestinal barrier 0.01673
establishment of Golgi localization 0.0316
extrinsic apoptotic signaling pathway in absence of ligand 0.04939

fatty acid alpha-oxidation 0.04657
fatty acid beta-oxidation 0.01581
fatty acid biosynthetic process 0.02627
fatty acid transport 0.02867
fatty-acyl-CoA biosynthetic process 0.00619
ferrous iron transmembrane transport 0.01168
G2/M transition of mitotic cell cycle 0.01334
ganglioside biosynthetic process 0.04657
glomerular filtration 0.02994
glyceraldehyde-3-phosphate metabolic process 0.04696
glycerol-3-phosphate catabolic process 0.01168
glycoside metabolic process 0.0331
Golgi to plasma membrane CFTR protein transport 0.03253
hepatocyte apoptotic process 0.01673
immunological synapse formation 0.00387
interleukin-2-mediated signaling pathway 0.02408
interleukin-4-mediated signaling pathway 0.01962
Kit signaling pathway 0.03253
L-methionine salvage from S-adenosylmethionine 0.03253
lipid droplet organization 0.02994
lipid glycosylation 0.01962
lymphocyte homeostasis 0.04876
macrophage cytokine production 0.0331
maintenance of protein location in mitochondrion 0.01066
medium-chain fatty acid metabolic process 0.0316
mitochondrion to lysosome transport 0.03253
modified amino acid transport 0.01733
N-terminal peptidyl-methionine acetylation 0.03253

necroptotic signaling pathway 0.01066
negative regulation of autophagosome assembly 0.01097
negative regulation of complement activation 0.04657
negative regulation of endoplasmic reticulum stress-induced neuron intrinsic apoptotic signaling pathway 0.01168
negative regulation of exocytosis 0.00877
negative regulation of fatty acid biosynthetic process 0.01743
negative regulation of Fc-gamma receptor signaling pathway involved in phagocytosis 0.01168
negative regulation of innate immune response 0.03022
negative regulation of membrane permeability 0.0316
negative regulation of protein kinase C signaling 0.01168
negative regulation of receptor recycling 0.03253
negative regulation of Rho-dependent protein serine/threonine kinase activity 0.03253
negative regulation of serotonin secretion 0.01168
negative regulation of T cell apoptotic process 0.04382
negative regulation of transposition 0.04657
negative regulation of viral-induced cytoplasmic pattern recognition receptor signaling pathway 0.01673
neuromuscular junction development 0.00685
neurotransmitter catabolic process 0.04696
neutrophil clearance 0.03253
neutrophil degranulation 0.00402
ovulation from ovarian follicle 0.04657
oxygen transport 0.00127
parkin-mediated stimulation of mitophagy in response to mitochondrial depolarization 0.01962

pentose metabolic process 0.04657
peptidyl-lysine N6-acetylation 0.01168
phosphatidylethanolamine biosynthetic process 0.00897
platelet dense granule organization 0.0316
positive regulation of ARF protein signal transduction 0.01168
positive regulation of autophagy of mitochondrion in response to mitochondrial depolarization 0.03145
positive regulation of cell proliferation involved in kidney development 0.0316
positive regulation of cold-induced thermogenesis 0.02085
positive regulation of dephosphorylation 0.04876
positive regulation of eosinophil migration 0.03253
positive regulation of ER-associated ubiquitin-dependent protein catabolic process 0.0316
positive regulation of gluconeogenesis 0.0331
positive regulation of histone deacetylase activity 0.00464
positive regulation of membrane potential 0.0331
positive regulation of mitochondrial electron transport, NADH to ubiquinone 0.03253
positive regulation of nuclear receptor transcription coactivator activity 0.01168
positive regulation of phosphatidylcholine biosynthetic process 0.03253
positive regulation of protein linear polyubiquitination 0.01168
positive regulation of protein localization to endoplasmic reticulum 0.01168
positive regulation of retrograde protein transport, ER to cytosol 0.03253
positive regulation of tumor necrosis factor secretion 0.04382
proteasome assembly 0.01673
protein geranylgeranylation 0.01962
protein localization to cell junction 0.04657
protein targeting to peroxisome 0.03181

pyrimidine dimer repair 0.01962
pyruvate biosynthetic process 0.00457
regulation of ARF protein signal transduction 0.00602
regulation of cell shape 0.02852
regulation of cellular amine metabolic process 0.04549
regulation of cofactor metabolic process 0.0124
regulation of dendritic cell cytokine production 0.0316
regulation of fibroblast migration 0.01218
regulation of hepatocyte proliferation 0.04382
regulation of I-kappaB kinase/NF-kappaB signaling 0.04152
regulation of nitric-oxide synthase biosynthetic process 0.02314
regulation of norepinephrine uptake 0.01168
regulation of oligodendrocyte apoptotic process 0.01168
regulation of p38MAPK cascade 0.04711
regulation of PERK-mediated unfolded protein response 0.01097
regulation of RNA binding 0.02408
regulation of ruffle assembly 0.0463
regulation of supramolecular fiber organization 0.04666
regulation of tetrapyrrole metabolic process 0.03253
regulation of thyroid hormone mediated signaling pathway 0.03253
response to exogenous dsRNA 0.03859
response to hydrogen sulfide 0.03253
response to interleukin-18 . . . . . 0.04657
response to unfolded protein 0.0262
retrograde trans-synaptic signaling by endocannabinoid 0.03253
RNA 5'-cap (guanine-N7)-methylation 0.01168
selenium compound metabolic process 0.00126
sequestering of iron ion 0.03253

siderophore transport	0.03253
sulfur compound metabolic process	0.02612
surfactant homeostasis	0.04657
tooth mineralization	0.02314
trabecula morphogenesis	0.03765
trans-synaptic signaling by endocannabinoid, modulating synaptic transmission	0.01168
tryptophan catabolic process	0.04657
vesicle organization	0.04868
xenobiotic metabolic process	0.00094
xenophagy	0.04657

GO unique signature for AFB1 5 Days *in vivo* exposure from Advaita

GO:Biological Process	**Adj.p-value**
adherens junction organization	0.00645
alveolar primary septum development	0.0447
AMP biosynthetic process	0.01823
antigen processing and presentation of endogenous peptide antigen via MHC class I	0.03607
antigen processing and presentation of endogenous peptide antigen via MHC class Ib via ER pathway, TAP-dependent	0.0447
antral ovarian follicle growth	0.02654
astral microtubule organization	0.03607
base conversion or substitution editing	0.04669
basement membrane organization	0.01111
bile acid metabolic process	0.03173
Cajal body organization	0.00034
cell-matrix adhesion	0.03111

cellular response to caloric restriction 0.0447
cellular response to carbon monoxide 0.0447
cellular response to cholesterol 0.03222
cellular response to cold 0.03607
cellular response to electrical stimulus 0.0112
cellular response to epinephrine stimulus 0.01817
cellular response to estrogen stimulus 0.0112
cellular response to interleukin-3 ..... 0.01823
cellular response to lipoprotein particle stimulus 0.03867
cellular response to reactive oxygen species 0.04907
cellular response to UV-C 0.02654
cellular response to virus 0.00816
cellular response to vitamin B1 ..... 0.0447
cellular sodium ion homeostasis 0.01445
ceramide transport 0.03607
chorio-allantoic fusion 0.02654
chromatin silencing at telomere 0.02654
cofactor biosynthetic process 0.03458
coreceptor-mediated virion attachment to host cell 0.0447
dibasic protein processing 0.0447
establishment or maintenance of polarity of embryonic epithelium 0.0447
eye pigment granule organization 0.0447
fatty acid beta-oxidation using acyl-CoA oxidase 0.01445
ferrous iron export across plasma membrane 0.0447
fibroblast proliferation 0.03628
fructose 6-phosphate metabolic process 0.02654
fusion of virus membrane with host plasma membrane 0.03607
glucose catabolic process 0.0478

glucoside transport 0.0447

glycosaminoglycan catabolic process 0.01757

helper T cell enhancement of adaptive immune response 0.0447

heme catabolic process 0.02654

heme oxidation 0.0447

histone H3-K4 acetylation 0.0447

histone H4 deacetylation 0.04669

induction by virus of host cell-cell fusion 0.0447

interleukin-3-mediated signaling pathway 0.01823

intermembrane lipid transfer 0.01127

intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator 0.00358

lateral ventricle development 0.03607

lung vasculature development 0.02654

membrane protein ectodomain proteolysis 0.02601

membrane raft organization 0.04396

metanephric collecting duct development 0.03607

microglial cell activation involved in immune response 0.00581

mitochondrial translational initiation 0.0447

mitochondrial tRNA modification 0.03607

mRNA methylation 0.01445

mucopolysaccharide metabolic process 0.04036

negative regulation of adenylate cyclase activity 0.02237

negative regulation of apoptotic process 9.2E-05

negative regulation of apoptotic signaling pathway 0.04747

negative regulation of carbohydrate metabolic process 0.01292

negative regulation of cell differentiation 0.04918

negative regulation of cellular biosynthetic process 0.04635

negative regulation of cellular response to hepatocyte growth factor stimulus 0.0447
negative regulation of coenzyme metabolic process 0.02705
negative regulation of constitutive secretory pathway 0.0447
negative regulation of endocytic recycling 0.0447
negative regulation of hepatocyte differentiation 0.0447
negative regulation of hepatocyte growth factor receptor signaling pathway 0.0447
negative regulation of histone H2A K63-linked ubiquitination 0.00581
negative regulation of keratinocyte migration 0.0447
negative regulation of lung blood pressure 0.0447
negative regulation of mast cell activation involved in immune response 0.02654
negative regulation of nucleotide catabolic process 0.02705
negative regulation of oligodendrocyte progenitor proliferation 0.0447
negative regulation of platelet activation 0.03222
negative regulation of protein processing involved in protein targeting to mitochondrion 0.0447
negative regulation of respiratory burst involved in inflammatory response 0.01127
negative regulation of response to DNA damage stimulus 0.01564
negative regulation of RNA interference 0.0447
negative regulation of synaptic vesicle clustering 0.0447
negative regulation of transforming growth factor beta1 production 0.01127
nerve maturation 0.0447
neuron-neuron synaptic transmission 0.04669
nose morphogenesis 0.0447
nuclear inner membrane organization 0.0447
nucleolus organization 0.02654
outer medullary collecting duct development 0.0447
p38MAPK cascade 0.04166

peptidyl-lysine deacetylation 0.02654
peptidyl-proline hydroxylation to 3-hydroxy-L-proline 0.0447
phospholipid translocation 0.02237
pigment biosynthetic process 0.04922
plasma membrane organization 0.02445
positive regulation of alkaline phosphatase activity 0.01127
positive regulation of astrocyte differentiation 0.03785
positive regulation of cardiac muscle cell apoptotic process 0.0112
positive regulation of cell proliferation 0.0296
positive regulation of chemokine secretion 0.02237
positive regulation of cholesterol esterification 0.04669
positive regulation of complement-dependent cytotoxicity 0.0447
positive regulation of DNA ligase activity 0.0447
positive regulation of DNA repair 0.04483
positive regulation of DNA strand elongation 0.0447
positive regulation of double-strand break repair via homologous recombination 0.03607
positive regulation of execution phase of apoptosis 0.01445
positive regulation of glutathione biosynthetic process 0.0447
positive regulation of hematopoietic stem cell proliferation 0.01127
positive regulation of histone H3-K27 acetylation 0.0447
positive regulation of interferon-alpha secretion 0.02654
positive regulation of intracellular signal transduction 0.04416
positive regulation of IP-10 production 0.0447
positive regulation of JNK cascade 0.02597
positive regulation of L-arginine import across plasma membrane 0.0447
positive regulation of L-lysine import across plasma membrane 0.0447
positive regulation of lysosome organization 0.0447

positive regulation of mRNA 3'-UTR binding 0.0447
positive regulation of muscle hyperplasia 0.0447
positive regulation of myeloid leukocyte differentiation 0.00921
positive regulation of NAD(P)H oxidase activity 0.00272
positive regulation of neuron apoptotic process 0.0326
positive regulation of nucleocytoplasmic transport 0.04409
positive regulation of peptidyl-serine phosphorylation of STAT protein 0.01823
positive regulation of protein folding 0.0447
positive regulation of protein tyrosine kinase activity 0.01436
positive regulation of proton-transporting ATP synthase activity, rotational mechanism 0.0447
positive regulation of RNA polymerase II transcriptional preinitiation complex assembly 0.01823
positive regulation of smooth muscle cell proliferation 0.03376
positive regulation of sodium:potassium-exchanging ATPase activity 0.0447
positive regulation of SUMO transferase activity 0.0447
positive regulation of superoxide anion generation 0.00276
positive regulation of T cell differentiation in thymus 0.04669
positive regulation of transcription from RNA polymerase II promoter in response to arsenic-containing substance 0.0447
positive regulation of transcription from RNA polymerase II promoter in response to oxidative stress 0.01823
positive regulation of transforming growth factor beta1 activation 0.0447
pre-mRNA catabolic process 0.0447
production of miRNAs involved in gene silencing by miRNA 0.02878
protein K33-linked ubiquitination 0.0447
protein K69-linked ufmylation 0.01823
protein localization to endoplasmic reticulum tubular network 0.0447

protein transport into membrane raft 0.01127
proteoglycan catabolic process 0.00581
pyruvate metabolic process 0.03112
regulation of calcium import into the mitochondrion 0.0447
regulation of cytoplasmic mRNA processing body assembly 0.01823
regulation of intracellular estrogen receptor signaling pathway 0.04309
regulation of macrophage differentiation 0.03785
regulation of myelination 0.02342
regulation of neuroinflammatory response 0.02101
regulation of nucleus organization 0.0447
regulation of pyruvate dehydrogenase activity 0.0447
regulation of RNA splicing 0.04036
regulation of transcription from RNA polymerase II promoter in response to copper ion 0.00199
regulation of transcription from RNA polymerase II promoter in response to iron 0.00199
regulation of vascular smooth muscle cell proliferation 0.03007
regulation of viral entry into host cell 0.01777
response to acetate 0.0447
response to decreased oxygen levels 0.04162
response to formaldehyde 0.0447
response to hydroperoxide 0.00732
response to lipoprotein particle 0.03063
response to superoxide 0.02059
RNA folding 0.0447
smoothened signaling pathway involved in regulation of secondary heart field cardioblast proliferation 0.0447
snRNA (adenine-N6)-methylation 0.0447

spleen trabecula formation	0.0447
sterol homeostasis	0.03877
sulfur amino acid transport	0.03607
synapse pruning	0.0061
telomere formation via telomerase	0.0447
trans-Golgi network to recycling endosome transport	0.0447
transcription-dependent tethering of RNA polymerase II gene DNA at nuclear periphery	0.0447
tricellular tight junction assembly	0.0447
very long-chain fatty-acyl-CoA catabolic process	0.0447
viral protein processing	0.0447
vitamin E biosynthetic process	0.0447
Wnt receptor catabolic process	0.0447

Table B.2: Unique BP signatures for AFB1 *in vivo* exposures from iPathway Guide (Advaita).

GO terms shared by both AFB1 *in vivo* exposure from Advaita		
	3 Days	**5 Days**
GO:Biological Process	**adj.p-value**	**adj.p-value**
positive regulation of transcription involved in G1/S transition of mitotic cell cycle	0.01168	0.00199
regulation of epidermal growth factor receptor signaling pathway	0.01037	0.04407

Table B.3: Shared BP by AFB1 *in vivo* exposures from iPathway Guide (Advaita)

GO unique signature for EE2 6 Hours *in vivo* exposure from Advaita	
GO:Biological Process	**Adj.p-value**
2'-deoxyribonucleotide biosynthetic process	0.04781
adipose tissue development	0.01219
ADP metabolic process	0.01874
alanyl-tRNA aminoacylation	0.01602
alditol catabolic process	0.04401
aminoacyl-tRNA metabolism involved in translational fidelity	0.00469
AMP catabolic process	0.01602
angiogenesis involved in coronary vascular morphogenesis	0.04401
apoptotic DNA fragmentation	0.0146
aspartate catabolic process	0.04401
auditory receptor cell fate determination	0.01602
axonogenesis involved in innervation	0.04401
base-excision repair, gap-filling	0.03012
binding of sperm to zona pellucida	0.02007
cardiolipin biosynthetic process	0.04781
cellular response to fatty acid	0.00148
cellular response to hydrogen peroxide	0.01
cellular response to interleukin-4	0.01235
cellular response to ionizing radiation	0.04489
cellular stress response to acid chemical	0.04401

circadian regulation of translation 0.04401
comma-shaped body morphogenesis 0.03012
commissural neuron axon guidance 0.04054
coumarin catabolic process 0.01602
cristae formation 0.03266
cysteine biosynthetic process via cystathionine 0.04401
deoxyribonucleotide biosynthetic process 0.01597
DNA damage checkpoint 0.01025
DNA ligation involved in DNA repair 0.04781
DNA recombination 0.0422
double-strand break repair via classical nonhomologous end joining 0.04401
endosome transport via multivesicular body sorting pathway 0.00771
fatty acid homeostasis 0.0173
fatty-acyl-CoA catabolic process 0.01602
female meiosis chromosome separation 0.01602
forebrain anterior/posterior pattern specification 0.04401
free ubiquitin chain polymerization 0.03012
GDP metabolic process 0.04781
GDP-mannose metabolic process 0.04781
glucocorticoid mediated signaling pathway 0.04401
gluconeogenesis 0.00575
glutathione biosynthetic process 0.02416
GMP metabolic process 0.00469
Golgi disassembly 0.01661
guanine metabolic process 0.04401
hexitol metabolic process 0.04401
histone H3-R17 methylation 0.01602
histone H3-S10 phosphorylation 0.01661

hypoxanthine biosynthetic process 0.01602
immunoglobulin transcytosis in epithelial cells 0.04401
IMP biosynthetic process 0.0286
insulin receptor signaling pathway 0.03433
low-density lipoprotein particle clearance 0.03524
maintenance of DNA repeat elements 0.00733
maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA) 0.04781
mitochondrial acetyl-CoA biosynthetic process from pyruvate 0.03012
mitochondrial DNA metabolic process 0.03255
mitochondrial mRNA catabolic process 0.04401
mitochondrial respiratory chain complex III assembly 0.04781
mitotic cell cycle checkpoint 0.01164
mitotic centrosome separation 0.00652
mitotic nuclear envelope disassembly 0.04054
mitotic spindle assembly 0.03729
mitotic spindle organization 0.0092
negative regulation of astrocyte activation 0.01602
negative regulation of DNA duplex unwinding 0.04401
negative regulation of ER-associated ubiquitin-dependent protein catabolic process 0.04401
negative regulation of forebrain neuron differentiation 0.01602
negative regulation of glucocorticoid secretion 0.01602
negative regulation of helicase activity 0.04401
negative regulation of hypoxia-induced intrinsic apoptotic signaling pathway 0.01661
negative regulation of inner ear receptor cell differentiation 0.04401
negative regulation of mitotic cell cycle 0.04852

negative regulation of oligodendrocyte differentiation 0.04054
negative regulation of oxidative stress-induced intrinsic apoptotic signaling pathway 0.00315
negative regulation of pro-B cell differentiation 0.04401
negative regulation of protein binding 0.02923
negative regulation of protein neddylation 0.01602
negative regulation of ubiquitin-dependent protein catabolic process 0.00375
Notch receptor processing, ligand-dependent 0.04781
nuclear pore complex assembly 0.01904
nucleoside catabolic process 0.04119
oxaloacetate(2-) transmembrane transport 0.01602
peripheral nervous system myelin maintenance 0.04781
phagosome acidification 0.00732
plus-end-directed vesicle transport along microtubule 0.04781
pore complex assembly 0.00308
positive regulation by virus of viral protein levels in host cell 0.04401
positive regulation of cardiac muscle myoblast proliferation 0.04401
positive regulation of cell proliferation by VEGF-activated platelet derived growth factor receptor signaling pathway 0.04401
positive regulation of cellular amide metabolic process 0.03706
positive regulation of cyclin-dependent protein serine/threonine kinase activity 0.02532
positive regulation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway 0.01904
positive regulation of cytokinesis 0.02911
positive regulation of DNA ligation 0.04401
positive regulation of DNA replication 0.0151
positive regulation of histone H3-K79 methylation 0.01602

positive regulation of histone H3-K9 methylation 0.04781
positive regulation of ligase activity 0.03012
positive regulation of lysosomal protein catabolic process 0.04401
positive regulation of miRNA catabolic process 0.01602
positive regulation of mitochondrial ATP synthesis coupled electron transport 0.01661
positive regulation of mitochondrial membrane potential 0.01174
positive regulation of mitochondrial outer membrane permeabilization involved in apoptotic signaling pathway 0.03524
positive regulation of mitochondrial RNA catabolic process 0.01602
positive regulation of mitophagy in response to mitochondrial depolarization 0.01904
positive regulation of mitotic cytokinetic process 0.01602
positive regulation of mitotic nuclear division 0.02874
positive regulation of protein dephosphorylation 0.02687
positive regulation of protein import into nucleus 0.02687
positive regulation of protein targeting to mitochondrion 0.00112
positive regulation of release of cytochrome c from mitochondria 0.03846
positive regulation of retinal ganglion cell axon guidance 0.01602
positive regulation of telomerase RNA reverse transcriptase activity 0.01602
positive regulation of transcription elongation from RNA polymerase II promoter 0.00469
positive regulation of translational termination 0.04401
post-embryonic camera-type eye morphogenesis 0.04401
post-translational protein acetylation 0.04401
postreplication repair 0.01933
protein import into nucleus 0.01263

protein insertion into mitochondrial membrane involved in apoptotic signaling pathway 0.04645
protein K29-linked ubiquitination 0.04401
protein K6-linked ubiquitination 0.04781
protein localization to chromosome 0.03814
protein localization to cytoplasmic stress granule 0.03012
protein localization to nuclear envelope 0.01661
protein stabilization 0.02668
protein-cofactor linkage 0.01904
protoporphyrinogen IX biosynthetic process 0.04781
purine nucleotide salvage 0.04781
purine ribonucleoside biosynthetic process 0.04033
purine ribonucleoside salvage 0.01661
pyrimidine deoxyribonucleotide metabolic process 0.01186
pyrimidine-containing compound metabolic process 0.03395
regulation of cell cycle arrest 0.01068
regulation of chaperone-mediated protein folding 0.01602
regulation of G1/S transition of mitotic cell cycle 0.04104
regulation of Golgi to plasma membrane protein transport 0.03012
regulation of IRE1-mediated unfolded protein response 0.04054
regulation of nitrogen utilization 0.04401
regulation of sister chromatid cohesion 0.04255
regulation of translation 0.01324
response to antineoplastic agent 0.01143
response to leptin 0.04119
response to osmotic stress 0.04226
response to UV-A 0.0031
retinal cell apoptotic process 0.04401

ribonucleoside diphosphate catabolic process	0.03012
RNA splicing	0.01802
rRNA import into mitochondrion	0.04401
T cell proliferation involved in immune response	0.01602
toxin transport	0.03524
transcription initiation from RNA polymerase II promoter	0.02761
translational frameshifting	0.04401
VEGF-activated neuropilin signaling pathway	0.01602
ventricular cardiac muscle cell development	0.01904
viral RNA genome packaging	0.01602

GO unique signature for EE2 9 Hours *in vivo* exposure from Advaita

GO:Biological Process	**Adj.p-value**
'de novo' AMP biosynthetic process	0.01516
'de novo' IMP biosynthetic process	0.03628
'de novo' pyrimidine nucleobase biosynthetic process	0.00093
2-oxoglutarate metabolic process	0.0363
[2Fe-2S] cluster assembly	0.04939
amide biosynthetic process	0.04068
angiogenesis involved in wound healing	0.02694
aspartate metabolic process	0.03628
ATP-dependent chromatin remodeling	0.04608
attachment of spindle microtubules to kinetochore	0.00957
base-excision repair	0.00739
cell motility involved in cerebral cortex radial glia guided migration	0.03628
cellular amino acid biosynthetic process	0.04861
cellular detoxification of aldehyde	0.03628
cellular modified amino acid biosynthetic process	0.01008

cellular response to camptothecin 0.04939
cellular response to exogenous dsRNA 0.01344
cellular response to misfolded protein 0.02297
cellular response to UV-A 0.01516
centriole replication 0.01795
centrosome cycle 0.00533
chaperone cofactor-dependent protein refolding 0.02297
chaperone-mediated protein transport 0.00711
chromosome segregation 0.00282
complement activation, classical pathway 0.00192
copper ion homeostasis 0.03272
coumarin metabolic process 0.01504
CTP biosynthetic process 0.01875
cytoplasmic sequestering of transcription factor 0.00947
cytoskeletal anchoring at plasma membrane 0.00947
deoxyribonucleoside monophosphate metabolic process 0.04698
DNA biosynthetic process 0.03516
DNA damage induced protein phosphorylation 0.03628
DNA damage response, signal transduction by p53 class mediator resulting in transcription of p21 class mediator 0.00199
DNA ligation 0.0363
DNA repair 0.02157
DNA replication-independent nucleosome assembly 0.04639
DNA strand elongation involved in DNA replication 0.00399
DNA strand renaturation 0.02614
double-strand break repair via homologous recombination 5.1E-05
electron transport chain 0.0424
embryonic neurocranium morphogenesis 0.03628

establishment of integrated proviral latency 0.03628
establishment of mitotic spindle orientation 0.02752
establishment of protein localization to chromosome 0.00702
establishment or maintenance of microtubule cytoskeleton polarity 0.04939
female meiosis chromosome segregation 0.01504
female meiosis I 0.04939
folic acid-containing compound biosynthetic process 0.02614
formation of translation preinitiation complex 0.00093
glial cell apoptotic process 0.0363
glucocorticoid receptor signaling pathway 0.02297
glutathione metabolic process 0.04254
glycine metabolic process 0.00707
glycoprotein transport 0.00375
glycosyl compound biosynthetic process 0.03741
Golgi vesicle transport 0.03439
GTP biosynthetic process 0.01504
guanosine-containing compound biosynthetic process 0.01512
heme biosynthetic process 0.01285
histone arginine methylation 0.02614
histone H3-K9 dimethylation 0.03628
histone phosphorylation 0.00535
homologous recombination 0.02554
inner cell mass cell proliferation 0.02297
interstrand cross-link repair 0.00769
IRE1-mediated unfolded protein response 0.04888
iron-sulfur cluster assembly 0.0011
ketone body catabolic process 0.01516
kinetochore organization 0.00388

L-alanine catabolic process 0.04939
L-cysteine catabolic process 0.00375
L-methionine salvage 0.0123
L-proline biosynthetic process 0.04939
L-threonine catabolic process to glycine 0.00375
leukotriene B4 catabolic process 0.01516
meiosis I 0.04004
meiotic chromosome condensation 0.01516
meiotic chromosome segregation 0.03595
mismatch repair 0.01975
mitochondrial DNA repair 0.04939
mitochondrial gene expression 0.04752
mitochondrial respiratory chain complex assembly 0.02452
mitochondrial ribosome assembly 0.01504
mitochondrion morphogenesis 0.01807
mitotic chromosome movement towards spindle pole 0.04939
mitotic sister chromatid cohesion 0.02554
modulation of signal transduction in other organism involved in symbiotic interaction 0.04939
mRNA cleavage 0.01153
mRNA splicing, via spliceosome 3.9E-07
negative regulation of axon extension involved in axon guidance 0.01504
negative regulation of centriole replication 0.02614
negative regulation of endothelial cell apoptotic process 0.03733
negative regulation of fat cell proliferation 0.04939
negative regulation of fibrinolysis 0.02614
negative regulation of histone H3-K9 methylation 0.04698
negative regulation of mRNA splicing, via spliceosome 0.0363

negative regulation of plasminogen activation 0.01504
negative regulation of proteasomal ubiquitin-dependent protein catabolic process 0.02719
negative regulation of protein export from nucleus 0.00093
negative regulation of response to endoplasmic reticulum stress 0.04026
negative regulation of RNA splicing 0.00373
negative regulation of smooth muscle cell migration 0.03815
negative regulation of synapse assembly 0.03628
negative regulation of telomere maintenance in response to DNA damage 0.04939
negative regulation of viral entry into host cell 0.0363
nuclear DNA replication 0.02244
nucleoside diphosphate phosphorylation 0.0297
nucleotide-excision repair 0.01558
nucleotide-excision repair, DNA damage recognition 0.00589
nucleotide-sugar biosynthetic process 0.0036
peptide biosynthetic process 0.03587
peptidyl-arginine methylation, to symmetrical-dimethyl arginine 0.04939
peptidyl-histidine modification 0.04698
peripheral nervous system neuron development 0.03272
phagosome maturation 0.02575
positive regulation by host of viral genome replication 0.01504
positive regulation of anoikis 0.03628
positive regulation of cell cycle G2/M phase transition 0.00823
positive regulation of cell growth involved in cardiac muscle cell development 0.02614
positive regulation of cellular response to insulin stimulus 0.02694
positive regulation of cyclic-nucleotide phosphodiesterase activity 0.00375
positive regulation of DNA-dependent DNA replication 0.04698

positive regulation of glucose mediated signaling pathway 0.01516
positive regulation of inositol 1,4,5-trisphosphate-sensitive calcium-release channel activity 0.04939
positive regulation of interferon-beta production 0.03815
positive regulation of mitotic cell cycle spindle assembly checkpoint 0.03628
positive regulation of myelination 0.0363
positive regulation of myeloid dendritic cell cytokine production 0.04939
positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway 0.01975
positive regulation of protein localization to kinetochore 0.04939
positive regulation of proteolysis involved in cellular protein catabolic process 0.03351
positive regulation of RNA binding 0.02614
positive regulation of sister chromatid cohesion 0.03628
positive regulation of synaptic vesicle recycling 0.02614
positive regulation of telomere capping 0.0363
positive regulation of transcription of nucleolar large rRNA by RNA polymerase I 0.00125
positive regulation of translation 0.02016
positive regulation of viral genome replication 0.00419
positive regulation of viral transcription 0.00769
pressure natriuresis 0.01516
protein autoprocessing 0.02554
protein deneddylation 0.01875
protein export from nucleus 0.03504
protein K11-linked ubiquitination 0.03815
protein localization to chromatin 0.01975
protein localization to extracellular region 0.03628

protein localization to kinetochore 0.0011
protein localization to microtubule cytoskeleton 0.01901
protein maturation by iron-sulfur cluster transfer 0.00038
protein peptidyl-prolyl isomerization 2.4E-05
protein quality control for misfolded or incompletely synthesized proteins 0.01624
protein repair 0.03628
protein sumoylation 0.00355
protein targeting to mitochondrion 0.03768
protein-DNA complex assembly 0.02208
purine nucleobase biosynthetic process 0.0123
pyrimidine deoxyribonucleotide biosynthetic process 0.04939
pyrimidine nucleobase metabolic process 0.03612
regulation of cytokinetic process 0.01516
regulation of DNA damage response, signal transduction by p53 class mediator 0.03815
regulation of double-strand break repair via homologous recombination 0.01443
regulation of endoribonuclease activity 0.04939
regulation of fatty acid beta-oxidation 0.01807
regulation of fibroblast apoptotic process 0.0363
regulation of glycolytic process 0.03974
regulation of Golgi inheritance 0.04939
regulation of intracellular steroid hormone receptor signaling pathway 0.04888
regulation of mitotic cell cycle spindle assembly checkpoint 0.01225
regulation of mitotic cytokinesis 0.01516
regulation of mitotic metaphase/anaphase transition 0.01836
regulation of mRNA binding 0.00947
regulation of peroxisome organization 0.01516
regulation of protein localization to cell cortex 0.04939

regulation of protein neddylation 0.01516
regulation of Rho-dependent protein serine/threonine kinase activity 0.04939
regulation of superoxide dismutase activity 0.04939
regulation of translation at synapse, modulating synaptic transmission 0.04939
regulation of translational initiation 0.04105
replicative senescence 0.00947
response to folic acid 0.04698
response to hydrostatic pressure 0.01504
response to light intensity 0.03272
RNA phosphodiester bond hydrolysis, endonucleolytic 0.0282
SCF complex assembly 0.01504
selenocysteine incorporation 0.03628
serine family amino acid biosynthetic process 0.02297
serine family amino acid catabolic process 0.01495
short-chain fatty acid biosynthetic process 0.04939
spindle organization 0.0352
spliceosomal complex assembly 0.01843
sulfur compound biosynthetic process 0.03733
t-circle formation 0.0363
tail-anchored membrane protein insertion into ER membrane 0.03628
telomerase holoenzyme complex assembly 0.04939
telomere maintenance via recombination 0.03272
transepithelial chloride transport 0.04939
tRNA wobble uridine modification 0.0123
urea cycle 0.01875
UTP biosynthetic process 0.01504
viral translational termination-reinitiation 0.00093
vitamin E metabolic process 0.04939

Table B.4: Unique BP signatures for EE2 *in vivo* exposures from iPathway Guide (Advaita).

GO terms shared by both EE2 *in vivo* exposure from Advaita		
	6 Hours	**9 Hours**
GO:Biological Process	**adj.p-value**	**adj.p-value**
'de novo' UMP biosynthetic process	0.04401	0.01516
5-phosphoribose 1-diphosphate biosynthetic process	0.00203	0.01516
actomyosin contractile ring assembly	0.00203	0.01516
ATP hydrolysis coupled proton transport	0.01193	0.00589
bile acid and bile salt transport	0.00315	0.03799
carnitine shuttle	0.04054	0.03272
cell cycle	0.01262	0.02615
cell division	2.9E-06	2.1E-12
cell redox homeostasis	0.00858	0.04391
cellular respiration	0.04667	0.03333
CENP-A containing nucleosome assembly	0.00649	1E-05
chaperone-mediated protein complex assembly	0.00078	0.00965
ciliary basal body-plasma membrane docking	0.03703	0.00064
cytoplasmic translation	8.1E-08	5.1E-06
DNA damage response, detection of DNA damage	0.00032	3.2E-09
DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	0.01694	0.00411
DNA duplex unwinding	0.00649	0.00035
DNA geometric change	0.00023	0.04848

DNA protection	0.00733	0.04939
DNA replication	0.00257	0.01645
DNA replication initiation	0.00339	3.3E-08
DNA unwinding involved in DNA replication	0.01174	0.00399
double-strand break repair via nonhomologous end joining	0.03029	0.02305
error-free translesion synthesis	0.00208	3.4E-05
error-prone translesion synthesis	0.00131	1.1E-05
establishment of mitotic spindle localization	0.03846	0.04917
exonucleolytic nuclear-transcribed mRNA catabolic process involved in deadenylation-dependent decay	0.02407	0.00509
formation of cytoplasmic translation initiation complex	0.00316	5.2E-07
G1/S transition of mitotic cell cycle	0.0047	0.00334
glucuronate catabolic process to xylulose 5-phosphate	0.01661	0.01504
histone mRNA metabolic process	0.02565	0.00208
hydroxylysine biosynthetic process	0.04401	0.01516
import into nucleus	0.00619	0.01202
intracellular transport of virus	0.00037	0.00044
IRES-dependent viral translational initiation	0.0013	1.4E-05
L-serine catabolic process	0.04401	0.01516
liver regeneration	0.00764	0.01443
mitochondrial respiratory chain complex IV assembly	0.04054	0.00711
mitotic chromosome condensation	0.0173	0.00527
mitotic metaphase plate congression	0.03193	0.01901
mitotic sister chromatid segregation	0.01763	0.00493
mitotic spindle assembly checkpoint	0.04645	0.00031
mitotic spindle midzone assembly	0.01661	0.01504
mRNA 3'-end processing	0.01861	0.00064
mRNA export from nucleus	0.00221	1.7E-07

negative regulation of G0 to G1 transition	0.00762	0.0001
negative regulation of release of cytochrome c from mitochondria	0.0196	0.00499
negative regulation of ubiquitin protein ligase activity	0.01174	0.00399
nucleobase-containing small molecule interconversion	2.4E-05	3E-06
nucleotide-excision repair, DNA gap filling	7.3E-06	3.3E-06
nucleotide-excision repair, DNA incision, 5'-to lesion	0.02687	0.00017
oxidation-reduction process	0.02652	0.0034
phenylalanyl-tRNA aminoacylation	0.04401	0.01516
polyamine biosynthetic process	0.0286	0.01875
positive regulation of ATPase activity	0.01021	0.02174
positive regulation of DNA directed DNA polymerase activity	0.0031	0.00023
positive regulation of endodeoxyribonuclease activity	0.00203	0.01516
positive regulation of establishment of protein localization to telomere	2.3E-06	2.7E-05
positive regulation of intrinsic apoptotic signaling pathway by p53 class mediator	0.00733	0.04939
positive regulation of mitochondrial translation	0.0146	0.03799
positive regulation of mRNA splicing, via spliceosome	0.01174	0.02614
positive regulation of protein localization to Cajal body	5.2E-07	9.8E-05
positive regulation of telomerase RNA localization to Cajal body	1.1E-07	1.1E-05
positive regulation of translational elongation	0.00733	0.04939
positive regulation of type I interferon production	0.03936	0.01383
positive regulation of ubiquitin protein ligase activity	0.01174	0.02614
protein folding	0.01119	0.00442
protein import into mitochondrial matrix	0.00203	0.01516

protein import into mitochondrial outer membrane	0.00026	0.00375
protein neddylation	0.01169	0.01226
purine ribonucleoside monophosphate catabolic process	0.01598	0.04939
pyrimidine nucleotide catabolic process	0.01904	0.04698
Ran protein signal transduction	0.00203	0.01516
regulation of attachment of spindle microtubules to kinetochore	0.0286	0.01875
regulation of cellular response to heat	0.01865	1.4E-05
regulation of centrosome duplication	0.0224	0.03707
regulation of cyclin-dependent protein serine/threonine kinase activity	0.00232	0.03357
regulation of G2/M transition of mitotic cell cycle	0.03734	1.8E-05
regulation of gene silencing by miRNA	0.03797	0.0068
regulation of signal transduction by p53 class mediator	0.02736	0.00345
regulation of transcription involved in G1/S transition of mitotic cell cycle	0.01193	0.00033
replication fork processing	0.00649	0.00208
response to cadmium ion	0.00366	0.02174
response to thyroxine	0.04401	0.01516
ribosomal large subunit assembly	0.00469	0.00329
ribosomal protein import into nucleus	0.00733	0.04939
ribosomal small subunit biogenesis	0.00465	0.02543
RNA processing	8E-05	0.00054
scaRNA localization to Cajal body	0.01661	0.01504
spindle assembly involved in meiosis	0.04401	0.01516
spliceosomal snRNP assembly	0.03524	0.00032
T cell receptor signaling pathway	0.00526	0.00157
telomere maintenance	0.00559	0.04064

telomere maintenance via semi-conservative replication	0.03143	1E-07
termination of RNA polymerase II transcription	0.03846	0.00136
toxin catabolic process	0.04401	0.01516
transcription-coupled nucleotide-excision repair	0.00549	4.3E-06
translational elongation	0.00512	0.00095
translocation of molecules into host	0.01904	0.04698
tRNA export from nucleus	0.03524	0.00367
tRNA splicing, via endonucleolytic cleavage and ligation	0.00131	0.01624
viral process	0.00045	9.9E-05

Table B.5: Shared BP by EE2 *in vivo* exposures from iPathway Guide (Advaita).