ERα AND AHR MEDIATED GENOMIC EFFECTS OF KARANJIN IN BREAST CANCER CELLS

A THESIS

Submitted By

GAURAV BHATT

for the award of the degree

of

DOCTOR OF PHILOSOPHY

Department of Biosciences and Bioengineering

Indian Institute of Technology Guwahati

Guwahati, Assam-781039, India

October 2023

DEPARTMENT OF BIOSCIENCES AND BIOENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI, ASSAM-781039

DECLARATION

The thesis titled **"ERα and AhR mediated genomic effects of karanjin in breast cancer cells"** represents my original research work carried out in the Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, India, under the supervision of Dr. Latha Rangan and Dr. Anil Mukund Limaye.

Sincere efforts have been made to acknowledge contributions from other investigators that helped in conceptualizing and executing the research work. Those who have provided suggestions and technical help have been duly acknowledged. All the research articles and resources used have been cited in the reference section.

3rd October 2023

Gaurav Bhatt
Roll No. 166106017
Department of Biosciences and Bioengineering
Indian Institute of Technology Guwahati

DEPARTMENT OF BIOSCIENCES AND BIOENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI, ASSAM-781039

CERTIFICATE

This is to certify that the work described in the thesis titled, **"ERα and AhR mediated genomic effects of karanjin in breast cancer cells"** submitted by Gaurav Bhatt (Roll No. 166106017) to the Indian Institute of Technology Guwahati, India, for the award of the degree of Doctor of Philosophy is an authentic record of the research work carried out under my supervision in the Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, India.

This thesis or any part thereof has not been submitted elsewhere to award any other degree or diploma.

3rd October 2023

Dr. Anil Mukund Limaye
Supervisor

Dr. Latha Rangan
Supervisor

DEPARTMENT OF BIOSCIENCES AND BIOENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI, ASSAM-781039

Acknowledgments

With deepest respect and esteem regards, I owe my profound gratitude to my research supervisors Dr. Latha Rangan and Dr. Anil Mukund Limaye. I consider it an honor to work under their invaluable guidance and to learn from them. Their persistent guidance, encouragement, and faith in me is the reason I could write this thesis. I thank them for shaping me into a better researcher and human being. They taught me how to be punctual and insightful, interpret minute details and analyze things differently—my true GURUs. I feel so grateful to my supervisors for always being supportive and for allowing me the freedom to plan and execute my experiments, which nurtured me to be a better researcher and human being.

I want to express my sincere thanks to the Doctoral Committee members, Prof. Ajaikumar B. Kunnumakkara, Prof. Shankar P. Kanaujia, Prof Aditya N. Panda, and past DC member Prof. Vikas K. Dubey for their constructive suggestions and direction provided during my progress seminar reviews, which enabled me to complete my Ph.D. thesis successfully.

I thank the past and present HODs of the Department of Biosciences and Bioengineering (BSBE), Prof. Kannan Pakshirajan, Prof. Latha Rangan, and Prof Rakhi Chaturvedi, for their valuable support.

I am thankful for infrastructural support from BSBE and DBT-funded Bioinformatics Facility (BIF), IIT Guwahati. I want to thank the funding agencies that gave me the financial aid to perform my research. I thank the Ministry of Education and SERB DST, Government of India.

I want to thank the staff of BSBE (Nurul sir, Dipankar sir, Prarthana Ma'am, Binoy sir, and Niranjan sir) BSBE for their ever-present support.

My sincere thanks to my previous Gurus, Dr. Preeti, Prof. Hemant Sood, Prof. Sudhir Syal, Prof. Rajinder Chauhan, and Dr. Vimarsh Raina, for inculcating my passion in science and research.

My sincere thanks to Dr. Sachin Kumar, Dr. Rajkumar Thummer, Dr. Kusum Kumari Singh, and Dr. Shririsha Nagatu for providing access to their lab facilities.

I would like to express my deepest gratitude to my seniors, Dr. Mohan, Dr. Ajay, Dr. Sheeba, Dr. Arun, Dr. Anuma, Dr. Ishani, Dr. Rahul, Dr. Reshmi, and Dr. Sanjana, for their constant support, care, and for helping me to make the right decisions in crucial moments of my research and teaching me everything from scratch.

I am particularly grateful for the support provided by Uttariya Pal, Dr. Manish Kumar Gupta, and Akshita Gupta in designing my experiments and providing me assistance in all ways possible.

It is an immense pleasure to thank our Molecular Endocrinology lab members, Dr. Snigdha, Sonya, Juana, Musfica, Pavan, Harsha, Ayesha, Sreeja, Swati, Sarbojit, and Chinni, for all their help in lab work and their constructive comments, discussions.

Thank you, ABL lab members Nuzelu, Bimal sir, Heera, Alok, Rashmi, Anamika, Rubeka, and Sonu, for all the good moments we had together and for the constant encouragement during tough times and for supporting me.

Special thanks to Prajakta Ma'am for her concern and encouragement for giving us memorable lab gatherings.

Special thanks to Dr. Sudip Mitra and my younger brother Shreyas for being my family here in Guwahati. For wonderful trips, Football matches, and delicious food outings.

To my football family here in Guwahati for giving me so much respect, love, and friends for a lifetime.

I thank Sukumar ji, Mahesh, Saddam, Riddhi, Angshu, Satakshi, Darshana, Dr. Debojit, and all my friends whom I couldn't mentions but hold significant place in my Ph.D. journey for their incessant support and for being part of my sound and challenging times.

Dabbu, Neelesh, and Jiban, the three musketeers and gem I found in IITG, my best friends and love.

Finally, I want to thank Ma and Papa for being so enduring and my family for their constant support and encouragement. What I am today is all because of my Ma and Papa. I feel so grateful to have such loving and caring parents. Ma, you are my first teacher, my first Guru, and it trickles down to you, whose constant encouragement and love is the reason I could write this thesis today.

I am incredibly grateful to the Almighty. God Bless all.

Gaurav Bhatt

To my beloved

Ma, Papa

and

My family

Table of contents

CHAPTER 1

Introduction

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

1.1 Introduction

Karanjin is a member of the furanoflavonoid class of natural compounds and occurs widely in the Leguminosae family. It is the most abundant bioactive compound in *Pongamia pinnata* seed oil and produces a myriad of biological effects *in vitro* and *in vivo* [1]. The seed oil has numerous medicinal properties and has been utilized in Ayurveda, Siddha, and traditional medicine for centuries. The powdered seed and oil of the plant are used in ameliorating bronchitis, chronic fever, whooping cough, ulcer, leukoderma, chronic skin diseases, abdominal tumors, keloid tumors, and painful rheumatic joints [2,3]. *Pongamia* seed extracts found in various herbal preparations are widely available in the market, and the majority of the therapeutic properties observed are attributed to karanjin [4].

Karanjin harbors a multitude of health benefits and applications, such as antioxidant, anti-bacterial, anti-cancer, anti-ulcer, antihyperglycemic, and anti-inflammatory activities [4–7]. Recent demonstrations of cell-cycle inhibitory and pro-apoptotic effects of karanjin *in vitro* have fueled speculations about its anticancer potential [8–10]. Recently, Yu and coworkers demonstrated that karanjin exhibited PI3K/Akt-mediated inhibition of cell viability in MDA-MB-231 breast cancer cells [11]. All cancer cell lines tested so far are growth-inhibited by karanjin with varying inhibitory concentration (IC_{50}) values. The argument that karanjin exerts low toxicity towards standard mouse embryonic fibroblasts [9,12] and animal models [1] perpetuates its use as a drug or active constituent in anticancer therapy. Given the flavonoid structure of karanjin, its detrimental effect on cell proliferation and growth is not surprising. Moreover, flavonoids are known to possess dual nature, both proliferative and antiproliferative, owing to their binding to diverse receptors [13]. Thus, the universality of the anti-tumor efficacy of karanjin cannot be concluded merely based on its antiproliferative effects demonstrated on a few cancer cell lines. The flavonoid structure of karanjin presents a caveat to its potential in anticancer therapy, lest it be counterproductive in hormone-dependent tumors.

Flavonoids exert their hormone-like effects through nuclear receptors (e.g., estrogen receptors [ERs]), non-nuclear steroid hormone receptors (e.g., membrane ERs), and other orphan receptors (e.g., aryl hydrocarbon receptor [AhR]) [14]. Flavonoids, including genistein, daidzein, S-equol, and liquiritigen, have been reported

to act via both ERα and AhR receptors, with their AhR activity moderated by negative crosstalk through ERα [15]. Several research groups have established the effects of flavonoids on gene regulation and target cell behavior via modulating ERα and AhR. Both indicate that receptors may play an essential regulatory role in the pharmacology and biology of flavonoids [15,16].

Due to the intrinsic property of flavonoids to modulate ERα, flavonoids are deemed natural Selective Estrogen Receptor Modulators (SERMs) [17,18]. SERMs are valuable therapeutic agents (e.g., tamoxifen, raloxifene) and have also served as tools to understand mechanisms of estrogen signaling and its effect on target cells and tissues. The actions of SERMs are known to be cell/tissue-dependent or even gene/promoter context-dependent, which makes them molecules of great utility [19]. This property of SERMs has led to their use in various therapeutic applications, including breast cancer prevention and treatment, osteoporosis management, and cardiovascular disease prevention [20,21]. Some SERMs, such as raloxifene, act as AhR agonists or selective AhR modulators (SAhRMs). There lies an intricate interplay between AhR and ERα, and drugs targeting ERα and AhR are of great interest in health and therapeutics [22,23]. Flavonoids have been indicated to modulate AhR and ERα signaling pathways, making them potential therapeutic agents. In addition, flavonoids can be found in many natural sources and deemed safe, making them an attractive alternative to synthetic drugs. Flavonoids such as berberine, quercetin, and resveratrol have been reported to be natural ER and AhR modulators [16] though their mechanism of action is still not well understood.

Karanjin, a major flavonoid from *P. pinnata*, has been reported to have various therapeutic applications [8,10,11]. However, the underlying mechanism behind karanjin-mediated gene modulation and the impact on global gene expression in breast tumors remains unaddressed. In summary, the present investigation provides valuable insights into the role of karanjin on breast cancer cells by delineating the possible influence and involvement of both ERα and AhR in karanjin-induced molecular and cellular effects. This study, while exposing a caveat to the anticancer potential of karanjin, will inspire further investigations into the possible use of karanjin or its derivatives in endocrine or combination therapies.

1.2 Aim and scope of the present work

Since estrogen is known to be tumor promoting agent [24,25], it increases the risk of breast and uterine cancer in women administered estrogen replacement therapy [26]. Flavonoids possess weak estrogenic properties [18,27,28] and may have a protective effect against breast and uterine cancer by competing with estrogen for binding to receptors [13,29]. For these beneficial effects, flavonoids are considered a safe alternate and dietary component, but their mechanisms of action are poorly understood [30]. Further research is needed to fully understand the potential benefits and risks of utilizing flavonoids as a safer alternative.

The effect of karanjin on hormone-responsive breast cancer cells and its impact on the modulation of gene expression is poorly understood. Hence, describing karanjin's cellular actions is critical, particularly considering its flavonoid nature. The mechanism of action of karanjin is yet to be studied extensively. However, karanjin has shown promising results in inhibiting the growth of cancer cells in numerous preclinical studies, much of our understanding deals with karanjin's anti-proliferative and cell-cycle inhibitory effects [1,4]. The proposed study seeks insights into karanjin's selective ER modulator (SERM) and selective AhR modulator (SAhRM)-like characteristics utilizing *in silico* and *in vitro* experimental approaches. It also attempts to explore the mechanistic details of karanjin-mediated regulation of target genes employing various molecular biology techniques.

An estrogen-like effect of karanjin implies that there is a possibility of an increased risk of breast cancer if used indiscriminately. Therefore, further studies are needed to determine the safety and efficacy of karanjin as a potential therapeutic agent. It is crucial to consider the potential risks and benefits before using karanjin for medical purposes. On the other hand, karanjin or its modified derivatives may also provide valuable SERMs that may serve as an alternative to the current estrogen replacement therapy. Furthermore, efforts can be made to harness karanjin's potential AhR binding property to deliver pharmacophores and generate novel AhR modulators, which could be used for various therapeutic interventions.

1.3 Objectives

The work embodied in the thesis is based on the following objectives:

1. To study global gene expression profiling of estrogen-responsive cells treated with karanjin.
2. To investigate the partial estrogen-like effects of karanjin and the mechanism involved.
3. To delineate the role of AhR in karanjin-mediated effects.

CHAPTER 2

Review of literature

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

2.1 Introduction

The plant kingdom contains a plethora of therapeutic or pharmacologically active constituents that can be used to combat various life-threatening diseases [31]. More than 65% of people still use phytochemicals as a source of medicine [32,33]. Nearly half of all pharmaceuticals are derived from natural sources, and several well-known medications, such as aspirin and artemisinin, have their roots in alternative medicine [33]. Herbal remedies, which are based on traditional and natural therapies, have recently gained popularity due to their inclusion in alternative, complementary, and holistic medical applications [34]. In India, natural products account for around 70% of current drugs, and many synthetic counterparts have been developed using inspiration from herbal products [33]. Roughly 75% of anticancer drugs are made from natural bioactive compounds, most of which come from plants [32]. India is known for its rich biodiversity and is home to approximately 7-8% of the world's total plant species. The traditional knowledge of using plant-based remedies has been passed down through generations and continues to play a significant role in the healthcare practices of India. There are 45,000 plant species, among which 15,000-20,000 have medicinal value, and 7,500 are used in traditional medicine practices in India [35] (Fig. 2.1). A plethora of active agents from medicinal plants can be extracted and employed in the pharmaceutical and industrial sectors.

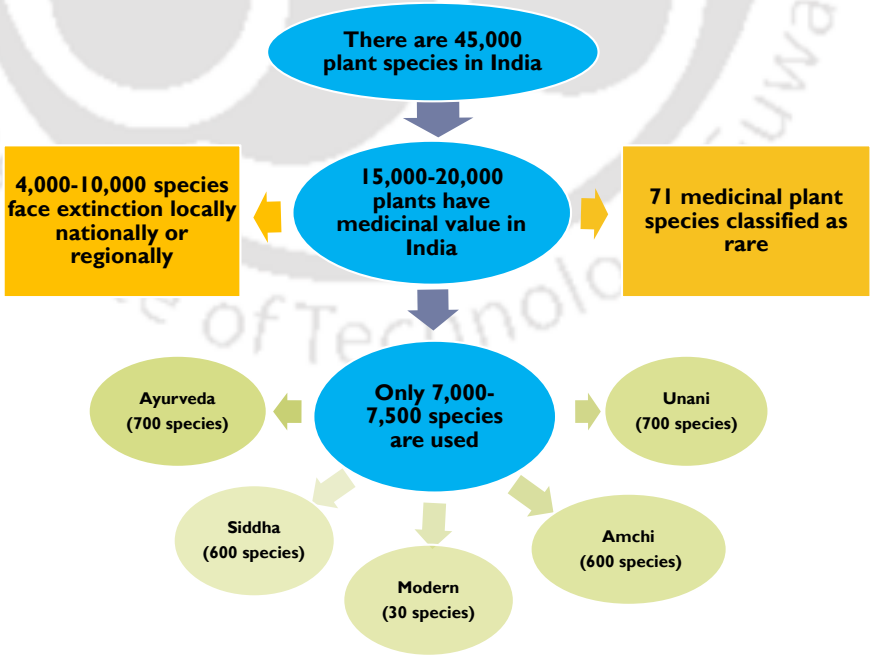


Fig. 2.1: A schematic of details of plant-species in India. The figure shows the total estimated number of plant species in India based on the latest available data from the Botanical Survey of India. It also describes plant species used in different folk and traditional medicine.

2.2 Plant-derived natural compounds

Plants produce primary and secondary metabolites to help them adapt to various environmental conditions. Primary metabolites are amino acids, nucleic acids, lipids, and other essential components that plants produce and utilize for their growth and development. On the other hand, secondary metabolites are assumed to be metabolic by-products that might induce biochemical and molecular effects on biological systems. Plants produce specialized metabolites for several reasons, including defense against biotic (diseases and pests) and abiotic (environmental) stress. Specialized metabolites act as an intrinsic modulator of plant growth hormone and have multiple commercial applications such as a coloring agent, essential oil, flavoring component, insecticide, medicine, tanning agent, perfume, or cosmetic [36–38]. Plant-based natural compounds are classified into terpenes, phenols, organosulfur compounds, organic acids and polysaccharides, and lipids (Fig. 2.2).

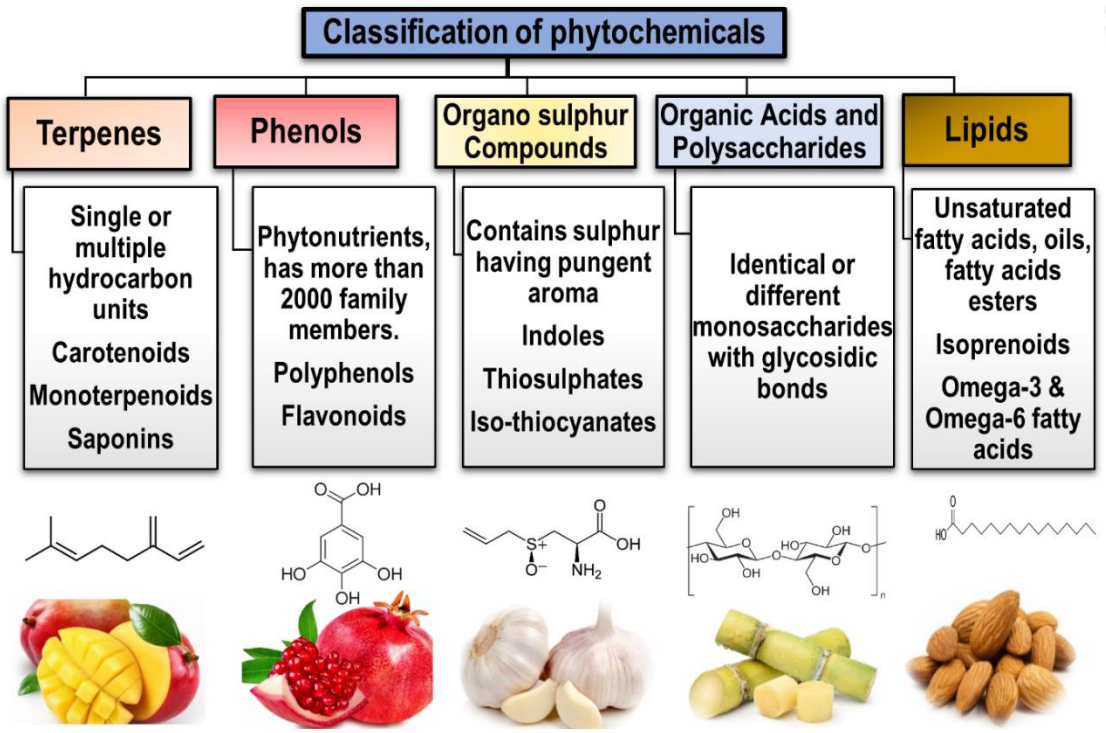


Fig. 2.2: Classification of phytochemicals. Phytochemicals are classified into several categories based on their chemical structure and biological activity. These categories include terpenes, phenols, organosulfur compounds, organic acids and polysaccharides, and lipids.

2.3 Flavonoids and their biological activity

Flavonoids are natural components plants produce in response to defense against different stresses. Flavonoids are the most common and essential polyphenolic compounds, as they are vital components responsible for human and animal health. Their major sources include fruits, vegetables, tea, wines, cereals, etc. [39]. Hence their occurrence is only through dietary intake. More than 10,000 flavonoids have been identified, many of which are responsible for the attractive colors of flowers, fruits, and leaves [40].

Numerous studies have investigated the efficacy of flavonoids and their metabolites in the treatment and prevention of a wide range of illnesses, including malignancies, obesity, diabetes, hypertension, hyperlipidemia, cardiovascular diseases, neurological disorders, and osteoporosis [41]. Flavonoids demonstrate antibacterial [5], antioxidant [37], anti-viral [42], and anticancer [9] properties. Plants produce flavonoids in response to environmental stress and microbial infection; hence, they act as potent antimicrobials *in vivo* and *in vitro* [43,44]. Several pathways have been hypothesized for flavonoids' impact on cancer onset and promotion, including a critical role in the development and hormonal activity [45].

2.4 Karanjin, a major flavonoid present in seeds of *Pongamia pinnata*

The bioactive furanoflavonoid karanjin [IUPAC: 3-methoxy-2-phenylfuro-(2,3-h-chrome-4-ol)] is a potent biomolecule with a wide range of biological effects [1]. It is one of the most potent compounds present in Karanj seed oil [4]. Karanjin exhibits multiple health benefits and applications, with evident anti-diabetic, anti-cancer, anti-viral, anti-inflammatory, anti-hyperglycemic, antioxidant, anti-colitis, anti-ulcer, and anti-Alzheimer properties [1,4]. Investigation into the structure-activity relationship indicated that karanjin posed a minimal danger to human health across four primary characteristics: mutagenicity, tumorigenicity, irritancy, and reproductive efficacy. Toxicological assessments have determined that karanjin is non-toxic under physiological conditions [12], supported by detailed *in vitro* and *in vivo* analysis [1].

Karanjin is a characteristic of Leguminosae and has been reported in the roots of several legumes like *Tephrosia purpurea* [46], *Fordia cauliflora* [8,47], *Millettia pulchra* [47], *Lonchocarpus latifolius* [48], *Millettia pachycarpa* [49] and *Desmodium*

sequax [50]. It is well documented that *P. pinnata* seeds are the traditional and popular extraction source, yielding up to 2% karanjin, the highest for any plant source [5,51]. Sriphana and co-workers reported karanjin yields as low as 0.00064% from the fruit of *Millettia leucantha* [52]. The fact that furanoflavonoids like karanjin are found in many legumes demonstrates that they are essential for the plant's metabolism and survival since they protect it from herbivores and pathogens. Karanjin's production and distribution in Leguminosae may be viewed as an adaptation to deal with biotic and abiotic stresses and enhance modulatory effects [2,53,54]. The structure of karanjin comprises of a furan ring and methoxy group on a flavone backbone, which imparts hydrophobicity and semi-polarity (Fig. 2.3). Hence, karanjin is soluble in various solvents, such as dimethyl sulfoxide (DMSO), ethyl acetate, isopropanol, dichloromethane, ethanol, methanol, benzene, and petroleum ether, among several others [5,55].

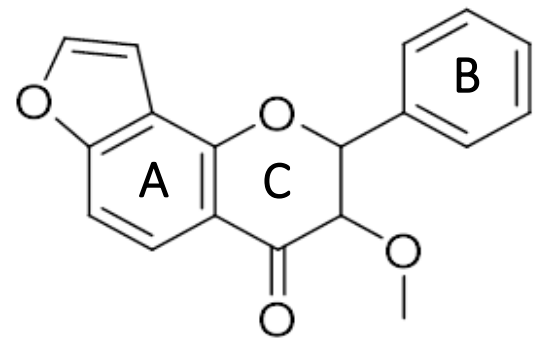


Fig. 2.3: Structure of karanjin. Karanjin comprises a furan ring and a methoxy substituent group tethered to the flavonoid backbone. The flavonoid backbone of karanjin consists of two aromatic rings, ring A and ring B, linked by a three-carbon chain that forms an oxygenated heterocyclic ring C.

2.4.1 Karanj or Pongamia pinnata

P. pinnata is a versatile perennial plant that belongs to the third largest family of flowering plants, Fabaceae. Other names include *Millettia pinnata*, *P. pinnata*, *Pongamia glabra, Derris indica, etc.* It is commonly known as Karanja, Honge, Karanj, Pongam oil tree/malva nut, etc. [2]. *P. pinnata* is a fast-growing, oil-yielding legume tree indigenous to the Indian subcontinent. Its seed contains 30-40% oil and is a rich source of protein (20-28%). Its oil consists of 5-6% flavonoids [56]. In folk medicine, fruits are used against abdominal tumors [57] and keloid tumors [58]. Karanj has been broadly applied in Ayurveda and traditional medicinal practices for various diseases, including skin ailments and rheumatism [2]. Moreover, Karanj is also used for anti-

inflammatory, anti-plasmodial, anti-nociceptive, anti-hyperglycemic, anti-lipid peroxidative, anti-diarrheal, anti-ulcer, antioxidant, and antibacterial activity [1]. *P. pinnata* contains various secondary metabolites such as flavonoids, terpenoids, alkaloids, etc., which are responsible for their medicinal properties, and almost all parts of the plants are frequently employed for a wide range of applications [3] (Fig. 2.4). Due to its multiple utilities, Karanj has gained much attention from scientists and has encouraged towards the development of potential therapeutics. Karanjin is the most potent and abundant bioactive in Karanj and is suggested to be the major contributor to its diverse pharmacological effects [4].

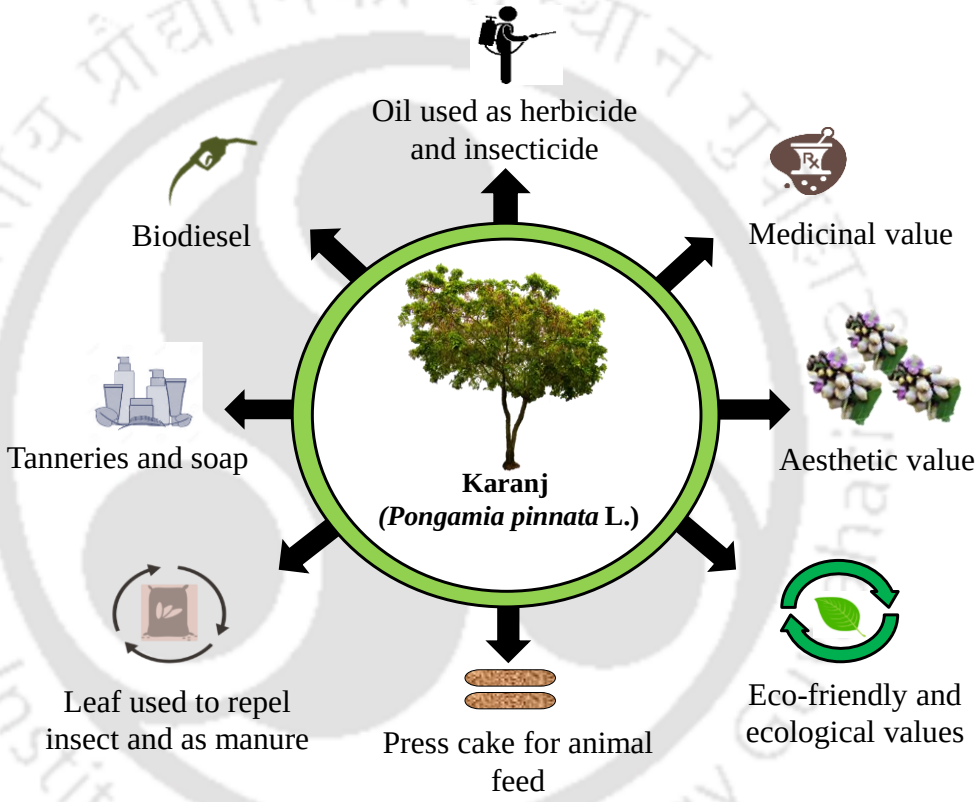


Fig. 2.4: Karanj applications in diverse fields (Source:[1]).

2.4.2 Extraction and isolation of karanjin

Karanjin was first isolated from *P. pinnata*. It has been close to a century since the isolation and characterization by Prof. Limaye in 1925, which became a vital compound of this species due to its high yield (2%) from the seed oil [59]. Later, many investigators isolated karanjin from other parts of the Karanj and other legumes such as the flower, stem bark, and root bark [1,51,60–62].

Rao *et al*. (1939) isolated karanjin by precipitation from the alcoholic extract obtained by country-pressed kernels for 30 h [63]. Row *et al*. (1952) extracted *P.*

pinnata oil using light petroleum ether and separated the products by fractional crystallization and chromatography, followed by extraction with hydrochloric acid to obtain karanjin [64]. Chromatographic techniques like thin-layer chromatography (TLC), column chromatography, and reverse-phase high-performance liquid chromatography (RP-HPLC) have been employed to isolate and purify karanjin using methanol, water, and acetic acid as the mobile phase [65–67]. Karanjin isolation by liquid-liquid extraction method from various crude extracts such as petroleum ether seed extract [36,66,68–70], hexane extract [71], methanolic extract [72–74], ethyl acetate extract [5,51] followed by appropriate purification of the compound by chromatographic techniques are well documented. These extraction methods proved to be efficient in terms of yield, which ranged from 0.32-1.5%. These advanced chromatographic techniques are rapid, sensitive, and accurate for the efficient isolation of karanjin compared to conventional methods.

2.4.3 Biological activities of karanjin

Studies have shown numerous pharmacological actions of crude extracts from different parts of *P. pinnata*. Karanjin is believed to be one of the most studied and active ingredients with a broad spectrum of pharmacological properties. Interestingly, karanjin is well examined for a deeper understanding of its biological activity in both *in vitro* and *in vivo* systems (Fig. 2.5 figure placed two pages below). A recent review has also highlighted the multiple biological properties of karanjin [1]. Utilizing structure-based prediction tools such as OSIRIS Property Explorer and Molinspiration, physicochemical or drug-like properties such as lipophilicity (cLogP), solubility (LogS), molecular weight, drug-likeness, overall drug-score, and percentage of absorption of karanjin were determined [5]. They play an important role in absorption, distribution, and desired physiological actions. The structure of karanjin demonstrated desirable physicochemical features and nontoxic behavior, disclosing its potential therapeutic value. Karanjin also confirmed compliance with the five Lipinski criteria, indicating that the molecule would have no concerns with oral bioavailability. An investigation by Shejawal and co-workers (2014) on the pharmacokinetics of karanjin on female Sprague-Dawley rats (aged 10-12 weeks; b.w. 200-250 g) demonstrated the absorption and elimination pattern of karanjin. Karanjin from *P. pinnata* ethanolic extract equivalent to a dosage of 10 mg kg^{-1} showed mean maximal plasma

concentrations (C_{max}) and elimination half-time ($t_{1/2}$) values of 0.498 ± 0.01 µg ml^{-1} and 3.78 h, respectively [75]. The study also showed karanjin to be eliminated within 24 h from systemic circulation.

Yi and co-workers (2015) have attempted to perform preclinical pharmacokinetic studies of karanjin on eighteen male Sprague-Dawley rats. The drug in plasma peaked at ~1.5 h post-oral dose and diminished gradually with a $t_{1/2}$ between 1.9 and 2.2 h. Karanjin at 5, 10, and 20 mg kg^{-1} exhibited C_{max} of 187.5 ± 25.2, 347.0 ± 38.2, and 772.0 ± 121.0 ng $mL^{-1,}$ respectively. These findings suggest no significant adverse reaction, indicating karanjin's safety and nontoxic potential within the dosage range, thus potentiating its use as a drug candidate in the future [1].

Studies have shown numerous medicinal uses of different parts of Karanj and have encouraged researchers to explore novel applications of the extracts or purified compounds as therapeutic agent for an effective medication for treating infectious or deadly diseases (Fig 2.5). Karanjin is the most studied compound in Karanj and possesses broad-spectrum biological activities, viz., antioxidant, antimicrobial, anti-ulcerogenic, antiviral, and anticancer. At the cellular level, karanjin induces apoptosis and cell cycle arrest by inhibiting cell proliferation [8,9,11]. It significantly inhibits the activity of the human cytochrome P450 family of enzymes and PTP-1β enzyme activity; karanjin thus elicits chemo preventive nature and anti-hyperglycemic activity. Since karanjin inhibits NF-κB, alterations in TNFα signaling are a distinct possibility [74,76]. In addition, karanjin is known to possess anti-ulcer, anti-colitis, and anti-Alzheimer activity [7,77,78]. Most of these activities of karanjin are attributed to its antioxidant nature and ability to quench free radicals, thus preventing DNA damage. Karanjin also plays a role in drug efflux and metabolism via regulating ABC transporters and cytochromes [74,79]. Therefore, in combination with known drugs, karanjin could be used to enhance their efficacy and sensitivity. Various *in vitro* and *in vivo* studies have revealed low toxicity, ensuring its safety and eligibility to be used in modern drug systems. Also, Karanj crude extracts and karanjin are being used as biopesticides in integrated pest management. It serves as a potential alternative to its chemical counterparts due to its environmentally non-pollutive, renewable, and easily biodegradable nature, thereby resulting in lesser toxic residues and largely avoiding pollution problems associated with chemical pesticides [1].

Its efficacy and low toxicity towards mammalian cells and animal models perpetuate its use in various formulations and as a drug. Multiple studies have already reported the role of karanjin against cancer, ulcers, Alzheimer's, psoriasis, and diabetes. However, a deeper insight into the phytochemistry and exact mechanism of action needs to be elucidated [1].

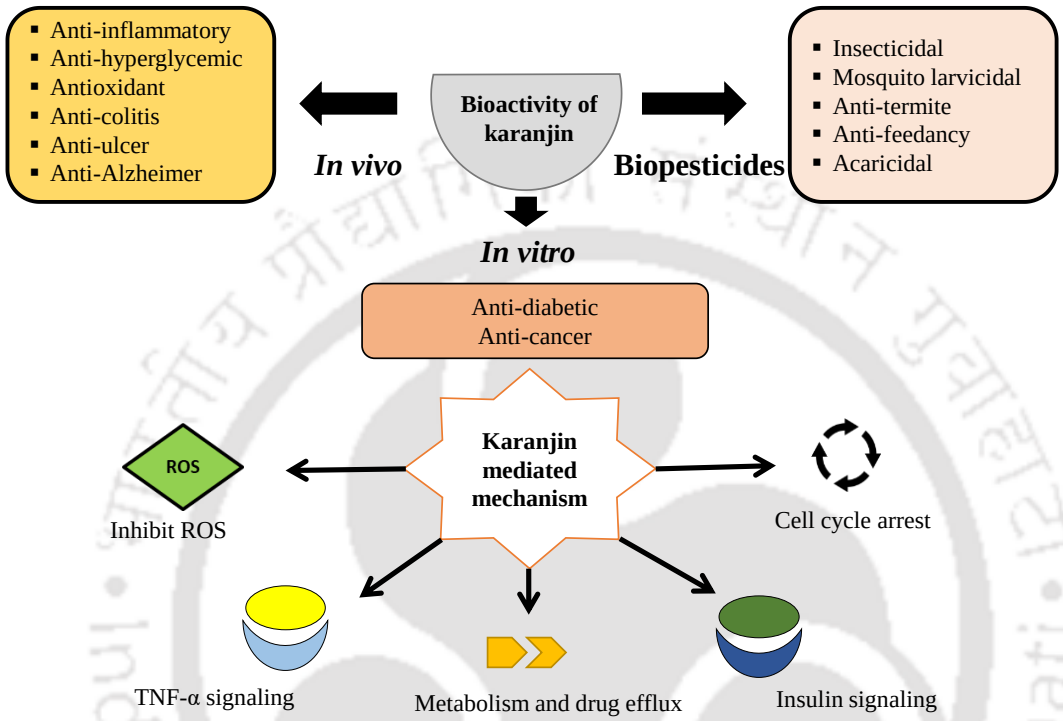


Fig. 2.5: Karanjin effect *in vitro* and *in vivo* and regulation of signaling pathway. Karanjin possesses a wide range of biological activities *in vitro* and *in vivo*, including antimicrobial, anti-inflammatory, and insecticidal properties. The mechanism of action for these effects involves the modulation of various signaling pathways in cells (Source: [1]).

Guo and coworkers have investigated the anticancer effects of karanjin against the human lung adenocarcinoma cell line (A549), human hepatocellular carcinoma cell line (HepG2), and human acute promyelocytic leukemia cell line (HL-60). Karanjin inhibits proliferation and viability in three human cancer cell lines (HepG2, A549, and HL-60 cells). Karanjin also induced cell cycle arrest in the G2/M phase and enhanced apoptosis, as confirmed by a flow cytometric study. This study was the first to reveal the potential of karanjin to cause cancer cell death by inducing cell cycle arrest and apoptosis [8]. Recently, Yu and co-workers reported the antitumor activity of karanjin in MDA-MB-231 cells by blocking the PI3K/Akt signaling pathway [11]. According to Michaelis *et al*. (2014), karanjin interferes with drug efflux mediated by ABCB1,

ABCC1, and ABCG2 transporters by enhancing their ATPase activity; therefore, it might have a role in anticancer therapy [79]. In a study on MCF-7 and HeLa cell lines with Karanj seed oil, the cells showed morphological changes such as cell shrinkage in the dose and time-dependent manner. MTT assay result exhibited a decrease in cell viability [80]. DNA fragmentation was found to be more in the case of Karanj oil-treated samples. All these data may be attributed to the inhibition of growth and multiplication by Karanj oil [80]. Roy and coworkers have reported a decrease in ROS level in HeLa cells treated with karanjin, inhibiting degradation of the nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor (I-κB), thus restricting nuclear translocation of NF-κB. Moreover, karanjin treatment leads to an increase in p53 expression, thus exhibiting low DNA damage. Karanjin induces G2/M arrest and an increase in the SubG1 population; it also induces caspase-dependent apoptosis by induction of the Bax/Bcl-2 ratio through a low expression of Bcl-2 [9,10] (Fig. 2.6).

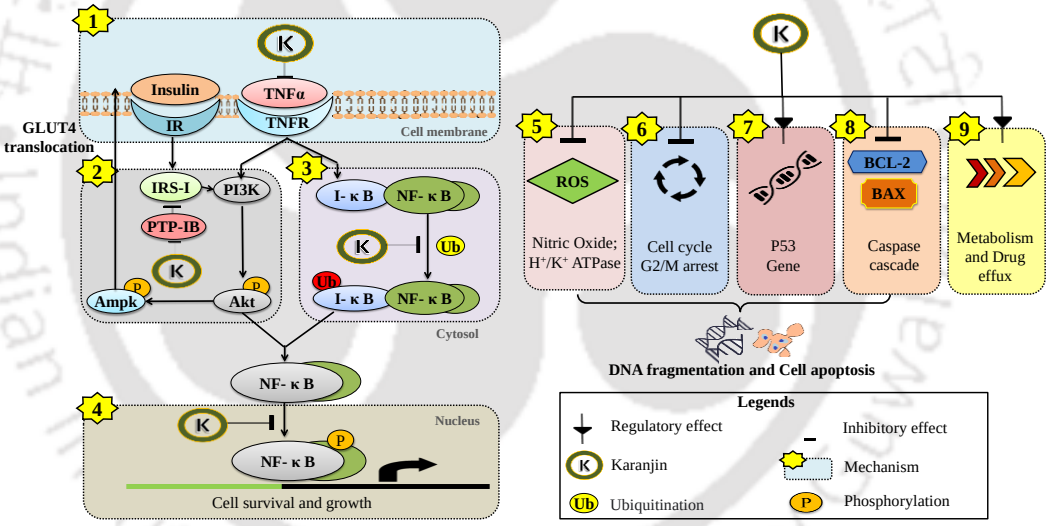


Fig. 2.6: Mechanism of karanjin action: 1. Karanjin suppresses TNFα production[81]**.** Karanjin inhibits Protein Tyrosine phosphatase 1B (PTP-1B), which suppresses IRS-1 [82], activates AMPK in PI3K/AKT independent manner, and induces GLUT4 translocation [76]. **3.** Karanjin inhibits the degradation of I-κB [9] **4.** Karanjin inhibits NF-κB translocation to the nucleus [9,83] **5.** Karanjin suppresses ROS [9,68,77,84–86] **6.** Induces cell cycle arrest, G2/M arrest, and increase in subG1 population [8–11] **7.** Karanjin-induced P53 expression [9] **8.** Karanjin-induced BAX/BCL2 ratio by inhibiting BCL2 expression [9]. **9.** Involved in drug efflux via ABC transporters and metabolism via Cytochrome P450 family enzymes [74,79] (Source:[1]).

Crude Karanj extract is used in the pesticide industry with different trade names, such as Derisom^{TM} as a potential bio-insecticide and bio-pesticide. Though karanjin is

reported to be an active constituent for its pesticidal activity, profound studies are required to elicit the exact mechanism of karanjin inhibition against various pests and insects. Traditionally, Karanj oil is used in multiple formulations in Ayurvedic and Siddha systems of medicine against various skin ailments and has recently been marketed under the tradename Devinez^{TM} and Naturalis^{TM}. Hence there is a scope to study the role of Karanj oil and karanjin in various skin diseases and disorders. Lastly, karanjin holds significant potential to be explored further for its biological activities against various chronic diseases and ailments.

2.5 Flavonoids: Estrogenic and anti-estrogenic effect

Since estrogen is known to be tumor promoting agent, it increases the risk of breast and uterine cancer in women administered estrogen replacement therapy [87]. Various flavonoids, due to structural similarity to estrogen, possess weak estrogenic properties. This property has led to research on their potential use in hormone replacement therapy and as a natural alternative to synthetic estrogen. Flavonoids affinity for binding to estrogen receptors is 10,000 times less than Estradiol (E2). Some flavonoids show an affinity for ER β [88]. On binding to ER, flavonoids act as agonists or antagonists. As estrogen mimetic, dietary flavonoids such as soy isoflavones showed protective functions like lowering total serum cholesterol, reducing the incidence of cardiovascular diseases, osteoporosis, and neuro-protective and anticancer effects [89]. Hence, there is further scope of research wherein flavonoids can be explored as an alternative to estrogen in hormone replacement therapy. Many legume-derived compounds possess this phytoestrogen-like potential [30]. For these beneficial effects, flavonoids are considered safe dietary components, but their mechanism of action is poorly understood.

In mammals, flavonoids have dual effects: anti-estrogenic and estrogen-mimetic [13]. This means that flavonoids can either block or mimic the effects of estrogen in the body, which can have important implications for hormone-related conditions such as breast cancer and osteoporosis. The estrogenic effects of flavonoids are attributed to their structural similarity with E2, due to which they bind to nuclear receptors [90,91].

2.6 Estrogen receptor and signaling

Estrogen receptors are ligand-activated transcription factors which belongs to the nuclear hormone receptor superfamily [92]. They serve an essential role in regulating the growth and development of various tissues and organs, including breasts, uterus, and bones [93]. Estrogen receptors are classified into ERα and ERβ isoforms encoded by *ESR1* and *ESR2* genes and located on chromosomes 6 and 14, respectively [94], and exhibit distinct ligand binding, tissue distribution and function [95]. ERα is abundant in many organs, including the uterus, liver, vagina, and pituitary. ERα is known to exert growth proliferation and differentiation, whereas ERβ is known to exert anti-proliferative effects [95]. The structure of ER is composed of mainly three functional domains, having specific roles and mechanisms of action, N terminal domain (NTD), DNA binding domain (DBD), and ligand binding domain (LBD). NTD and LBD comprise the Activation Function domain AF1 and AF2, respectively (Fig. 2.7). Both AF1 and AF2 are essential, independently, or synergistically, for regulating the comprehensive transcriptional activity of ER. AF2 is ligand/hormone-dependent, whereas AF1 is independent [94]. To date, 266 crystal structures of ERα are available, among which 252 comprise the ligand binding domain (LBD).

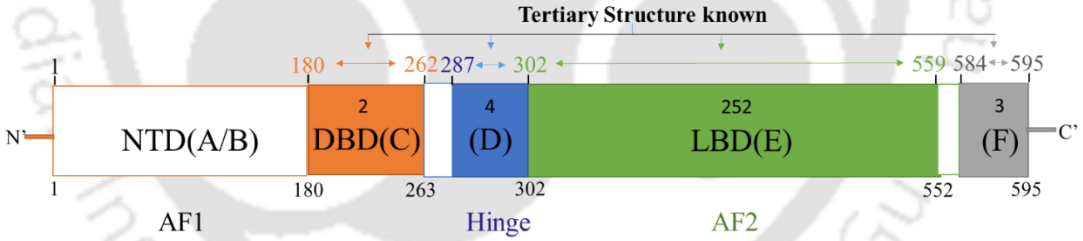


Fig. 2.7: Schematic of PDB structures available for ERα (white region denotes structure unavailable). Orange color denotes DBD, blue-hinge domain, green-LDB, and grey-F domain, respectively.

The ligand binding domain encompasses 12 alpha helices and is responsible for ligand binding. Its globular structure harbors hormone binding sites, homo, or hetero dimerization interfaces, and co-regulatory (activator/repressor) interaction sites [96]. Interestingly, despite low sequence homology with NHR family LBD, the 3D structure organization is strikingly similar. Ligand binding pocket (LBP) is flexible and hydrophobic in nature. Upon Ligand binding, LBP reshapes around the ligand and stabilizes the complex by complementary hydrophobic interactions and hydrogen bonds [97]. The first crystal structure bound to E2 is shown to have an ellipsoidal cavity where E2 is buried in a highly hydrophobic environment [98]. This pocket comprises

of 22 residues. The Hydroxyl group of A and D rings of E2 are H bonded, with Glu 353, Arg 394, and His 524 [99]. The study of ligand-bound crystal structure plays an essential role in understanding the mechanism of interaction through different structure and conformation changes induced by ligands leading to varied cellular responses [90,91].

Estrogen, a steroid hormone, diffuses across the membrane and interacts with its nuclear receptors (ERα and ERβ), causing these receptors to alter conformation, dimerize, and translocate into the nucleus [100]. These ligand-activated receptors function as transcription factors, controlling the expression of numerous genes by binding to the Estrogen Response Element (ERE) or tethering to other transcriptional factors bound to their specific recognition sequences [101] (Fig. 2.8).

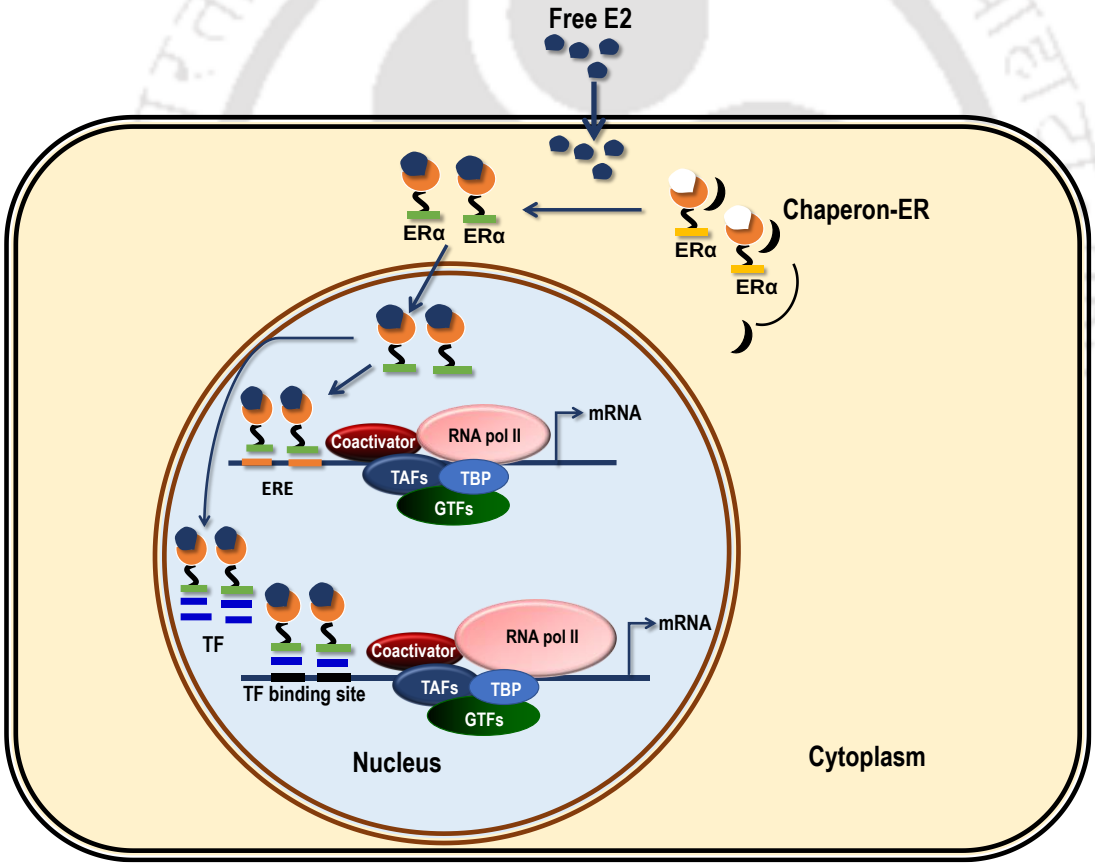


Fig. 2.8: The mechanism of estrogen action and estrogen receptor signaling. E2 diffuses through the plasma membrane binds and activates cognate Estrogen receptors, which translocate into the nucleus. In classical E2 signaling, the E2-ERα complex binds in the ERE region or other transcription factors binding sites such as AP-1 and SP1 and recruits various coregulators, leading to gene expression and cellular response.

2.7 AhR and AhR signaling

The aryl hydrocarbon receptor (AhR), encoded by the *AHR* gene, is a ligand-dependent transcription factor that acts as a dioxin or xenobiotic receptor. AhR belongs to the basic helix-loop-helix family of transcription factors that contains a ligand-binding domain (Fig. 2.9) [102]. Upon activation, AhR translocates to the nucleus, where it mediates the transcription of downstream target genes with AhR-responsive elements (also called dioxin or xenobiotic response elements) in a cell type- and context-dependent manner [103]. AhR is well known to bind to a number of exogenous and endogenous ligands, which can be divided into the following three categories: agonists (xenobiotics, dietary compounds, endogenous compounds, and microflora); antagonists (xenobiotics and dietary) and selective AhR modulators (xenobiotics). Unliganded AhR in the cytoplasm forms a complex with a dimer of HSP90 and the co-chaperone protein X-associated protein. AhR can undergo nucleocytoplasmic shuttling owing to the presence of a nuclear localization and a nuclear export signal. The AhR complex translocases into the nucleus after binding an agonist, and ARNT facilitates HSP90 displacement, creating an AhR/ARNT heterodimer. This heterodimer can then bind to dioxin response elements of target genes such as *CYP1A1* and *CYP1B1* [104].

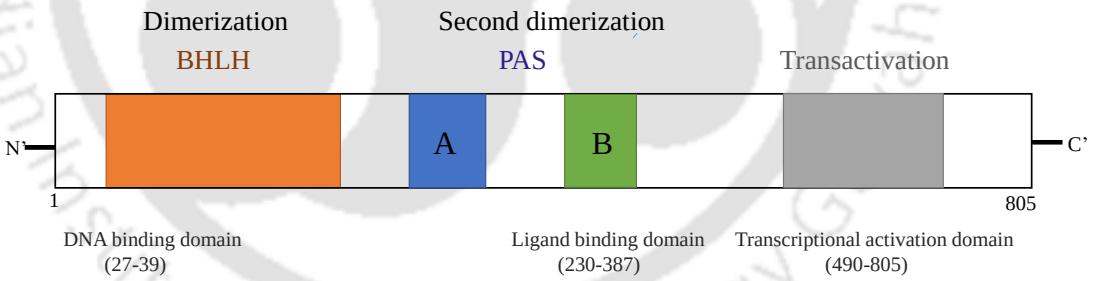


Fig. 2.9: Schematic of AhR structure. AhR is an 805 amino acid (aa) long transcription factor. The orange color denotes the basic helix loop helix (BHLH) domain, which encompasses the DNA binding domain (23-39 aa). This BHLH domain allows the dimerization, binding of DNA and interactions with chaperones such as Hsp90 (Heat Shock Protein 90). Blue-PAS A is involved in dimerization, green-PAS B comprises ligand binding domain (230-387 aa), and the grey-transcription activation domain (490-805 aa) allows coactivator binding, respectively.

AhR serves as an environment sensor and either detoxifies harmful chemicals or activates prodrugs to an active form [22]. AhR activation by xenobiotics such as polycyclic aromatic hydrocarbons (PAHs) and dioxins can lead to cell proliferation and cancer. In contrast, the activation of AhR by natural compounds such as indoles and

flavonoids can have anti-inflammatory and anti-tumor effects [105,106]. In addition to its role in response to environmental pollutants and natural compounds, the AhR also plays a vital role in the immune response, metabolism, and energy balance [102].

The effects of AhR activation can be positive and negative, depending on the tissue and the type of ligand. For example, in immune cells, AhR activation can lead to the differentiation of T-helper 17 cells, which can have anti-inflammatory effects, while in breast cancer cells, AhR activation can promote cell proliferation and increase resistance to chemotherapy. In addition, AhR can also interact with other signaling pathways, such as the PI3K/Akt pathway and the Wnt signaling pathway, which can modulate its activity and downstream effects [105]. AhR has also been implicated in the development and progression of breast cancer. AhR activation by environmental pollutants such as polycyclic aromatic hydrocarbons (PAHs) and dioxins has been shown to promote breast cancer cell growth and increase breast cancer incidence in animal models. Additionally, AhR activation has been linked to the development of resistance to anticancer therapies such as tamoxifen. It has also been found that AhR expression is upregulated in breast cancer cells and that its activation correlates with poor prognosis and increased risk of recurrence in breast cancer patients. Furthermore, studies have shown that inhibiting AhR activity can inhibit the growth of breast cancer cells and sensitize them to chemotherapy.

On the other hand, some natural compounds such as indoles and flavonoids have been found to have anti-inflammatory and anti-tumor effects on breast cancer cells by activating AhR [105]. In summary, AhR is a potential therapeutic target in breast cancer, and ongoing research is needed to understand the mechanism of action and the most effective ways to target it [106–108].

AhR is a ligand-activated transcription factor whose activity can be modulated by polyphenols [16]. Transcriptional activation of the human CYP1A1 and CYP1B1 gene is mediated via AhR (Fig. 2.10). Activation of AhR is implicated in many critical stages of tumorigenesis, such as initiation, promotion, progression, and metastasis. However, whether it acts as a pro-tumorigenic or anti-tumorigenic factor remains debatable [104]. The target genes of AhR, namely CYP1A1 and CYP1B1, are implicated in cancer. CYP1A1 is involved in the conversion of toxicants such as benzo(α)pyrene (B(a)P) into genotoxic agents and regulation of breast cancer

proliferation and survival [109]. CYP1B1 is considered a "universal tumor marker" since it is overexpressed in several types of cancer, including renal, breast, and colon cancers [110]. CYP1B1 is involved in tumor progression due to its ability to convert estrogen into genotoxic metabolites such as 4-hydroxy estradiol [111].

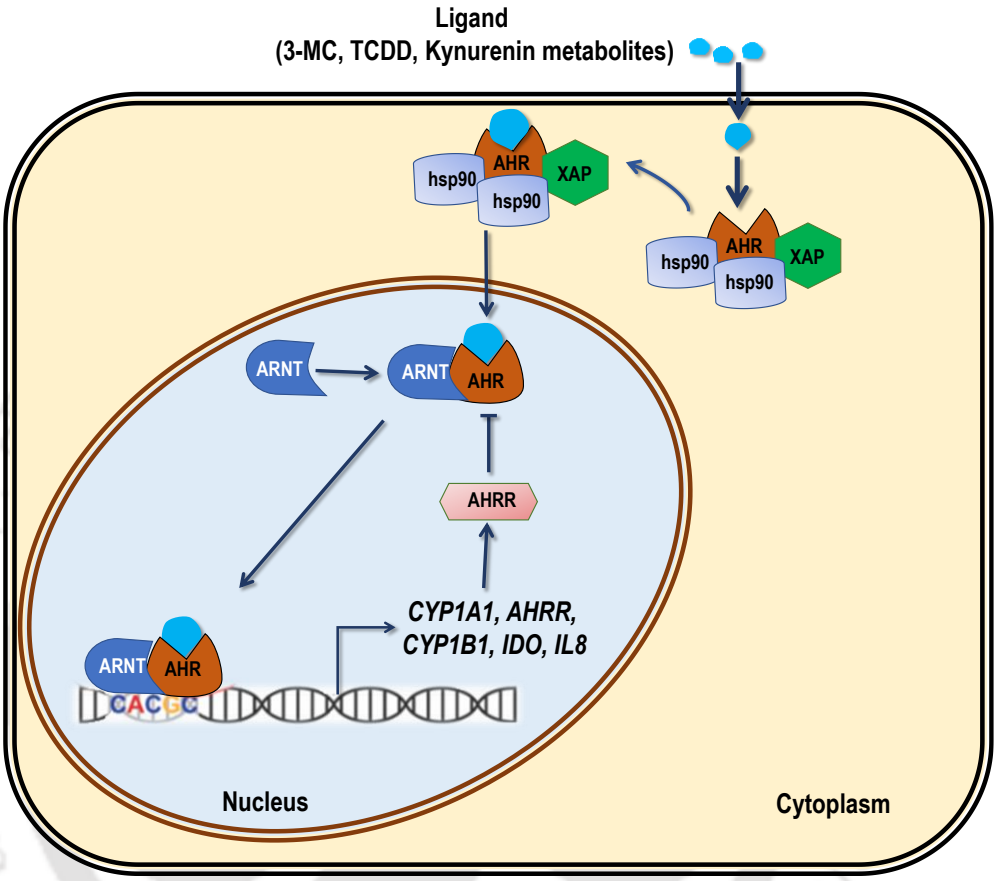


Fig. 2.10: The mechanism of AhR signaling. Upon activation, AhR translocates to the nucleus, forms a heterodimer with its partner protein, ARNT, and binds to specific DNA sequences called xenobiotic response elements (XREs), leading to the transcriptional regulation of various genes involved in cellular processes such as metabolism, immune response, and development.

2.8 Breast cancer

Breast cancer is a malignant neoplasm that starts in the breasts and spreads via the lobules (the glands that produce milk) and the connecting ducts. Changes in the size or form of the breast, as well as changes to the skin of the breast, such as dimpling or redness, may be symptoms of breast cancer [87,112]. Breast soreness, drainage from the nipple, and enlarged lymph nodes in the armpit are additional symptoms. Breast cancer is the second most frequent cancer overall and the most prevalent cancer in women. According to the International Agency for Cancer Research (IARC) estimates

[113] breast cancer accounts for the highest number of cancer-related deaths in women worldwide [114]. In India, it is the cause of 27% of all new cancers, with comparatively high mortality due to factors such as lack of awareness, delayed diagnosis, and socioeconomic status [115]. Breast cancer treatment strategies include surgical removal of the breast (lumpectomy or mastectomy), radiation therapy, chemotherapy, hormone therapy, and targeted therapy.

Endocrine therapy, also known as hormone therapy, targets estrogen and progesterone and their receptors, which can fuel the growth of hormone receptor-positive breast cancer. Endocrine treatment works by blocking the production or action of estrogen and progesterone by altering the estrogen signaling pathway, which can delay or inhibit the growth of the cancer cells [116,117]. Several types of endocrine therapy can be used to treat breast cancer, including Aromatase inhibitors (AIs) such as anastrozole or letrozole, Selective estrogen receptor modulators (SERMs) such as tamoxifen or raloxifene, ER degraders such as Fulvestrant, and Luteinizing hormone-releasing hormone (LHRH) agonists such as Leuprolide [118,119]. Endocrine therapy may be used alone or in combination with other treatments, such as surgery or radiation therapy. Increased interest in endocrine therapy has been shown because 60-75% of breast cancers are ER or PR positive, making it a desirable therapeutic choice for patients[120–122]. For aggressive triple-negative breast cancer (TNBC), AhR has emerged as a viable biological target for the therapy of this fatal illness. AhR agonists such as amino flavone have progressed to phase I clinical trials [123].

2.8 Selective estrogen receptor modulators

Selective estrogen receptor modulators (SERMs) are drugs that bind to the ER and can either stimulate or inhibit the activity of the receptor and thus behave as an agonist or antagonist, respectively. They possess cell/tissue- and ligand-specific effects due to variations in coregulator recruitment, ligand-induced receptor conformations, and epigenetic changes [124,125], which allows them to exert distinct effects on varied tissues or organ type. For example, raloxifene is a SERM used to prevent and treat osteoporosis in postmenopausal women by mimicking the effects of estrogen on bone tissue. In contrast, tamoxifen is used to treat breast cancer by blocking the effects of estrogen on breast tissue. Other examples of SERMs include toremifene and lasofoxifene (non-steroidal SERM) [19]. Due to the structural plasticity of ER LBD, a

diverse set of ligands can modulate ER and elicit SERM-like behavior [96]. Certain flavonoids possess a structural resemblance to estrogen, allowing them to interact and bind to estrogen receptors, hence garnering the term natural SERMs [17]. Natural SERMs have been studied for their potential use in treating menopausal symptoms and reducing the risk of certain cancers [126].

2.9 Selective aryl hydrocarbon receptor modulators

AhR ligands that exhibit diverse gene, promoter, tissue, or cell type-specific effects at the genomic or functional level are term Selective AhR modulators (SAhRMs) [108]. They elicit a spectrum of effects depending on the environment and setting in which they are applied. Some AhR modulators behave as an agonist in one condition while antagonists in another. SAhRMs are broadly classified into endogenous, chemical compounds, or natural compounds. Tryptophan derivates, including Kynurenic acid, tryptamine, and serotonin, were important endogenous ligands that behave SAhRMs [125]. Chemical halogenated toxic compounds such as 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and substituted di-diindolylmethane (DIMs) have shown potential in breast cancer treatment and thus termed anticancer SAhRMs [127]. Natural AhR modulators are essentially plant-derived compounds with AhR modulatory functions. Flavonoids belong to the largest class of natural SAhRMs. Due to the structural plasticity of AhR, these SAhRMs can bind and modulate the downstream genomic and cellular effects. Some AhR modulators have toxic effects, while others exhibit beneficial effects such as anti-inflammatory or cancer-preventive nature. Thus, these compounds are embarked as vital therapeutics leads. Predicting the nature and activity of SAhRMs is difficult, and necessitates rigorous testing and thorough analysis [127,128]. The use of genomic analysis, namely through the identification of disparities in gene regulation, facilitates the ability to predict the characteristics of SAhRMs [129].

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

CHAPTER 3

Materials and Methods

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

3.1 Chemicals, reagents, and plasticware

Karanjin was isolated from *P. pinnata* seed oil or standard was procured from Yucca Enterprises (CAS no. 521-88-0, Batch No. Yucca/KG/2019/04/21, Mumbai, India), and validated by NMR, HRMS, and HPLC techniques as described in Appendix 1 and previously [5,130]. 17β-estradiol (E2, Cat. No. E8875) and AhR antagonist (CH223191, Cat. No. C8124) were purchased from Sigma-Aldrich (MO, USA). Dulbecco's Modified Eagle Medium (DMEM) with (Cat. No. AT007) or without phenol red (Cat. No. AT187), Dulbecco's Phosphate Buffered Saline (DPBS, Cat. No. TS1006), trypsin-EDTA (Cat. No. TCL014), antibiotic solution (Cat. No. A001), fetal bovine serum (FBS, Cat. No. RM10432), and charcoal-stripped FBS (cs-FBS, Cat. No. RM10416), were purchased from HiMedia (Mumbai, India). PowerUp SYBR Green PCR Master Mix (Cat. No. A25743), High-Capacity cDNA Reverse Transcription Kit (Cat. No. 4368814), ERα siRNA (Cat No.4392420), scrambled siRNA (Cat. No. AM4611), and Lipofectamine RNAiMAX (Cat. No. 13778-075) were from Invitrogen (CA, USA). AhR antibody (Cat. No. D5S6H**),** ERα antibody (Cat. No. 8644S), β-actin (Cat. No. 4970T), and anti-rabbit immunoglobin G (Cat. No. 7074S) were from Cell Signaling Technology (Massachusetts, USA). Polyclonal histone H3 antibody (Cat. No. BB-AB0055) was purchased from BioBharati LifeScience Pvt. Ltd. (Kolkata, India). All other chemicals and buffers were purchased from Merck (Mumbai, India), Sigma (St Louis, MO, USA), or SRL (Mumbai, India). All cell culture plasticware was purchased from Eppendorf (Hamburg, Germany).

3.2 Cell culture

MCF-7, MDA-MB-231, and T47D breast cancer cells were obtained from the National Centre for Cell Science (NCCS) (Pune, India). MCF-7 or MDA-MB-231 and T47D cells were routinely cultured and maintained under standard conditions of 37°C and 5% CO_2 in phenol red-containing DMEM or RPMI-1640, respectively, which were supplemented with 10% heat-inactivated (fetal bovine serum) FBS, 100 units/mL penicillin, and 100 µg/mL streptomycin (M1 medium).

3.3 Treatment

Cells were cultured in M1 medium till 70–80% confluence and then shifted to phenol red-free DMEM supplemented with 10% heat-inactivated charcoal-stripped

fetal bovine serum (cs-FBS,) 100 units/mL penicillin and 100 µg/mL streptomycin (M2 medium) for 24 h. Then, the cells were treated with vehicle (0.1% DMSO), 10 nM E2, or indicated concentrations of karanjin in the M2 medium for the indicated period. The treatment medium was replenished every 48 hours in experiments with extended treatment periods. For the AhR antagonist (CH223191) experiment, cells were treated with vehicle (0.1% DMSO) or 10 µM karanjin in M2 medium either in the absence or presence of CH223191 for 24 h. Following the completion of the experiment, the cells were processed for total RNA isolation.

3.4 Cell viability assays

MCF-7, MDA-MB-231, and T47D breast cancer cells were seeded with a density of 40,000, 20,000, and 60,000, respectively, in 35-mm dishes with M1 medium. After 48 h, cells were washed twice with DPBS and incubated in M2 medium for 3 h. The cells were treated with vehicle (0.1% DMSO), 10 nM E2, or the indicated concentrations of karanjin in M2 medium for 0, 24, or 120 h. Thereafter, the cells were washed with DPBS, trypsinized, and viable cells were counted based on trypan blue dye exclusion [131]. 10 nM E2 treatment was used as a standard reference. The 0 h treatment group provided the starting viable counts. The treatment medium was replenished every 48 h. Post-completion of the experiment, cells were trypsinized and collected in 1 mL medium. The homogenous cell suspension was mixed with 0.4% trypan blue due (Sigma-Aldrich, St. Louis, USA, Cat No. T8154) in equal ratio, and viable cell count were obtained using hemocytometer.

3.5 Total RNA isolation and cDNA synthesis

Total RNA was isolated using an RNA extraction reagent made in-house that was identical to TRIzol. RNA integrity was determined using agarose gel electrophoresis and quantified using a BioSpectrometer® (Eppendorf, Hamburg, Germany). 2 µg of total RNA was reverse transcribed in accordance with the manufacturer's instructions using the High-Capacity cDNA Reverse Transcription kit.

3.6 Transcriptome profiling

The quality of total RNA isolated from control and treated cells was assessed using Bioanalyzer 2100 (Agilent Technologies, CA, US). RNA samples that were used

for library preparation using TruSeq RNA Kit (Illumina, USA) for MCF-7 or KAPA RNA Hyper prep kit (Roche, Switzerland) for T47D samples had a concentration of 1 µg/ml and RIN values above 9.5. cDNAs reverse transcribed from mRNA were amplified and purified using AMPureXP beads (Beckman Coulter, USA) or KAPA Pure Beads (Roche, Switzerland). The double-stranded cDNAs were end-repaired, polyadenylated, and ligated with adapter sequences followed by size selection for approximately 250-450 bp using AMPure XP beads or KAPA Pure Beads for MCF-7 or T47D samples, respectively. Uracil-containing strands were degraded by treatment with USER Enzyme (New England Biolabs, USA). Sequencing was performed using the Illumina Novaseq 6000 platform or the Nextseq2K platform in paired-end format for the MCF-7 and T47D samples, respectively. Approximately 40 million reads per sample were sequenced. The FASTQ files obtained after sequencing were subjected to QC using the FASTQC tool [132]. Trimming to remove the adapter sequences and short reads was performed using Trimmomatic [133]. The trimmed reads were aligned using the STAR aligner [134]. BAM files generated were used to obtain the read counts using the featureCounts tool [135]. The read count data from all samples were processed further using the DESeq2 package in R [136]. The quality of the normalized count data was assessed and visualized using unsupervised clustering analysis, principal component analysis (PCA) plots, and correlation heatmaps. Following the quality check, the normalized counts were fitted on the negative binomial model. After the estimation of dispersion values, the normalized counts were subjected to statistical analyses to obtain differentially expressed genes by applying the Wald statistic with α set to 0.05, followed by FDR correction with a 5% cut-off. RNA-seq data generated in the MCF-7 and T47D study have been submitted to NCBI Gene Expression Omnibus (GEO) with accession number GSE183913 & GSE233159, respectively.

3.7 Quality assessment of RNA-seq data

Total RNA isolated from MCF-7 and T47D cells treated with DMSO (control) or 10 µM karanjin were subjected to next-generation sequencing (RNA-seq) to identify karanjin-modulated genes. The RIN values of RNA samples were greater than 9.5 (Appendix 2). Sequencing libraries prepared using the RNA samples were of acceptable quality, all having Q30 values greater than 95 (Appendix 2). Paired-end reads were generated, and the quality of reads before and after trimming was assessed using

FastQC. Less than 1% of data was lost due to trimming (Appendix 2). The sequences were aligned with the human genome build GRCh38 using STAR, and feature read count files (Appendix 2) were analyzed using DESeq2. In MCF-7 RNA seq data, unsupervised clustering analysis of read count data from three controls (G1, G2, and G3) and three karanjin-treated (K1, K2, and K3) samples revealed that K1 was an outlier. K1 was found to cluster with the control samples (Appendix 2). Hence, K1 was omitted from differential gene expression analysis.

3.8 Gene set enrichment analysis

The hallmark gene sets enriched upon karanjin simulation were identified by employing fGSEA package in R with FDR correction of 25% [137]. Enrichment plots, and normalized enrichment scores (NES) plots were generated using additional R packages.

3.9 Identification of genes commonly regulated by estrogen, tamoxifen, and karanjin.

Curated gene expression dataset GSE117942, corresponding to 1 nM E2 or 1 μM tamoxifen treatment in MCF-7 cells for 24 hours, was obtained from GEO [138]. The read count data were analyzed using the DESeq2 package [136] in R to identify genes regulated by 1 nM estrogen or 1 μM tamoxifen. The differentially modulated genes were identified by applying the Wald statistic with α equal to 0.05, followed by FDR correction with a 5% cut-off. They were then compared to the karanjin-regulated genes (this study) to generate sets of overlapping genes depicted in the form of a Venn diagram.

3.10 qRT-PCR

cDNA (1:5 dilution) was utilized as a template for PCR reactions with gene-specific primers (Appendix 3). Cyclophilin A or RPL35A served as internal control. Agilent AriaMx Real-time PCR System was used for the real-time PCRs (Agilent, CA, US). For relative gene expression analysis, the comparative ΔΔCt method [139] was utilized.

3.11 Western blotting

Total protein was isolated from the phenolic fraction of the RNA extraction reagent. Alternatively, total protein was extracted using the 1.5X Laemmli sample buffer. Protein was quantified by the Lowry [140] or TCA [141] method, respectively. Protein was fractionated on 10% SDS-PAGE gel and transferred to a nitrocellulose membrane. The blots were probed with ERα, AhR, β-actin, or histone (H3)-specific antibodies followed by HRP-conjugated anti-rabbit secondary antibody. Chemiluminescence signals obtained with Clarity Western ECL Substrate (Bio-Rad, California, US, Cat. No. 170-5060) were captured with ChemiDoc XRS+ system (Bio-Rad, California, US). Histone H3 or β-actin served as an internal control.

3.12 siRNA transfection

MCF-7 and T47D cells were seeded in six-well plates. Cells were transfected with siRNA for 24 h using Lipofectamine RNAiMAX transfection reagent according to the manufacturer's instructions. Each well received 25 pmol of scrambled (control), ERα, or AhR-specific siRNA.

3.13 Chromatin immunoprecipitation (ChIP) Assay

MCF-7 cells were treated with vehicle (DMSO), 10 nM E2, and 10 µM karanjin for 24 h and crosslinked with 0.75% (v/v) formaldehyde for 10 min at room temperature. The crosslinking reaction was stopped by adding 125 mM glycine. Cells were washed thrice with 5 ml ice-cold DPBS, and lysed with 1ml of ChiP lysis buffer (50 mM HEPES pH 7.5, 140 mM NaCl, 1 mM EDTA pH 8.0, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, and 1×Protease inhibitor cocktail). Lysates were then sonicated to shear the DNA at an amplitude of 30% for 35 cycles, each with a 10-sec pulse on and a 25-sec pulse off. The sheared chromatin was centrifuged at 14000 rpm for 10 min, and the supernatants containing chromatin were collected or stored at -80°C until use. 80 µg of sonicated chromatin in ChIP lysis buffer was precleared by incubation with Protein G Plus-Agarose beads, which were precoated with bovine serum albumin (BSA) and herring sperm DNA. The precoated beads were pelleted, and the precleared chromatin in the supernatants was collected. 5% of the chromatin samples were kept aside as input, and the remaining was incubated either with ERα antibody (Cell Signalling Technology, Massachusetts, USA) or normal rabbit IgG

antibody (BioBharati Life Science, Kolkata, India) 4°C for 4 h. 20 µl of precoated protein G Plus-Agarose beads were added to each sample and incubated at 4°C for 2 h. The beads were then pelleted by centrifugation and washed extensively with a series of wash buffers, as described earlier [142,143]. Immunoprecipitated chromatin was eluted by incubating the beads in 100 µl elution buffer (50 mM Tris-HCl pH 8.0, 1% SDS, 10 mM EDTA) at 65 °C for 30 min. The eluted immunoprecipitated chromatin was then centrifuged at 14000 rpm for 2 min at 4°C, and the supernatant was transferred to a fresh tube. 100 µl elution buffer was then again added to the beads, and the elution step was repeated. After that, the elutes were pooled and reverse crosslinked by heating overnight at 65°C. To the eluted chromatin, an equal volume of Tris EDTA (TE) buffer and Proteinase K (100 g/ml) was added and incubated for 3 hours at 37°C. Eluted chromatin was column purified using Nucleospin Gel and PCR clean-up kit from Machery-Nagel (Duren, Germany), and used in PCR reactions with primers to amplify specific regions of interest in the *PS2*, *CSTA HOXB2, TIPARP*, and *CYP1A1* locus. ERα or AhR immunoprecipitated chromatin was eluted, column purified, and used in PCR reactions with primers (Appendix 3, Table 3A2) to amplify specific regions of interest in the *PS2, CSTA*, *HOXB2, TIPARP,* and *CYP1A1* locus [143,144]. E2 was employed as a positive control to validate the experimental protocol [145]. Similarly, amplification of specific region of interest (XRE-Xenobiotic Response Element) in the upstream of *CYP1A1* locus was conducted to check AhR occupancy.

3.14 Docking study

The three-dimensional crystal structure of the ERα ligand binding domain was retrieved from the Protein Data Bank. The crystal structure of a fragment of AhR in complex with ARNT has been solved (PDB code: 5nj8). However, the fragment does not include the ligand binding domain. Hence, an attempt was made to model the residues 34-400 of human AhR using homology modeling techniques. AhR amino acid sequence was obtained from UniProt, and a 3D structure was obtained using HIF-2α (PDB: 4ZPK) as a template due to high sequence identity and similarity. Swiss model was used for homology modeling, and structure validation was performed using the server [146] SAVESv6.0 - Structure Validation Server (ucla.edu). Ligands karanjin, E2, and TCDD were retrieved in Standard Delay Format (SDF) from PubChem and were converted to PDB format using the online SMILES translator or Open Babel

software. Docking studies were carried out using AutoDock 4.2.6 [147]. For ERα and AhR protein, water molecules were removed, and polar hydrogen atoms were added. Gastieger and Kollman charges were added. All default torsions of ligands E2 and karanjin were allowed to rotate during the docking run. A grid box 40 x 40 x 40 was prepared, encompassing the ligand binding pocket of ERα or AhR. A grid spacing of 0.375 Å was used to compute affinity maps and electrostatic maps. A docking search was carried out using the Lamarckian genetic algorithm and RMSD (1 Angstrom). All remaining parameters were set as default. The docked conformations were sorted according to predicted binding energy, with the lowest energy conformation considered to be the most reliable one.

A schematic of the docking protocol is shown in Figure 3.1. 1GWR is representative of Eight wildtype E2-bound structures that were used for the study. The wild-type E2-bound crystal structures were taken as receptors, and the endogenous E2 from each structure was re-docked to get the same binding conformation. Docked conformations were visualized in PyMOL, LigPlot, or Discovery Studios software.

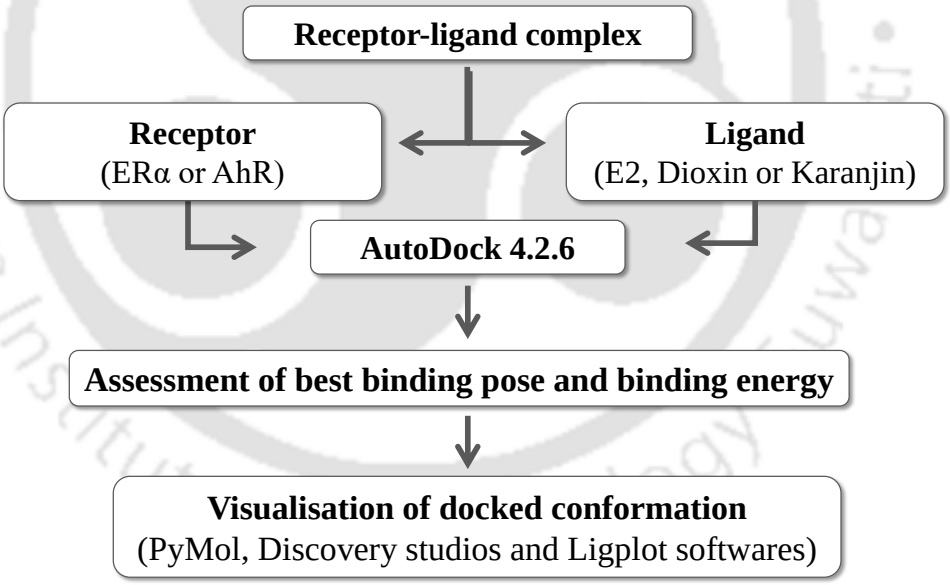


Fig. 3.1: Schematic of docking procedure. Human ERα LBD crystal structure retrieved from PDB was docked against E2 and karanjin using AutoDock 4.2.6. Assessment of binding, and visualization was carried out using PYMOL and LIGPLUS.

3.15 Molecular dynamics simulation

Models of AhR-karanjin and AhR-TCDD were obtained from AutoDock and all-atom classical Molecular Dynamics (MD) simulations were performed using GROMACS 2020.1 program. MD was run on in-house supercomputing facility Param-

Ishan at IIT Guwahati. Protein, water, and ions were described using the standard CHARMM36 force field [148]. The simulation box was slowly heated up (up to 310K) in the first phase of equilibration and then kept fixed during the production dynamics. This system was subjected to energy minimization for 50000 steps by using the steepest descent algorithm. The minimized system was equilibrated for a short MD run (100 ps) with position restraints until the potential energy converged and stabilized at 300 K. The TIP3P water model [149] was used to represent explicit waters. Berendsen temperature coupling was turned on with temperature set to 300 K[150]. The leap-frog algorithm was used for integrating Newton's equation in MD simulation. LINCS constraints were used on protein covalent bonds to maintain constant bond lengths [151]. Finally, the equilibrated systems were subjected to MD simulations for 100 ns at 300 K, and all atoms were permitted to move freely. In total, 100 ns of production MD simulations were considered for structural analysis of AhR-bound structures. A root-mean-square deviation (RMSD) and Root-mean-square fluctuation (RMSF) of the protein-heavy atoms relative to the homology-modeled structure were estimated (Appendix 4).

3.16 Statistical analysis

Two-group data were analyzed by one-tailed t-tests. Multiple group data were analyzed by one-way ANOVA followed by multiple comparison tests with Tukey HSD. To study the effect of karanjin concentration on viable cells after 24 or 120 h, and the effect of ERα or AhR knockdown, and AhR antagonist on karanjin-modulated gene expression, the data were analyzed by two-way ANOVA, only after confirming homogeneity of variances using the Levene test. For the identification of genes modulated by karanjin using DESeq2, we applied Wald statistic with α equal to 0.05, followed by FDR correction with a 5% cut-off. All statistical analyses were performed using the R statistical package.

CHAPTER 4

Genome-wide transcriptomic effects of karanjin in MCF-7 breast cancer cells

The results presented and discussed in this chapter are published in Bhatt et al., (2022), *Gene.*

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

4.1. Introduction

Karanjin, the most abundant flavonoid present in *P. pinnata* seed oil, harbours a multitude of health benefits and applications. It produces a wide range of biological effects *in vitro* or *in vivo* [1]. The cell cycle inhibitory and pro-apoptotic effects of karanjin *in vitro* have fuelled speculations about its anticancer properties [8–10]. Notably, Roy and co-workers [10] reiterated karanjin-mediated G2/M arrest and apoptosis in HeLa cells. Furthermore, they demonstrated karanjin-induced DNA damage associated with lowered reactive oxygen species (ROS) and restricted nuclear translocation of NF-κB via cytoplasmic I-κB induction [9]. The fact that karanjin exerts a much lower growth inhibitory effect on normal mouse embryonic fibroblasts [9] augurs in favor of its anticancer potential. All cancer cell lines tested so far are growth-inhibited by karanjin with varying inhibitory concentration (IC_{50}) values [1]. However, the impact of karanjin on gene expression and genomic correlates remains unaddressed. Also, the effect of karanjin simulation in hormone-responsive breast cancer cells is lacking.

In this chapter, attempts were made to delineate the transcriptomic footprint of karanjin in MCF-7 breast cancer cells using next-generation sequencing technology (RNA-seq). This study revealed that karanjin regulates a diverse set of genes and pathways. Karanjin modulated G2/M checkpoint genes in a manner consistent with cell cycle progression rather than cell cycle arrest. Besides, modulation of estrogen-response early and other estrogen-regulated genes suggests a partial estrogen-like effect of karanjin on gene expression in ER-positive breast cancer cells. This is the first study to demonstrate the proliferative and phytoestrogen-like potential of karanjin.

4.2 Results

4.2.1 Determining the optimal concentration of karanjin

A study of the genome-wide transcriptomic effects of karanjin necessitated the identification of an optimal concentration, and duration of treatment, which captured direct and early responses free from strong mitotic or toxic effects. The short-term and the long-term effect of varying concentrations of karanjin on MCF-7 cell viability was analyzed. MCF-7 cells were treated with 0 to 50 μM concentrations of karanjin, and viable cells were recorded after 0, 24, or 120 h (Fig. 4.1). We applied two-way ANOVA

to study the difference in viable cells after 0, 24, and 120 h of treatment as a function of karanjin concentration. At 24 h time-point, there was no interaction between concentration and time ($p \approx 1$). The viable cell counts were significantly higher after 24 h of treatment ($p < 0.0001$), and the effect of time was significant. There was no effect of concentration. The viable cell counts after 24 h of treatment with varying concentrations of karanjin (0–50 µM) were not significantly different (Fig. 4.1A). Thus, over a period of 24 h, karanjin had no significant impact on MCF-7 cell viability compared to control, although proliferation was evident from the significant increase in viable cell count at all concentrations. Hence, we chose to treat MCF-7 cells with 10 µM karanjin for 24 h for RNA-seq.

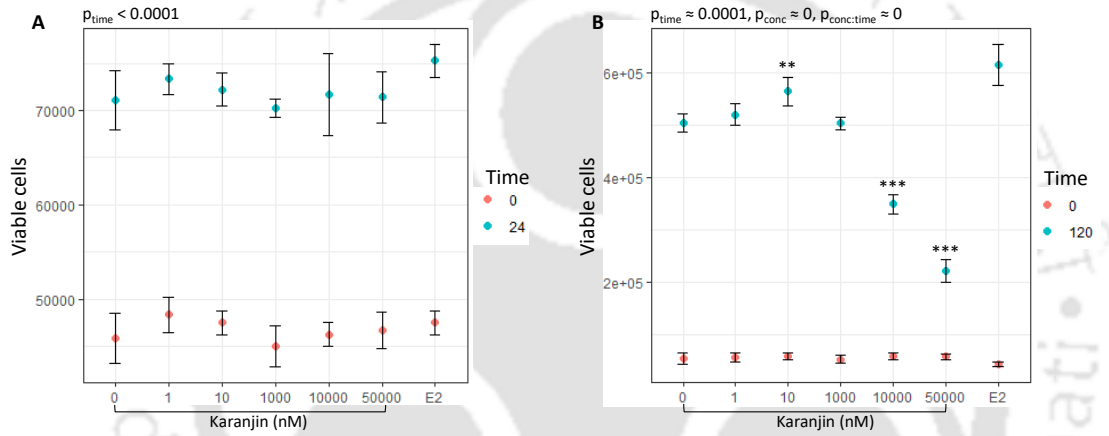


Fig. 4.1: Concentration-dependent effect of short- and long-term stimulation with karanjin on MCF-7 cell viability. 4 x 10^4 MCF-7 cells were seeded in 35 mm dishes and grown in M1 medium for a period of 36 h. The cells were then shifted to the M2 medium and incubated for 3 h. Thereafter, the cells were treated with vehicle or indicated concentrations of karanjin. The viable cells were counted after 24 h (A) and 120 h (B) based on trypan blue dye exclusion using a hemocytometer. Cells treated for 0 h provided the initial viable counts for each treatment group (red dots). Each dot represents the mean viable count ± sd (n = 3). The data were analysed by two-way ANOVA to test whether the increase in the viable count after 24 h was dependent on the concentration of karanjin. $p_{time} < 0.0001$.

The long-term effects of varying concentrations of karanjin on MCF-7 cell viability revealed a significant interaction between time and concentration. Although the viable cell counts increased at all concentrations of karanjin as revealed by significant main effects of time ($p \approx 0$), the number of viable cells after 120 h of treatment depended on the concentration of karanjin (Fig. 4.1B, p-value for interaction term ≈ 0). About 10 nM karanjin treatment yielded significantly higher viable cell count ($p = 0.002$), whereas 10 or 50 µM karanjin yielded significantly lower viable cell counts

with respect to control ($p \approx 0$). Thus, 10 μM karanjin had an intermediate effect on viability over a period of 120 h.

4.2.2 Karanjin-regulated genes

We identified karanjin-modulated genes by interrogating the read count data with DESeq2 package in R. The volcano plot in Figure 4.2A, shows that 736 genes (blue dots) were modulated by karanjin, based on a threshold of 0 for LFC, and $p < 0.05$ for Wald statistic and FDR. About 362 genes were repressed, and 374 genes were induced by karanjin, as illustrated by Figure 4.2B. Among the top 25 induced (Table 4.1) or repressed genes (Table 4.2) were those that encode proteins with diverse functions. These include receptors, signal transducers, enzymes, enzyme inhibitors, transcription factors, and non-coding RNAs. Selected genes from Table 4.1 and Table 4.2 were validated by qRT-PCR. As shown in Figure 4.2C, the significant down-modulation of *B3GALT5* ($n = 3$, $p = 0.009$), *BRINP2* ($n = 4$, $p = 0.002$), *CHST1* ($n = 3$, $p = 0.002$), and *CSTA* ($n = 3$, $p = 0.002$), and significant upregulation of *CYPIA1* ($n = 3$, $p = 0.002$), *CYPIB1* ($n = 4$, $p < 0.0001$), *MRVI1* ($n = 4$, $p < 0.001$), and *CREG2* ($n = 4$, $p < 0.0001$) following 10 μM karanjin treatment is consistent with the DESeq2 results.

Table 4.1: Top 25 up-regulated genes in MCF-7 cells when treated with 10 μM karanjin for 24 h.

Gene Symbol	Name	Nature	Function[@]	FC
CYP1A1	Cytochrome P450 Family 1 Subfamily A Member 1	Enzyme (Aryl Hydrocarbon Hydroxylase)	drug metabolism, steroidogenesis	72.4
ANKRD29	Ankyrin Repeat Domain 29	Peripheral protein	attachment of integral membrane proteins to spectrin-actin based membrane cytoskeleton	8.7
SLC7A5	Solute Carrier Family 7 Member 5	Transport protein (amino acid exchanger)	sodium independent transport of neutral amino acids across membrane	7.4
MRVI1	Inositol 1,4,5-trisphosphate receptor-associated 1	Enzyme substrate	regulates IP3 induced calcium release, involved in myeloid cell growth and differentiation	6.8

CYP1B1	Cytochrome P450 Family 1 Subfamily B Member 1	Enzyme (Aryl Hydrocarbon Hydroxylase)	drug metabolism, steroidogenesis	6.6
CREG2	Cellular Repressor of E1A Stimulated Genes 2	Enzyme (oxidoreductase)	repression of E1A-stimulated genes involved in transcriptional regulation	6.0
ALDH1A3	Aldehyde Dehydrogenase 1 Family Member A3	Enzyme (aldehyde dehydrogenase)	catalyzes the formation of retinoic acid, involved in the embryonal development of nasal and eye region	5.5
SLC7A11	Solute Carrier Family 7 Member 11	Transport protein (cystine/glutamate transporter)	sodium independent transport of anionic amino acids across the membrane	5.4
TENM4	Teneurin Transmembrane Protein 4	Transmembrane protein	establishes neuronal connectivity during development	4.8
AC015712.6	Gastric Cancer Associated WDR5 And KAT2A Binding LncRNA	Long non-coding RNA	gene silencing of transcriptional regulators such as KAT2A, WDR5	4.8
AC015712.2	ALDH1A3 Antisense RNA 1	Single-stranded anti-sense RNA	silencing of ALDH1A3 at the transcriptional level	4.7
ALDH3A1	Aldehyde Dehydrogenase 3 Family Member A1	Enzyme (aldehyde dehydrogenase)	detoxification of alcohol-derived acetaldehyde, confers resistance to UV radiation in cells	4.4
TIPARP	TCDD Inducible Poly (ADP-Ribose) Polymerase	Enzyme (ribosyl transferase)	catalyzes mono-ADP-ribosylation of basic amino acids, inhibits transcription activator activity of AHR	4.2
CENPE	Centromere Protein E	Kinesin-like motor protein	mediates stable spindle microtubule capture at kinetochores, required for	3.0

			chromosomal alignment during prometaphase	
LAMA3	Laminin Subunit Alpha 3	ECM remodeling protein	mediates the formation and function of basement membrane, cell migration, mechanical signal transduction	3.6
CENPF	Centromere Protein F	nuclear-matrix protein	associates with the centromere-kinetochore complex, regulate chromosomal segregation during mitosis.	3.5
VTCN1	V-Set Domain Containing T Cell Activation Inhibitor 1	Cell surface receptor	suppresses tumor-associated antigen-specific T cell immunity, promotes cell cytotoxicity	3.2
WSCD1	WSC Domain Containing 1	Enzyme (sulphotransferase)	involved in tissue development and distribution at different organ sites	3.2
MYO16	Myosin XVI	Actin-based motor protein	intracellular movement of membranous compartments, regulates cell cycle progression	3.2
LINC00930	Long Intergenic Non-Protein Coding RNA 930	Long non-coding RNA	implicated role in the development of Acute Megakaryocytic Leukaemia (AMKL)	3.0
RND1	Rho family GTPase 1	Enzyme (Rho GTPase)	regulates the organization of the actin cytoskeleton in response to extracellular growth factors.	3.0
ASPM	Assembly Factor for Spindle Microtubules	Calmodulin-binding messenger protein	involved in normal mitotic spindle function and co-ordination of mitotic	2.9

			processes, regulates neurogenesis	
AREG	Amphiregulin 2	Growth factor	part of EGFR and TGFβ Signaling pathways, involved in mammary, oocyte, and bone tissue development	2.8
SCARA5	Scavenger Receptor Class A Member 5	Ferretin receptor	mediates non-transferrin-dependent delivery of iron to cells by ferretin endocytosis	2.8
RGS22	Regulator of G Protein Signaling 22	GTPase activating protein	inhibits signal transduction by increasing GTPase activity	2.8

[@] Curated from GeneCards, KEGG, and Uniprot.

Table 4.2: Top 25 down-regulated genes in MCF-7 cells when treated with 10 µM karanjin for 24 h.

Gene Symbol	Name	Nature	Function[@]	FC
BRINP2	Bone Morphogenetic Protein/Retinoic Acid Inducible Neural-Specific 2	Regulatory protein	cell cycle, positive regulation of neuron differentiation, negative regulation of mitotic cell cycle	0.2
B3GALT5-AS1	B3GALT5 Antisense RNA 1	Long non-coding RNA	implicated role in development of Rectum Adenocarcinoma	0.2
B3GALT5	Beta-1,3-galactosyl-transferase 5	Enzyme (Transferase)	glycosphingolipid biosynthesis; protein glycosylation; acetylgalactosaminyltransferase activity	0.3
CHST1	Keratan sulphate 6-sulfotransferase 1	Enzyme (Sulfotransferase)	glycan biosynthesis; sulfotransferase activity and keratan sulfotransferase activity	0.3
CSTA	Cystatin A	Cysteine Protease Inhibitor	epidermal development and maintenance; cell-cell adhesion; keratinocyte differentiation	0.3
ADAM2	ADAM Metallopeptidase Domain 2	Enzyme (Peptidase)/membrane-anchored proteins	cell-cell and cell-matrix interactions, including fertilization, muscle development, and neurogenesis.	0.3

SLIT1	Slit Guidance Ligand 1	Glycosaminoglycan binding protein	calcium ion binding; axon guidance; negative chemotaxis	0.3
GHR	Growth hormone receptor	Cytokine receptor	PI3K-AKT, JAK-STAT signaling pathway, growth hormone synthesis, secretion, and action	0.3
CXCR4	C-X-C chemokine receptor type 4	Chemokine receptor	regulates calcium, chemokine signaling pathway	0.4
RET	Proto-oncogene tyrosine-protein kinase Ret	Enzyme (Tyrosine-protein kinase)	regulates ERK signaling, PI3K signaling, calcium signaling pathway, involved in central carbon metabolism in cancer	0.4
RASSF4	Ras association domain family member 4	KRAS effector protein	Regulates hippo signaling pathway, involved in tumor suppression	0.4
SIM1	SIM BHLH transcription factor 1	Transcription factor	Genetic information processing; DNA-binding transcription factor activity and obsolete signal transducer activity	0.4
NR4A3	Nuclear receptor subfamily 4 group A member 3	Nuclear receptor	Genetic information processing: signaling and cellular processes; Transcriptional mis-regulation in cancer	0.4
NELL2	Neural EGFL like 2	Laminin domain-containing proteins	Signaling and cellular processes; plays a role in neural cell growth and differentiation as well as in oncogenesis	0.4
DIO1	Iodothyronine deiodinase 1	Enzyme (Oxidoreductase)	Thyroid hormone metabolism and signaling pathway; cellular amino acid metabolic process	0.4
CLEC3A	C-type lectin domain family 3 member A	Carbohydrate-binding protein	promotes cell adhesion to laminin-332 and fibronectin	0.4
FUT2	Fucosyltransferase 2	Enzyme (Transferases)	membrane trafficking and genetic information processing	0.4
HOXB2	Homeobox B2	Transcription factor	involved in development and part of developmental regulatory system that provides cells with	0.4

			specific positional identities on the anterior-posterior axis.	
C10orf10	DEPP1 autophagy regulator	Transcription factor activator	critical modulator of FOXO3-induced autophagy via increased cellular ROS.	0.4
PLA2G10	Secretory phospholipase A2	Hydrolases (phospholipase A2)	phospholipid remodelling; maturation and activation of innate immune cells	0.4
KCNJ8	Potassium inwardly rectifying channel subfamily J member 8	Integral membrane protein	role in cGMP-PKG signaling pathway, potassium transport inside the cell	0.4
ADGRA2	Adhesion G protein-coupled receptor A2	Enzyme (Peptidase) and inhibitor; G protein-coupled receptor	metabolism; signaling and cellular processes	0.4
IL1R1	Interleukin 1 receptor type 1	Cytokine receptor	signal transduction; NF-kappa B signaling pathway; Cytokine-cytokine receptor interaction	0.4
SLC35C1	Solute carrier family 35 (GDP-fucose transporter), member C1	Transporters (Solute carrier family)	signaling and cellular processes; N-Glycan biosynthesis	0.4
RGS11	Regulator of G Protein Signaling 11	GTPase-activating protein	involved in cAMP Pathway and G-Alpha I Signaling	0.5

[@] Curated from GeneCards, KEGG, and Uniprot.

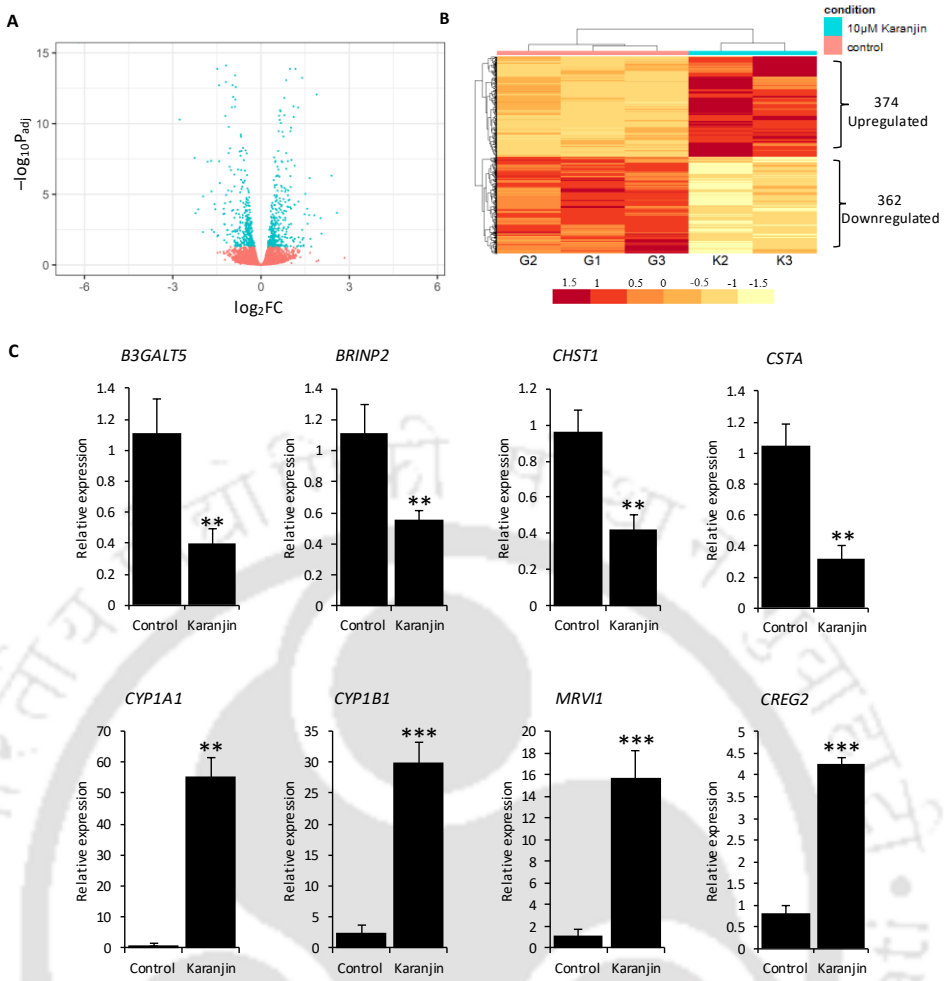


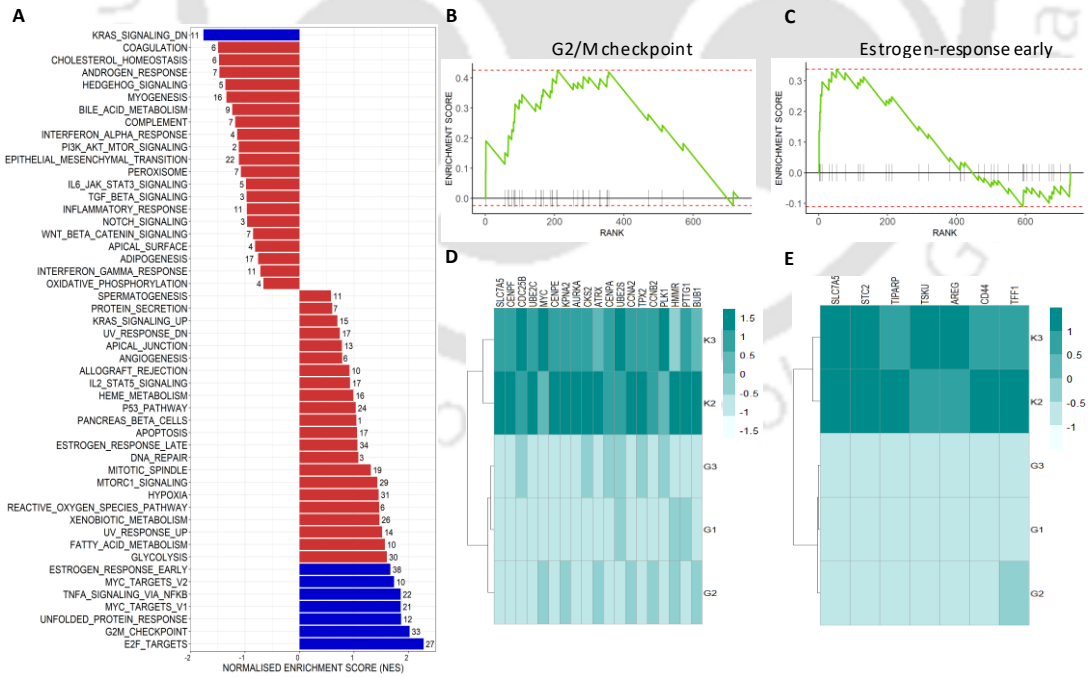
Fig. 4.2: Summary of the RNA-seq data. A. Volcano plot. Each point represents a gene, which is plotted according to its –log10Padj and log_2FC. The blue dots represent genes (n = 736) that are significantly modulated by 10 μM karanjin. The red dots represent non-significant genes. B. Expression heatmap of 736 significantly modulated genes. G1, G2, and G3 are control samples. K2 and K3 are karanjin-treated samples. The color for each gene on the heatmap represents its z-score obtained from the normalized count. C. Validation of RNA-seq data. Total RNA isolated from MCF-7 cells treated with vehicle or karanjin (10 μM) for a period of 24 h was subjected to qRT-PCR analysis of the indicated genes chosen from Appendix 3. The ΔΔCt method was used for analyzing the qRT-PCR data using *Cyclophilin A* as a normalizing control. The experiment had three replicated dishes of control or karanjin-treated cells. The mRNA expression in each sample was expressed relative to the expression of one of the control samples, which was set to one. Bars represent mean relative expression ± sd (n = 3). For each gene, the data were analyzed by one-tailed t-tests. Asterisks represent significant results. (**p < 0.01, ***p < 0.001).

4.2.3 Karanjin modulates G2/M checkpoint, and estrogen-response-early genes

The diversity of genes regulated by karanjin motivated the mining of enriched gene-sets using the fGSEA package. The karanjin-modulated genes were enriched in

several hallmark genesets (Fig. 4.3A), which include G2/M checkpoint and estrogen-response-early genes. The enrichment plots for G2/M checkpoint and estrogen-response-early genes are shown in Figure 4.3B, and 3C, respectively. The leading-edge genes in the G2/M checkpoint set were *SLC7A5*, *CENPF*, *CDC25B*, *UBE2C*, *MYC*, *CENPE*, *KPNA2*, *AURKA*, *CKS2*, *ATRX*, *CENPA*, *UBE2S*, *CCNA2*, *TPX2*, *CCNB2*, *PLK1*, *HMMR*, *PTTG1* and *BUB1*. The leading-edge genes in estrogen-response-early set were *SLC7A5*, *STC2*, *TIPARP*, *TSKU*, *AREG*, *CD44* and *TFF1* (also known as *pS2*). The manner in which karanjin regulated the leading-edge genes within the two-hallmark gene-sets are shown in Figure 4.3D and 3E, respectively.

Modulation of selected leading-edge genes from both hallmark gene-sets was validated by qRT-PCR (Fig. 4.3F). Among the G2/M checkpoint genes, we confirmed the induction of *MYC* (n = 3, p < 0.001), *CDC25B* (n = 3, p = 0.0042), and *CENPF* (n = 3, p = 0.046). Among estrogen-response-early genes, we confirmed the induction of *TFF1* (n = 4, p = 0.019), *CD44* (n = 3, p = 0.003), *STC2* (n = 3, p = 0.017), and *TIPARP* (n = 3, p < 0.001). *SLC7A5*, which belonged to both hallmark gene-sets, was significantly induced (n = 3, p = 0.004).



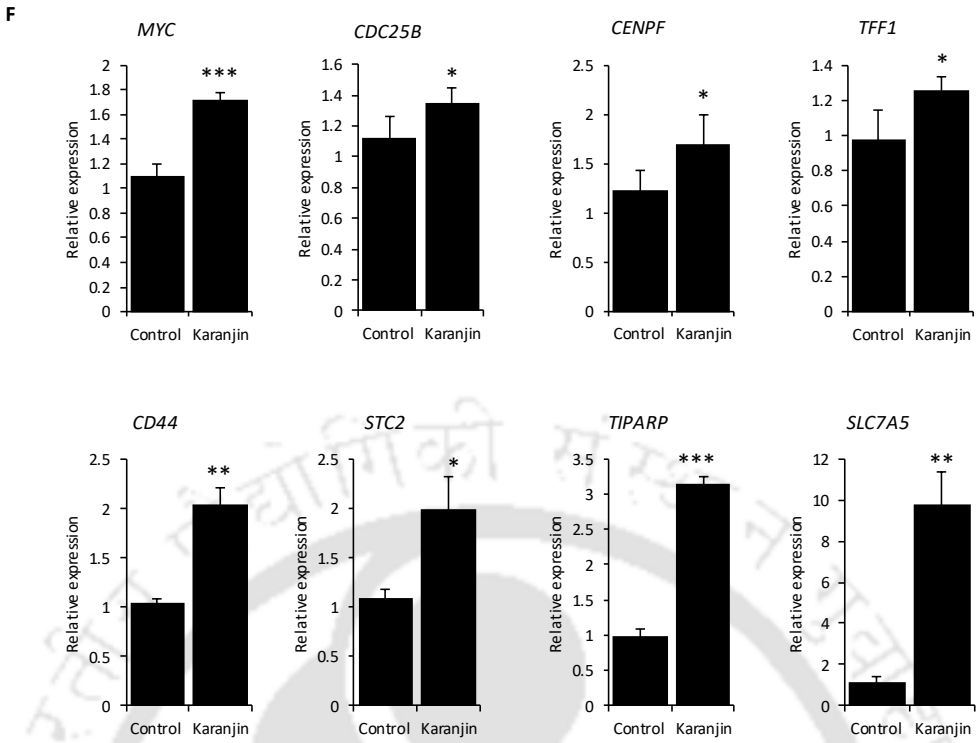


Fig. 4.3: Gene set enrichment analysis. A. Bar chart showing the fGSEA results. The fGSEA package in R was used to analyze the RNA-seq data to identify the enriched hallmark gene sets in MCF-7 cells following 10 μM karanjin treatment. Bars represent NES. Blue bars correspond to the significantly enriched gene sets based on an FDR cut-off 25%. The number of leading-edge genes in each gene set is indicated beside the bars. B, C. Enrichment plots for G2/M checkpoint, and estrogen-response-early genes, respectively. D, E. The pattern of expression of leading-edge genes are shown as heatmaps below the respective enrichment plots for G2/M checkpoint, and estrogen-response-early genes. G1, G2, and G3 are control samples. K2, and K3 are karanjin treated samples. The color on the heatmap represents the expression in terms of z-score obtained from normalized counts for genes (rows) across samples (columns). F. Validation of the karanjin modulated G2/M checkpoint, and estrogen-response-early genes. Total RNA isolated from control and karanjin treated cells were subjected to qRT-PCR analysis of the G2/M checkpoint, and estrogen-response-early genes. The ΔΔCt method was used for analyzing the qRT-PCR data using *Cyclophilin A* as the normalizing control. The experiment had three replicate dishes of control (DMSO) or karanjin (10 μM) treated cells. The expression in one control sample was set to 1, relative to which the expression in other samples were expressed. Bars represent the mean relative expression ± sd (n = 3). For each gene the data were analyzed by a one-tailed *t*-test. Asterisks represent significant results. (*p < 0.05, **p < 0.01, ***p < 0.001).

4.2.4 Karanjin-modulated genes overlap with E2- or tamoxifen-modulated genes

The identification and validation of estrogen-response-early genes instigated us to examine the similarity or differences between the genomic effects of karanjin, E2,

and tamoxifen, which is a SERM. Re-analysis of RNA-seq data (GSE117942) of MCF-7 cells treated with 1 nM E2, or 1 µM tamoxifen [138] yielded two overlapping sets of 3727 estrogen-regulated and 435 tamoxifen-regulated genes. Karanjin-modulated genes were matched with those regulated by estrogen or tamoxifen. Overall, 419 genes were regulated by karanjin and estrogen, whereas 94 genes were regulated by karanjin and tamoxifen. In total, 72 genes were regulated by estrogen, tamoxifen, and karanjin (Fig. 4.4A). Heatmaps in Figure 4.4B and 4C show that karanjin had similar or opposite effects on estrogen or tamoxifen-regulated genes, respectively. For instance, *STC2, TIPARP, SLC7A5,* and *AREG* were upregulated, whereas *CSTA, ADAMTS19, and NCOA3* were downregulated by karanjin and estrogen treatment. *CYP1A1, BMF, and RET* were regulated by karanjin and estrogen in opposite directions. Karanjin also induced similar or opposite effects on genes regulated by tamoxifen. *SLC7A5, FOSL2,* and *PHLDA1* genes were upregulated, whereas *IGFBP4, CHRD,* and *PCDH7* were downregulated by both karanjin and tamoxifen. *GREB1*, *TFF1*, and *CYP1B1* were regulated by both, albeit in opposite directions (Table A1 in Appendix 2). For a more objective assessment of the similarity between the effects of karanjin and 1 nM estrogen, we examined the correlation between the $\log_2$FC values with karanjin and 1 nM estrogen for the subset of 419 genes regulated by both. As shown in Figure 4.4D, the $\log_2$FC values with karanjin correlated significantly with those with 1 nM estrogen (R = 0.42, $p < 0.0001$). Terasaka *et al.*, (2004) have described a set of 154 genes (EstrArray) [153], that can be used to identify estrogenic activity in compounds. Out of these 154 E2-regulated genes (Estr Array), 93 and 28 genes overlapped with 1 nm E2 and 10 µm karanjin gene lists. When we further compared these 93 and 28 genes, and found 25 genes commonly regulated genes by both E2 and Karanjin. Furthermore, the $\log_2$FC values with 10 µm karanjin and 1 nM estrogen for these 25 genes were also significantly correlated (Fig. 4.4E, R = 0.61, p = 0.0012).

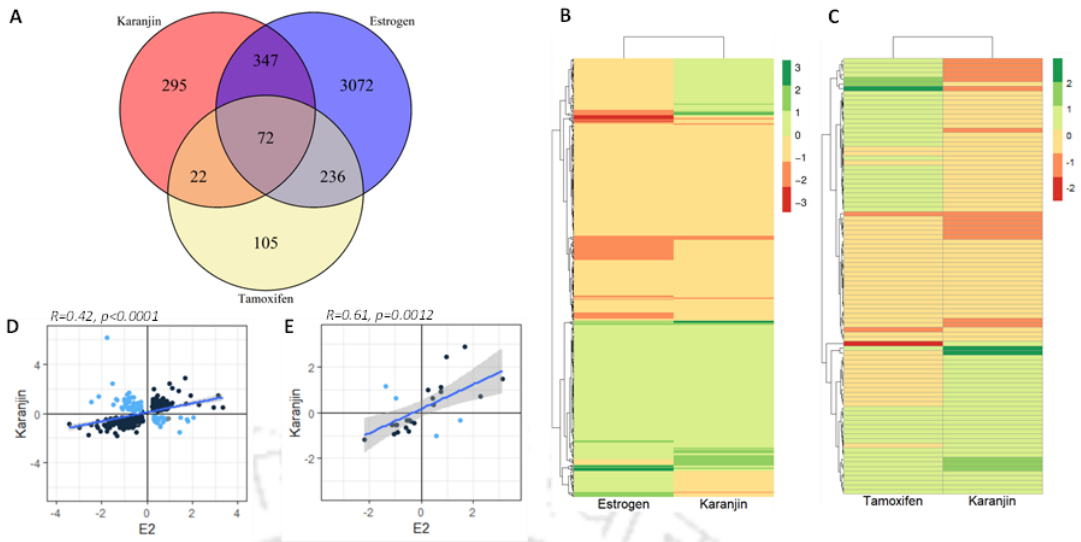


Fig. 4.4: Overlapping genomic targets of karanjin, estrogen and tamoxifen. A. Venn diagram showing the number of genes modulated by karanjin (10 µM), estrogen (1 nM), or tamoxifen (1 µM) in MCF-7 cells. The karanjin-modulated genes were from this study. Estrogen or tamoxifen-modulated genes were obtained by analyzing the GSE117942 dataset as described in Materials and Methods. Values in the areas of intersection are the number of genes commonly regulated by two or more compounds. B. and C. Heatmaps showing the patterns of gene modulation by karanjin compared with estrogen (n = 419), or tamoxifen (n = 94), respectively. Colors represent log_2FC expression compared to their respective controls. For better visualization of expression patterns across tamoxifen and estrogen-modulated genes, *CYP1A1* has been omitted during generation of the heatmap due to its high fold change. D. Correlation between the effects of karanjin and E2 on 419 genes regulated by both. Shown is a scatterplot of log_2FC values with 1 nM E2 and 10 µM karanjin. E. Correlation between the effects of karanjin and E2 on a subset of 25 genes within the EstrArray E2 responsive genes 94 [153]. Shown is a scatterplot of log_2FC values with 1 nM E2 and 10 nM karanjin for 25 genes, which are regulated by both 1 nM E2 and 10 µM karanjin. The dark blue and light blue spots in panels D and E represent genes regulated by both treatments in similar or opposite directions, respectively.

4.3. Discussion

Despite the well-documented *in vitro* and *in vivo* effects of karanjin, the paucity of evidence on its genomic consequences served as the impetus for the current investigation [1]. Here, we have explored the genome-wide alteration of gene expression induced by karanjin at the mRNA level in MCF-7 cells using RNA-seq. The transcriptomic- or gene-modulatory footprint of karanjin encompasses a multitude of cellular processes, such as metabolism, signaling, ROS-scavenging, unfolded protein response, modulation of transcription factor targets, and hormonal response. The 24 h-exposure of MCF-7 cells to 10 µM karanjin did not affect cell viability (Fig. 4.1A). Hence, the observed alteration in the transcriptome may be interpreted as an early and

direct effect that is evidently genomic and unlikely to be caused by a strong mitotic or toxic effect.

The enrichment of G2/M checkpoint genes among those modulated by karanjin is worthy of attention. Although over a period of 24 h, 10 μM karanjin did not significantly induce viable count, it modulated the G2/M checkpoint genes in a manner that is consistent with cell cycle progression, rather than cell cycle arrest. For example, *SLC7A5*, a sodium-independent transporter, is an amino acid exchanger that maintains intracellular levels of leucine. It is an established master regulator of the mTORC1 pathway and is overtly expressed in various cancers [154]. *MYC* is a proto-oncogene that encodes a nuclear phosphoprotein. The regulation of MYC expression in cells is intricately linked with cell proliferation. The elevated expression of MYC promotes the activation of cyclins and CDKs and impairs the functionality of cell cycle inhibitors [155]. CDC25B is a MYC target [156]. It facilitates the entry of cells into mitosis by dephosphorylating cyclin-dependent kinase 2 (CDC2) [157]. Thus, the induction of G2/M checkpoint genes captures the transcriptomic state of cells poised for cell cycle progression. Furthermore, these data contradict other reports of cell cycle arrest and apoptosis [8–11]. It is noteworthy that, these cell cycle inhibitory effects of karanjin on cell viability were shown in different cell lines but not in MCF-7 cells [8–11]. Arguably, the molecular phenotype of cells determines their response to karanjin in terms of alteration in growth, as well as global gene expression. Thus, the induction of G2/M checkpoint genes may be due to the intrinsic properties of MCF-7 cells.

The enrichment of estrogen-response-early genes by karanjin (Fig. 4.3B), the partial overlap of karanjin and 1 nM estrogen-regulated genes (Fig. 4.4B), and the significant correlation between the $\log_2$FC values for the genes commonly regulated by karanjin and 1 nM estrogen (n = 419, Fig. 4.4D) indicate partial estrogen-like effect of karanjin. Some genes, which are not a part of the estrogen-response-early gene signature but are proven to be estrogen-regulated in an ERα-dependent manner, were modulated by karanjin. For instance, *TFF1* [158], *CSTA* [143], and *HOXB2* [144], which are regulated by estrogen via ERα, are modulated by karanjin. However, it does not regulate the expression of PR, which is a classical estrogen-responsive gene. The partial overlap in karanjin and tamoxifen-regulated genes suggests a possible ERα modulatory effect. It remains to be explored whether karanjin directly binds ERα, and modulates its transactivation function in a concentration-, gene-, or cell-type-dependent

manner as exhibited by selective estrogen receptor modulators [18,159]. ERα ligands induce differential conformational changes in ERα protein, thereby impacting coactivator recruitment, ER transactivation [100,101] and target gene modulation [95,160]. This may partly explain the non-uniformity in the direction of gene modulation by karanjin compared to that brought about by estrogen or tamoxifen. Thus, insight into concentration and cell-type dependent effect of karanjin needs to be carried out.

On the premise of these observations, Chapter 5 focuses on to clarify whether the effects produced by karanjin are selective based on the molecular or cellular subtypes. Further studies are needed to investigate the effect of karanjin in other ER+ positive cell lines and to determine if the observed effects are consistent across other cell types.

Indian Institute of Technology Guwahati

CHAPTER 5

Comparative analysis of genomic, concentration- and cell-type dependent effect of karanjin in MCF-7 and T47D cells

Indian Institute of Technology Guwahati

5.1 Introduction

Numerous studies have reported the anti-proliferative, cell-cycle inhibitory, and apoptotic effects of karanjin [1]. Hence, karanjin has gained traction as a potent anticancer phytochemical. However, in the previous chapter, we revealed the proliferative and partial estrogen-like activity of karanjin in hormone-responsive MCF-7 breast cancer cells. The contradiction led by our report is backed by Karanjin's flavonoid potential. The flavonoid backbone of karanjin might bring about its proliferative potential and disseminate weak estrogen-like properties. Flavonoids are known to bind ERα [91,161], modulate its transactivation function [91], and differentially affect cell proliferation and gene expression in a concentration or cell context-dependent manner [162–165]. The dose- and cell-dependent role of flavonoids is well established. Compounds such as genistein and quercetin elicit both proliferative effects at lower concentrations and anti-proliferative effects at higher concentrations [166].

The transcriptomic footprint of 10 µM karanjin and partial overlap with estrogen and tamoxifen-modulated gene signatures in MCF-7 cells envisaged karanjin to be a novel phytoestrogen [130]. Uncertainty exists as to whether the estrogen-like nature is universal and evident in other cell types. Hence, the global transcriptomic effect of karanjin in yet another ERα-positive cell may elucidate the plausible mode of action, genomic signatures, and the pathways involved. To address the universality of genomic-correlates induced by karanjin, a global transcriptomic analysis was performed in T47D breast cancer cells. The impact of varying concentrations of karanjin on T47D cells' viability and gene expression patterns in MCF-7 and T47D cells were determined and illustrated a selective modulatory effect. Interestingly, karanjin-mediated regulation on T47D cells via ERα partly converged with MCF-7 cells and reiterated a partial estrogen-like effect.

5.2 Results

5.2.1 Global transcriptome profile of T47D breast cancer cells treated with karanjin

A partial-E2-like property of karanjin in a hormone-responsive MCF-7 breast cancer cells instigated us to explore the universality of this comportment. To mine the genomic footprint and explore the unique global transcriptome profile of T47D breast

cancer cells, an RNA-seq analysis was performed. We identified karanjin-regulated genes in T47D cells by analyzing the read count data with the DESeq2 package in R. The volcano plot in Figure 5.1A shows that 177 genes (blue dots) were modulated by karanjin, based on a threshold value of 0 for LFC and 0.05 for Wald statistic and FDR. In total, 93 genes were repressed, and 84 genes were induced by karanjin, as illustrated by the expression heatmap in Figure 5.1B. Among the top karanjin-regulated genes *CYP1B1*, *CENPE*, and *MYB* were validated by qRT-PCR.

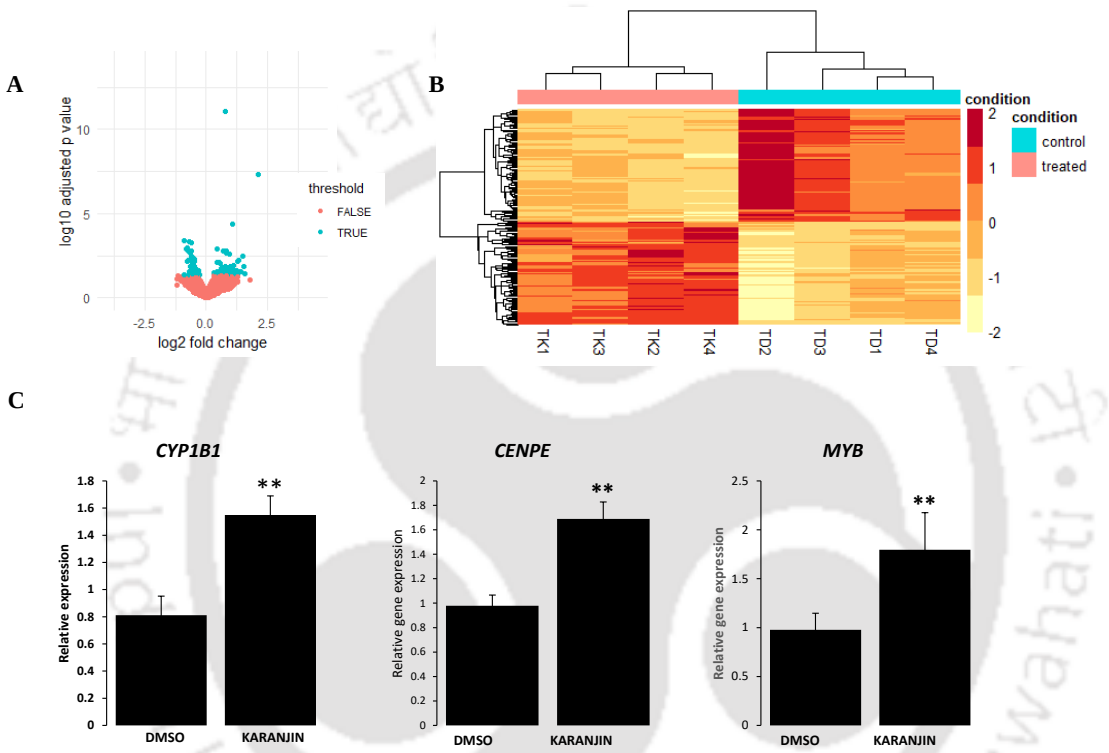


Fig. 5.1: Effect of karanjin on gene expression in T47D cells. A Volcano plot. Each point represents a gene, which is plotted according to its $-\log_{10}P_{adj}$ and $\log_2FC$. The blue dots represent genes (n = 177) that are significantly modulated by 10 µM karanjin, whereas the red dots represent non-significant genes. B Expression heatmap of 177 significantly modulated genes. TD1, TD2, TD3, and TD4 are control samples. TK1, TK2, TK3, and TK4 are karanjin-treated samples. The color for each gene on the heatmap represents its z-score obtained from the normalized count. C Validation of RNA seq data using qRT-PCR. Total RNA isolated from T47D cells treated with vehicle or karanjin (10 µM) for a period of 24 h was subjected to qRT-PCR analysis of the indicated genes. The ΔΔCt method was used for analyzing the qRT-PCR data using *Cyclophilin A* as a normalizing control. The experiment had three replicated dishes of control or karanjin-treated cells. The mRNA expression in each sample was expressed relative to the expression of one of the control samples. Bars represent mean relative expression ± sd (n = 3). For each gene, the data were analysed by one-tailed t-tests. Asterisks represent significant results. (**p < 0.01, ***p < 0.001).

Karanjin modulated diverse class of genes which belongs to transporters,

transcription factors, enzymes, inhibitors, coactivators, and receptors, which may be responsible for modulating various signalling pathways. Table 5.1 and 5.2 describes the top 25 upregulated and downregulated gene upon exposure to 10 µM karanjin for 24 h in T47D.

Table 5.1: Top 25 up-regulated genes in T47D cells when treated with 10 µM karanjin for 24 h.

Gene Symbol	Name	FC
CYP1A1	cytochrome P450 family 1 subfamily A member 1	5.1
CYP1A2	cytochrome P450 family 1 subfamily A member 2	4.3
ZBED6	zinc finger BED-type containing 6	3.1
CLGN	calmegin	3.0
CENPE	centromere protein E	2.6
PRPF18	pre-mRNA processing factor 18	2.6
CEP126	centrosomal protein 126	2.5
RPS6KA6	ribosomal protein S6 kinase A6	2.5
ARHGAP15	Rho GTPase activating protein 15	2.4
BFSP2	beaded filament structural protein 2	2.3
RESF1	retroelement silencing factor 1	2.3
ODR4	odr-4 GPCR localization factor homolog	2.3
ZNF569	zinc finger protein 569	2.3
FAM161A	FAM161 centrosomal protein A	2.3
ZNF449	zinc finger protein 449	2.2
TCEAL9	transcription elongation factor A like 9	2.2
LRRCC1	leucine rich repeat and coiled-coil centrosomal protein 1	2.2
SYNE1	spectrin repeat containing nuclear envelope protein 1	2.2
LURAP1L-AS1	LURAP1L antisense RNA 1	2.2
RGL1	ral guanine nucleotide dissociation stimulator like 1	2.2
CALCR	calcitonin receptor	2.2
TBC1D19	TBC1 domain family member 19	2.1
ZNF33B	zinc finger protein 33B	2.1
TBC1D8B	TBC1 domain family member 8B	2.0
PDZK1	PDZ domain containing 1	2.0

Table 5.2: Top 25 down-regulated genes in T47D cells when treated with 10 µM karanjin for 24 h.

Gene symbol	Name	FC
FAM38A-AS	novel transcript, antisense to FAM38A	0.53
L1CAM	L1 cell adhesion molecule	0.53
LINC01547	long intergenic non-protein coding RNA 1547	0.54
ANGPTL4	angiopoietin like 4	0.56
DMPK	DM1 protein kinase	0.59
SLC52A2	solute carrier family 52 member 2	0.59
SBK1	SH3 domain binding kinase 1	0.60
EGLN3	egl-9 family hypoxia inducible factor 3	0.60
TP53INP2	tumor protein p53 inducible nuclear protein 2	0.61
CRLF1	cytokine receptor like factor 1	0.61

NDUFS7	NADH:ubiquinone oxidoreductase core subunit S7	0.62
KCNC1	potassium voltage-gated channel subfamily C member 1	0.63
TNS4	tensin 4	0.64
ZDHHC12	zinc finger DHHC-type palmitoyltransferase 12	0.64
FASN	fatty acid synthase	0.64
FADS2	fatty acid desaturase 2	0.64
MPG	N-methylpurine DNA glycosylase	0.64
CCND3	cyclin D3	0.65
SREBF1	sterol regulatory element binding transcription factor 1	0.65
UNC13D	unc-13 homolog D	0.65
EFNB2	ephrin B2	0.65
CAPN5	calpain 5	0.66
APOL2	apolipoprotein L2	0.66
SRM	spermidine synthase	0.66
DYNC2I2	dynein 2 intermediate chain 2	0.66

5.2.2 Karanjin modulates estrogen-response late and xenobiotic-response pathway in T47D cells

Gene set enrichment analysis (GSEA) was applied to identify hallmark gene sets that were enriched upon karanjin treatment. The significantly enriched gene sets, based on FDR<25%, are indicated by red bars (Fig. 5.2A). Karanjin directed positive enrichment of E2-response late gene sets, similar to MCF-7, consolidating the partial E2-like nature of karanjin. There were 7 and 2 leading edge genes within E2 response late and Xenobiotic response enriched gene sets. We analysed and validated the significant induction of *PR* (n = 3, $p < 0.05$) and *CXCL12* (n = 4, $p < 0.05$) genes (Fig. 5.2B), and CYP1A1(n = 4, $p < 0.001$), and CYP1A2 (n = 4, $p < 0.001$) via qRT-PCR.

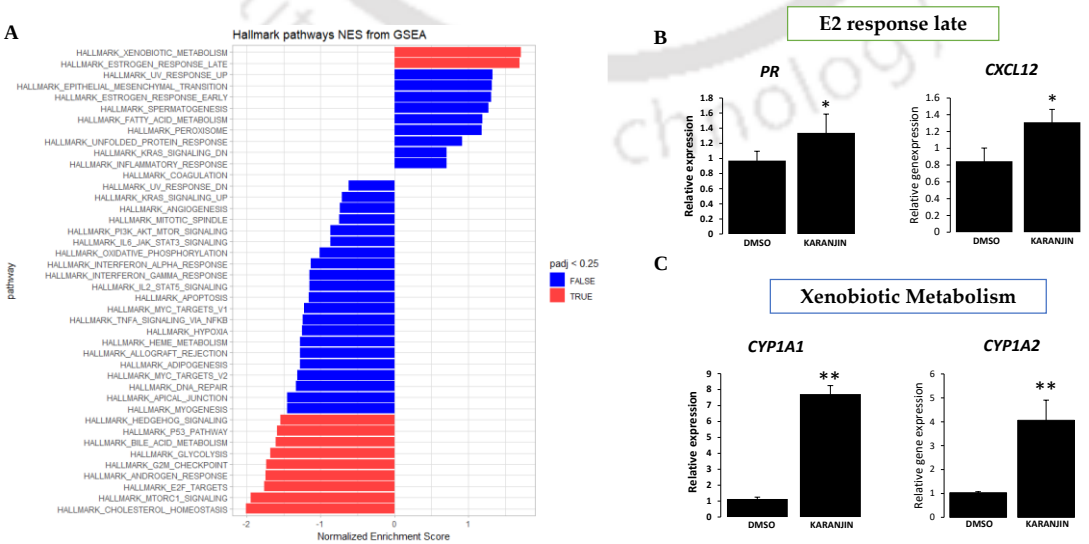


Fig. 5.2: Gene set enrichment analysis and validation of pathways regulated by karanjin in T47D cells. A. Bar chart showing enriched hallmark gene sets in T47D cells following 10 μM karanjin treatment using fGSEA. Bars represent NES. Blue bars correspond to the significantly enriched gene sets based on an FDR cut-off of 25%. B. Validation of the karanjin modulated E2-response late gene-sets. C. Validation of the karanjin modulated Xenobiotic metabolism gene set. The ΔΔCt method was used for analyzing the qRT-PCR data using *Cyclophilin A* as the normalizing control. The mRNA expression in each sample was expressed relative to the expression of one of the control samples. Bars represent the mean relative expression ± sd (n = 3). For each gene, the data were analyzed by a one-tailed *t*-test. Asterisks represent significant results. (*p < 0.05, **p < 0.01, ***p < 0.001).

5.2.3 Impact of karanjin on MCF-7 and T47D Transcriptome and gene expression signature.

Among the diverse set of genes modulated by karanjin, differential regulation of genes in MCF-7 and T47D was observed. Re-analysis of RNA-seq data (GSE117942) of MCF-7 cells treated with 1 nM E2 or 1 μM Tamoxifen [138] yielded two overlapping sets of 3727 estrogen-regulated and 425 tamoxifen-regulated genes. Analysis of the karanjin-regulated gene in MCF-7 and T47D cells yielded 736 and 177 genes, respectively. Karanjin-modulated genes in T47D and MCF-7 gene lists were matched with those regulated by estrogen or tamoxifen-regulated gene lists, and a Venn diagram was created using the Interacti Venn web-based tool [167]. Overall, 123 genes were regulated by karanjin and estrogen, whereas 19 genes were regulated by karanjin (MCF-7) and karanjin (T47D). In total, 11 genes were commonly regulated by estrogen, karanjin (MCF-7), and karanjin (T47D), while 20 genes were common in karanjin (T47D) and Tamoxifen gene set. Only 2 genes, belonged to the overlapping set of gene from Tamoxifen, E2, and karanjin (T47D and MCF-7) gene list. (Fig. 5.3A). The comparative analysis of the gene list suggests selective regulation of genes induced by karanjin and predicts the SERM-like nature of karanjin.

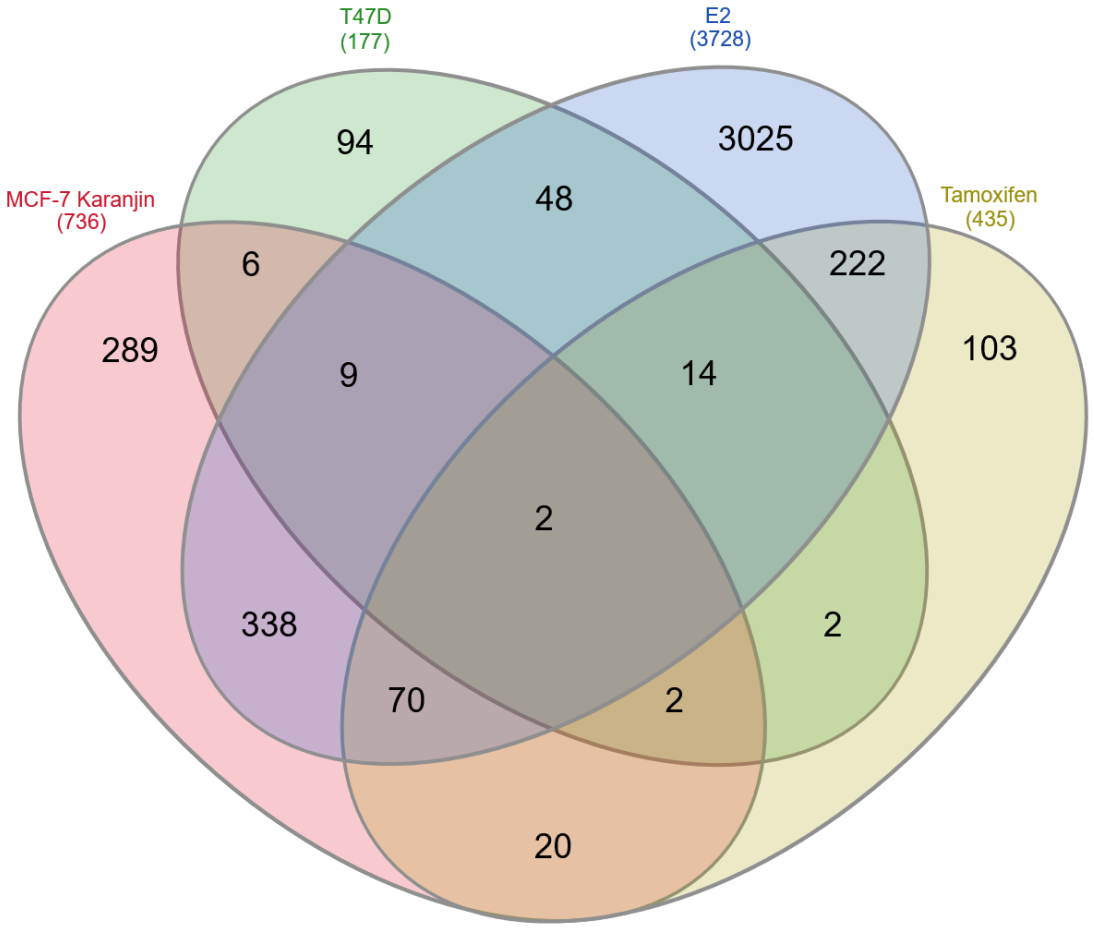


Fig. 5.3: Overlapping genomic targets of karanjin, estrogen, and tamoxifen in MCF-7 and T47D cells. Venn diagram representing the overlap of karanjin-treated T47D or MCF-7 and E2-or Tamoxifen-treated T47D cells gene list. Light orange color-gene list of karanjin-treated MCF-7 cells, light green-gene list of karanjin-treated T47D cells, light blue-gene list of E2-treated MCF-7 cells, and light brown-gene list of tamoxifen-treated MCF-7 cells.

To gain insights into gene expression patterns in MCF-7 and T47D cells, select leading-edge genes of the E2-response gene set were analyzed via qRT-PCR. Differential regulation of *PS2*, *PR*, *STC2*, *MYC*, *SLC7A5*, *MYC*, *NCOA3*, *B3GALT5*, and *AHR* were observed in T47D compared to MCF-7 cells. *TIPARP*, *CYP1A1*, *CYP1B1*, *CYP1A2*, *ANKRD29*, *CENPF*, and *CENPE* depicted similar gene regulation in both MCF-7 and T47D cells. The gene expression pattern of some of the E2 target genes is described in Table 5.3.

Table 5.3. Karanjin induced differential regulation of gene expression in MCF-7 and T47D cells

s	Gene	MCF-7	T47D
1	***TFF1/PS2***	Up	No regulation
2	***PR***	No regulation	Up
3	***TIPARP***	Up	Up
4	***STC2***	Up	No regulation
5	***CSTA***	Down	UP
6	***SLC7A5***	Up	No regulation
7	***AHR***	No regulation	Up
8	***CYP1A1***	Up	Up
9	***CYP1B1***	Up	Up
10	***CYP1A2***	Up	Up
12	***NCOA3***	Down	No regulation
13	***MYC***	Up	No regulation
14	***B3GALT5***	Down	UP
15	***ANKRD29***	Up	Up
16	***CENPE***	UP	UP
17	***CENPF***	UP	UP

5.2.4 Concentration- and cell type-dependent effect of karanjin on gene expression

The flavonoid nature of karanjin potentiates need to examine the concentration and cell-type dependent role of karanjin. The global transcriptome profile of MCF-7 and T47D cells treated with karanjin reveals E2 or SERM-like effects. Hence, the effects of varying concentrations of karanjin on the expression of leading-edge genes of estrogen-response-early genesets were addressed in two ER-positive cell lines, namely MCF-7 and T47D. Moreover, molecular phenotype and cell-type have a role in gene expression patterns [143,168]. The experimental design included cells treated with 10 nM E2 as a reference compound. E2 induces TFF1 expression in ER-positive breast cancer cells [169]. As expected, 10 nM E2 induced *TFF1* mRNA in both cell lines (Fig. 5.4A, B; $n = 3$, $p < 0.001$). 10 nM karanjin did not affect *TFF1* mRNA expression in MCF-7 cells. In the qRT-PCR-based validation of RNA-seq data, we confirmed the significant induction (~1.2 fold) of *TFF1* mRNA in MCF-7 cells by 10 μM karanjin (Fig. 4.3). The 1.2-fold induction of *TFF1* mRNA by 10 μM karanjin in Figure 5.4A (left panel) is consistent with RNA-seq data, although one-way ANOVA followed by Tukey HSD does not show significant result. 50 μM karanjin significantly induced *TFF1* mRNA expression in MCF-7 (Fig. 5A, left panel, $n = 3$, $p < 0.001$). None of the concentrations of karanjin modulated the expression of TFF1 in T47D cells (Fig. 5.4A, right panel). 10 nM E2 did not affect *TIPARP* mRNA expression in MCF-7 cells. In agreement with RNA-seq data, 10 μM karanjin induced *TIPARP* mRNA expression (Fig. 5.4B, left panel; $n = 3$, $p < 0.001$). In contrast, 10 nM karanjin repressed its

expression (Fig. 5.4B, left panel; n = 3, $p < 0.01$). In T47D cells, 10 nM E2 induced the expression of *TIPARP* mRNA (Fig. 5.4B, right panel; n = 3, $p < 0.01$). Karanjin did not modulate the expression of *TIPARP* mRNA in T47D cells. STC2 mRNA, another estrogen-response-early gene, was induced by 10 nM E2 in both cell lines (Fig. 5C, n = 3, $p < 0.001$ for MCF-7 cells, $p < 0.05$ for T47D). In MCF-7 cells, *STC2* mRNA was induced by karanjin only at a concentration of 10 µM (Fig. 5.4C, left panel, n = 3, $p < 0.01$). 10 nM and 50 µM karanjin induced *STC2* mRNA expression more than two-fold (Fig. 5C, left panel). However, the results were not statistically significant when analyzed by ANOVA followed by TukeyHSD ($p = 0.07$ for 10 nM, and $p = 0.09$ for 50 µM). None of the concentrations of karanjin significantly modulated the expression of *STC2* mRNA in T47D cells (Fig. 5.4C, right panel). *SLC7A5* mRNA expression was significantly induced by 10 nM E2 in T47D cells (Fig. 5.4D, right panel, n = 3, $p < 0.001$) but not in the MCF-7 cell (Fig 5.4D, left panel). In MCF-7 cells, there was significant induction of *SLC7A5* mRNA expression post 10 and 50 µM karanjin treatment (Fig. 5.4D, left panel, n = 3, $p < 0.001$ for 10 µM, and $p < 0.01$ for 50 µM). None of the concentrations of karanjin modulated its expression in T47D cells. *CD44* was not modulated by 10 nM E2 in both cell lines (Fig. 5.4E). Karanjin induced its expression only in MCF-7 cells at a concentration of 10 µM (Fig. 5.4E, left panel, n = 3, $p < 0.01$). Besides estrogen-response-early genes, we also studied the effect of karanjin on the expression of *CSTA*, a known estrogen-suppressed gene. As expected, 10 nM E2 suppressed *CSTA* mRNA expression in MCF-7 cells (Fig. 5.4F, left panel, n = 3, $p < 0.001$). Significant and progressive dose-dependent suppression of *CSTA* mRNA by karanjin was observed in MCF-7 cells (Fig. 5F, n = 3, $p < 0.001$). In T47D cells, 10 nM E2 did not affect *CSTA* mRNA expression, which is consistent with the previous findings. 10 nM and 10 µM karanjin did not modulate *CSTA* mRNA expression in T47D cells. However, in sharp contrast to MCF-7 cells, 50 µM karanjin significantly induced *CSTA* mRNA in T47D cells (Fig. 5.4F, right panel, $p < 0.001$). These data demonstrate that karanjin partially mimics estrogen-like effects on gene expression in a concentration- or cell-type-dependent manner.

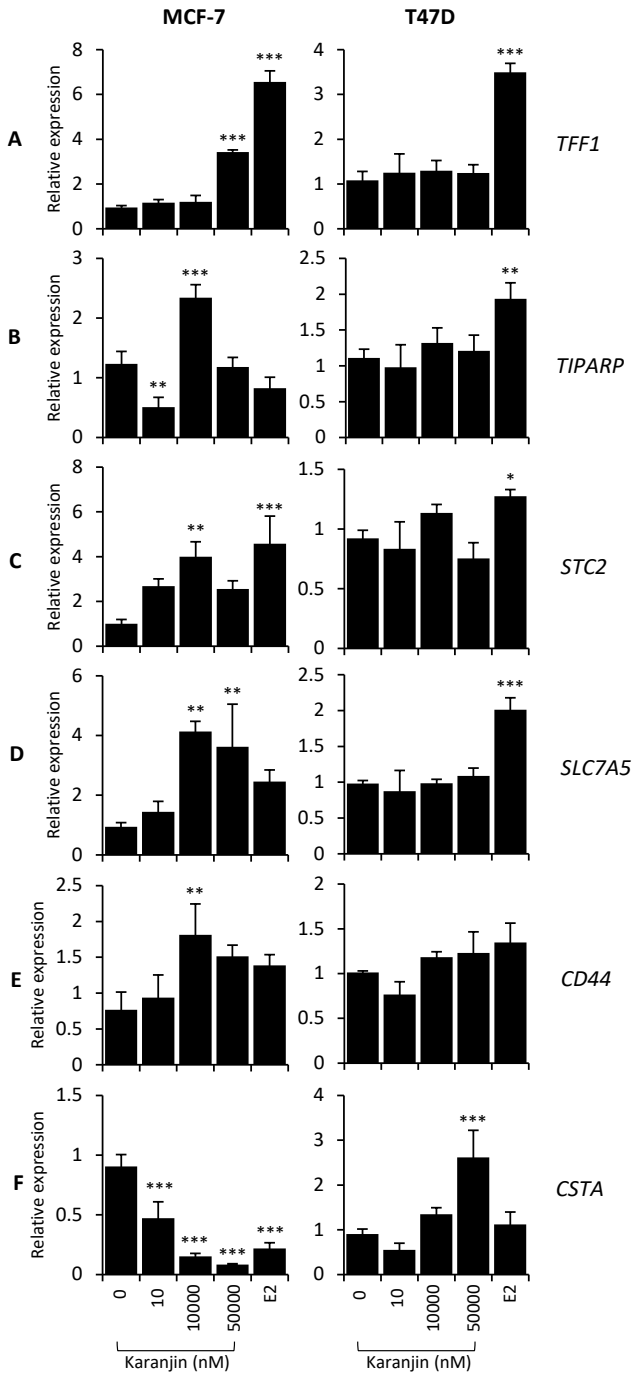


Fig. 5.4: Concentration-dependent effect of karanjin on gene expression in MCF-7 and T47D cells. Cells were treated with indicated concentrations of karanjin or 10 nM E2 for 24 h. Thereafter, total RNA was isolated and subjected to qRT-PCR analysis of the indicated genes using the ΔΔCt method. The experiment was done with three replicate dishes of control or karanjin-treated cells. The expression in each of the samples was expressed relative to one of the controls, which was set to 1. Bars represent mean relative expression ± sd (n = 3). For each gene, the data were analyzed by one-way ANOVA followed by TukeyHSD. Significantly different means with respect to control (0 µM karanjin) are indicated by asterisks. (*p < 0.05, **p < 0.01, ***p < 0.001).

5.2.5 Effect of varying dose of karanjin on cell viability

The short- and long-term impact of different concentrations of karanjin on the viable cell count was evaluated in ERα-positive T47D and ERα-negative MDA-MB-231 breast cancer cells to bring about the cell-context and concentration-dependent effect of karanjin. The cells were treated with karanjin at concentrations ranging from 0 to 50 µM, and viable cells were counted after 0, 24, and 120 hours. Cells treated with 10 nM E2, which stimulates the proliferation of ER-positive breast cancer cells [170], served as a control in the experiment. (Fig. 5.5).

We used a two-way ANOVA approach to analyze the effect of karanjin concentration on cell viability at time points of 0 h (no treatment: red dots), 24 h (short-term: green dots), and 120 h (long-term: blue dots) treatment. For the 24 h treatment, viable cell count, time, and concentration did not interact ($p > 1$). Significant changes in viable cell count have occurred over time. After 24 hours of treatment, the number of viable cells was much higher than 0 h ($p > 0.0001$). Concentration had no impact on viable cell count. After 24 h of treatment, there was no significant difference in the number of viable cells between the 0–50 µM karanjin treated groups represented by green dots in both T47D and MDA-MB231 cells (Fig. 5.5A, B). Our results revealed that after 24 h, karanjin had no appreciable effect on T47D or MDA-MB-231 cell viability, despite the fact that proliferation was visible and on parity with that caused by E2 (Fig. 5.5A, B).

Long-term exposure (120 h) to karanjin on T47D cell viability represented a significant interaction between time and concentration. Although the viable cell counts increased at all concentrations of karanjin as revealed by significant main effects of time ($p \approx 0$), the number of viable cells after 120 h of treatment depended on the concentration of karanjin (Fig. 5.5A, p-value for the interaction term ≈ 0). 1 nM karanjin treatment yielded a significantly higher viable cell count ($p = 0.02$), whereas 10 or 50 µM karanjin yielded significantly lower viable cell counts with respect to the control ($p \approx 0$). Thus, 10 µM karanjin had no effect on T47D cell viability over a period of 24 h and had an intermediate effect on viability over a period of 120 h, similar to MCF-7 cells reported in the previous chapter. On the contrary, in MDA-MB-231, there was no effect of concentration and no interaction between concentration and time ($p \approx 1$). The viable cell counts after 24 or 120 h of treatment with all concentrations of karanjin (0-

50 µM) were not significantly different (Fig. 5.5B). In MDA-MB-231, no reduction in cell number was observed even at higher dosages of 10 µM and 50 µM of karanjin. This differential effect on cell viability could be due to the proliferative actions of hormone-responsive receptors other than ERα, which act as a balancing factor to keep cell numbers under control.

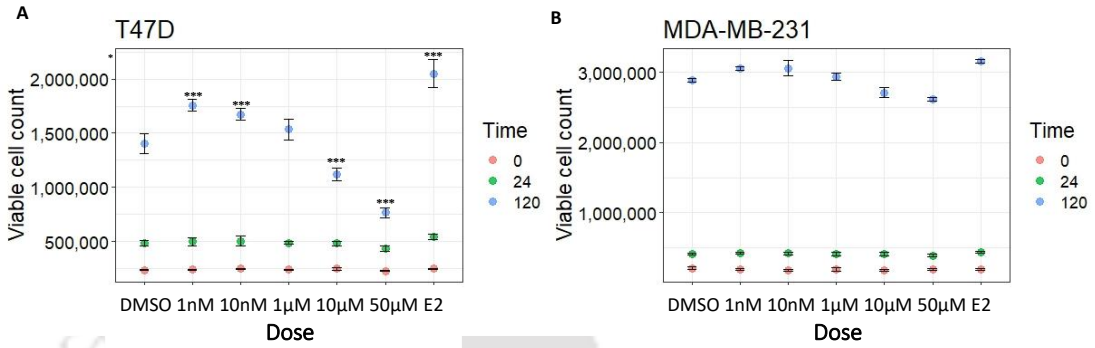


Fig. 5.5: Concentration-dependent effect of short and long-term stimulation with karanjin on T47D and MDA-MB-231 cell viability. Cells were treated with either vehicle or the indicated karanjin concentrations. A. T47D. B. MDA-MB231, viable cells were counted after 24 h (green dots) or 120 h (blue blots) based on trypan blue dye exclusion using a hemocytometer. Cells treated for 0 h provided the initial viable counts for each treatment group (red dots). Each dot represents the mean viable count ± sd (n = 3).

5.3 Discussion

Various research groups have previously demonstrated *in vitro* effects of karanjin over a wide concentration range. We found no evidence for any short-term (24 h) effect of karanjin (0-50 µM) on MCF-7 viable cell count [130]. The long-term (120 h) effect was significantly concentration-dependent, although the viable cell count increased at all concentrations of karanjin. The viable count was significantly lower at higher concentrations (10 and 50 µM), whereas it was modestly but significantly higher at lower concentrations (10 nM) and (1nM) for MCF-7 and T47D cells, respectively. We reported no evidence of short- and long-term exposure to karanjin on viable cell counts of ER-negative MDA-MB-231 breast cancer cells. These data suggest that the antiproliferative effect of karanjin may not be universal but rather be cell/receptor-context or concentration-dependent. Karanjin, like other flavonoids, such as genistein, quercetin, or daidzein, demonstrates a dual nature [13,166,171]. This presents a caution to the anticancer potential of karanjin and underscores the importance of dosage for various healthcare applications suggested in the literature.

Here, we present possible explanations behind the contradictory proliferative nature of karanjin in breast cancer cell type. Although 10 µM karanjin treatment for 24 or 120 h produces similar or significantly lower viable counts, respectively, compared to control, the proliferation of cells is evident from a significant (1.55-fold and 5.9-fold, respectively) increase compared to baseline (0 h) viable count. The cells could not have proliferated without cell cycle progression. In the event of cell cycle arrest followed by apoptosis, the viable counts are expected to go below the baseline. Thus, it is possible that at 10 µM, karanjin delays cell cycle progression rather than induces arrest. The delay could be due to other yet unknown effects of karanjin. It is to be noted that even at 50 µM, there is the proliferation of cells, although the final viable count after 120 h is even lower than that resulting from 10 µM. On the other hand, 10 nM and 1nM karanjin results in significantly more viable counts compared to control in MCF-7 and T47D cells, respectively. The concentration-dependent effect indicates that there is more than one receptor for karanjin in MCF-7 and T47D cells. The high-affinity receptors could be responsible for the proliferative actions of karanjin at lower concentrations. In contrast, low-affinity receptors could be responsible for cell cycle arrest at high concentrations.

In some studies, conclusions about the antiproliferative effects of karanjin were based on MTT assays, which are end-point assays. This obscures the initial viability at the start of the experiment, which makes it difficult to conclude whether the treatment caused cell death or cell cycle arrest. This distinction is possible with trypan blue dye exclusion assay. Furthermore, mechanisms of cell cycle arrest and pro-apoptotic effects of karanjin were shown in HeLa, HepG2, A549, HL-60, and MDA-MB-231 but not in MCF-7 and T47D cells [8–11]. Thus, the apparent contradictions could be due to the intrinsic properties of MCF-7 and T47D cells. Upon treatment with karanjin, the transcriptome profile of T47D breast cancer cells demonstrates an E2-like effect, at least in parts via ERα. PR gene, hallmark E2 regulated gene upregulated by 10uM karanjin. Similarly, positive enrichment of E2-response early and late pathways indicates an estrogen-like nature. However, the G2/M checkpoint and E2F gene set are negatively enriched, underscoring the delayed cell cycle progression in T47D at 10 µM karanjin dosage. Both G2M checkpoint and E2F target genes are involved in the cell cycle; however, each of them works on distinctively different phases of the cell cycle. The G2M checkpoint occurs between the G2 and M phases, whereas E2F plays a major

role during the G1/S transition [172].

The effect of varying concentrations of karanjin on gene expression of MCF-7 and T47D was observed. We found differential regulation of E2-response early pathway genes when we analyzed the gene expression changes in MCF-7 and T47D. The concentration range at which karanjin induces gene expression changes is within the physiological range of around 10 µM, making it potentially therapeutic. A therapeutic drug with a concentration over 5-10 µM is deemed inconsequential due to its potential for off-target and harmful consequences. Furthermore, attaining such elevated concentrations within the body for a specific agent is unattainable. The enrichment of E2-response genesets and ERα target genes in both MCF-7 and T47D is worthy of attention. ERα plays a major role in the estrogen-mediated proliferation of breast cancer cells [173]. The similar or differential effects of karanjin on estrogen or tamoxifen-regulated genes, coupled with its possible ERα modulatory effect, exposes a caveat to its consideration in anticancer therapy, lest it be counter-productive in hormone-responsive tumors. Thus, an in-depth mechanistic investigation into concentration and cell-type dependent effects of karanjin and possible ERα modulatory effect needs urgent attention. This aspect has been addressed in chapter 6 of the thesis.

Indian Institute of Technology Guwahati

CHAPTER 6

Partial estrogen-like effects of karanjin and the mechanism involved

6.1 Introduction

Flavonoids are known to exert their hormone-like effects through nuclear receptors [14]. Besides their wide range of biological and pharmacological properties, flavonoids possess weak estrogenic properties. Since estrogen is known to be a tumor-promoting agent, it increases the risk of breast and uterine cancer in women administered with estrogen replacement therapy [26]. Due to this estrogenic potential, flavonoids are considered a safe alternative and dietary component used in HRT and chemoprevention, such as Genistein, Daidzein and soy isoflavones [174–176].

In our previous chapters, we revealed that karanjin modulates gene expression in the ERα-positive cell lines, suggesting partial estrogen-like activity. Karanjin affects the proliferation of MCF-7 and T47D cells in a concentration-dependent manner. The global Transcriptomic footprint of 10 µM karanjin on MCF-7 and T47D breast cancer cells demonstrated cell cycle progression, estrogen signaling, and genomic effects akin to estrogen [130]. A partial overlap with E2 and Tamoxifen-regulated genes and the regulation of ER-target genes by karanjin suggests a possible involvement of ERα.

Mechanisms of karanjin regulation are yet to be studied extensively. However, in the context of cancer, much of our understanding deals with the anti-proliferative and cell cycle inhibitory effects of karanjin. The dose- and cell-type dependent karanjin effect on hormone-responsive breast cancer cells and its effects pertaining to modulation of gene expression are poorly understood. Describing karanjin's cellular actions is critical, particularly considering its flavonoid nature and ERα modulatory property. It remains to be explored whether karanjin directly binds ERα and modulates its transactivation function in a concentration-, gene-, or cell-type-dependent manner as exhibited by selective estrogen receptor modulators [18,159].

In summary, the present chapter provides valuable insights into the role of karanjin in breast cancer cells by delineating the possible role and involvement of ERα in karanjin-mediated effects. The concentration- and cell-type-dependent effects and the ERα modulatory activity envisage karanjin as a novel SERM. Furthermore, the mechanistic detail of ERα-mediated karanjin regulation of E2 target genes utilizing *in silico* and *in vitro* approaches have been addressed.

6.2. Results

6.2.1 In silico molecular modeling study to predict the involvement of ERα in karanjin-mediated effects.

Computational techniques have often been used to complement experimental studies. Docking is one of the popular techniques commonly used for several purposes, e.g., ligand pose prediction, ligand binding affinity prediction as well as identifying potential actives. The partial overlap in karanjin- and tamoxifen-regulated genes and E2-like nature prompted us to investigate possible ERα-modulatory effects. Whether karanjin directly binds ERα and modulates, its function remains to be explored. ERα plays a crucial role in the development and progression of breast cancer. Understanding the binding of karanjin to ERα can provide insights into its potential as a therapeutic agent. To predict how karanjin interacts with ERα structurally and energetically, we docked eight wild-type E2 (agonist) or 4-hydroxy Tamoxifen (antagonist) (Fig. 6.1) bound structures using Autodock 4.2.6.

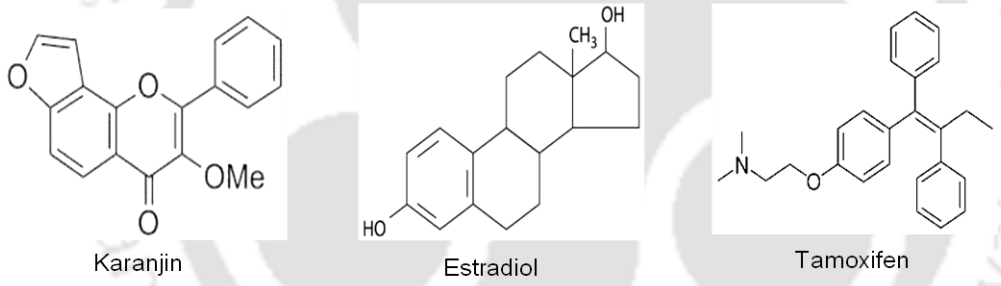


Fig. 6.1: Karanjin, Estradiol (E2), and Tamoxifen structure. Karanjin is a flavonoid, Estradiol (E2) is an active form of estrogen and endogenous ligand for ERα (agonist), and Tamoxifen is a selective estrogen receptor modulator (SERM) and antagonist to ERα.

The docking protocol was validated by docking the native ligand and calculating the root-mean-square deviation (RMSD) between the docked and native poses. The results showed that the docking protocol was reliable for further analysis of karanjin-ERα LBD interactions. E2 was docked to the ligand binding pocket, and the RMSD found was less than 1 Å in all eight E2 bound ERα LBD structures (Table 6.1).

Table 6.1: List of estrogen-bound PDB structure of wild type ERα LBD.

PDB	Resolution	Chains	Length	Ligand	RMSD
1A52	2.8	A/B	297-554	E2	0.6 A
1ERE	3.1	A/B/C/D/E/F	301-553	E2	0.6 A
1G50	2.9	A/B/C	304-550	E2	0.5 A
1GWR	2.4	A/B	305-549	E2	0.7 A
1QKU	3.2	A/B/C	301-550	E2	0.5 A
2YJA	1.82	B	299-551	E2	0.6 A
5GS4	2.4	A	305-547	E2	0.4 A
5GTR	2.8	A	305-547	E2	0.9 A

Figure 6.2A demonstrates the docked conformation of 1GWR E2-bound structure and the H-bonds associated with it. 1GWR served as a reference for the native conformation, denoted in pink. In contrast, the docked E2 conformation depicted in blue represents a predicted binding. The ligand in black color represents E2 from the 1GWR crystal structure, whereas the ligand in red represents E2 from the docked structure (Fig. 6.2B). The RMSD of the structures was found to be 0.723Å, indicating that the binding pose of the ligand is reliable and AutoDock protocol is suitable for virtual screening. However, experimental validation is still necessary to confirm the binding affinity and specificity of the ligand.

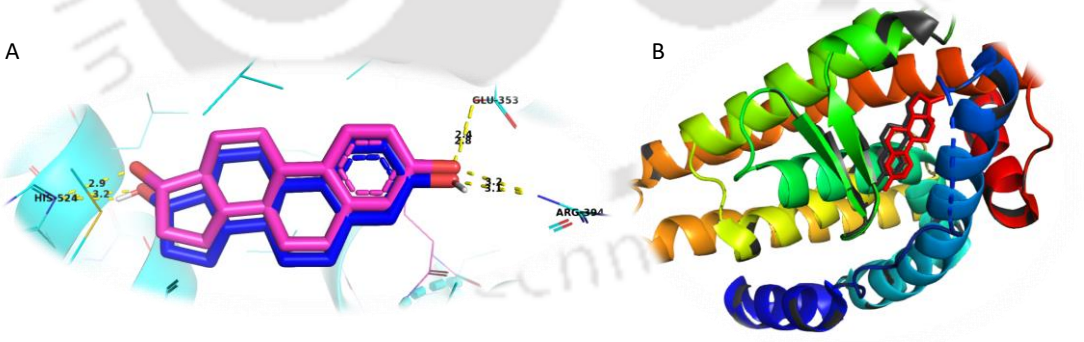


Fig. 6.2: Validation of docking protocol. A. Pink-1GWR from PDB; Blue-1GWR-E2 relocked B. Black corresponds to 1GWR, and Red represents redocked E2 conformation with AutoDock software.

After the validation of the protocol with the endogenous E2 ligand, each crystal structure was docked with karanjin. The docked conformations were visualized via LigPlot software (Fig. 6.3). The ERα LBD is composed of 12 helices and two sheets,

and amino acid residues such as Leu 346, Ala 350, Leu 384, Leu 387, Phe 404, Val 418, Met 421, Ile 424, His 524, Glu 353 and Leu 525 are essential for E2 interaction with the ERα LBD [177]. We observed similar binding site residues on docking of karanjin with ERα with Leu 346, Ala 350, Leu 387, Phe 404, Met 421, Ile 424, His 524, Glu 353, and Leu525. Val 418 and Leu 383 were missing contact residues in the karanjin-bound structure compared to the E2-bound ERα LBD. E2 formed hydrogen bonds with His 524, Arg 394, and Glu 353 in eight estrogen-bound, nonmutant ERα LBD structures. Similar to E2, docking study predicted His 524 forms an H-bond with the furan ring oxygen of karanjin (Fig. 6.3). Furan ring oxygen imparts hydrophilicity to the karanjin structure and thus could be involved in forming these polar contacts with the ERα LBD.

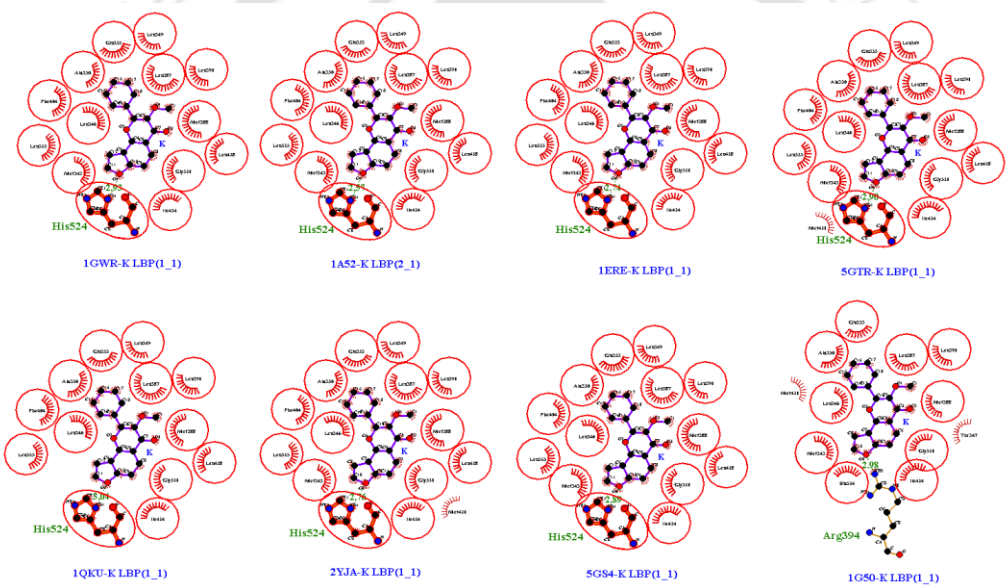


Fig. 6.3: Karanjin bound eight ERα LBD structures visualized in LigPlot. H-bond formed is shown with His 524, and other residues represent the contacting residues.

The conformation of karanjin in seven out of eight structures coupled to wild-type structures were identical, except for the 1G50 structure, whose docked conformation is upturned relative to other karanjin-docked structures (Fig. 6.4). The possible explanation lies in the alteration in the ligand binding pocket of 1G50 [178].

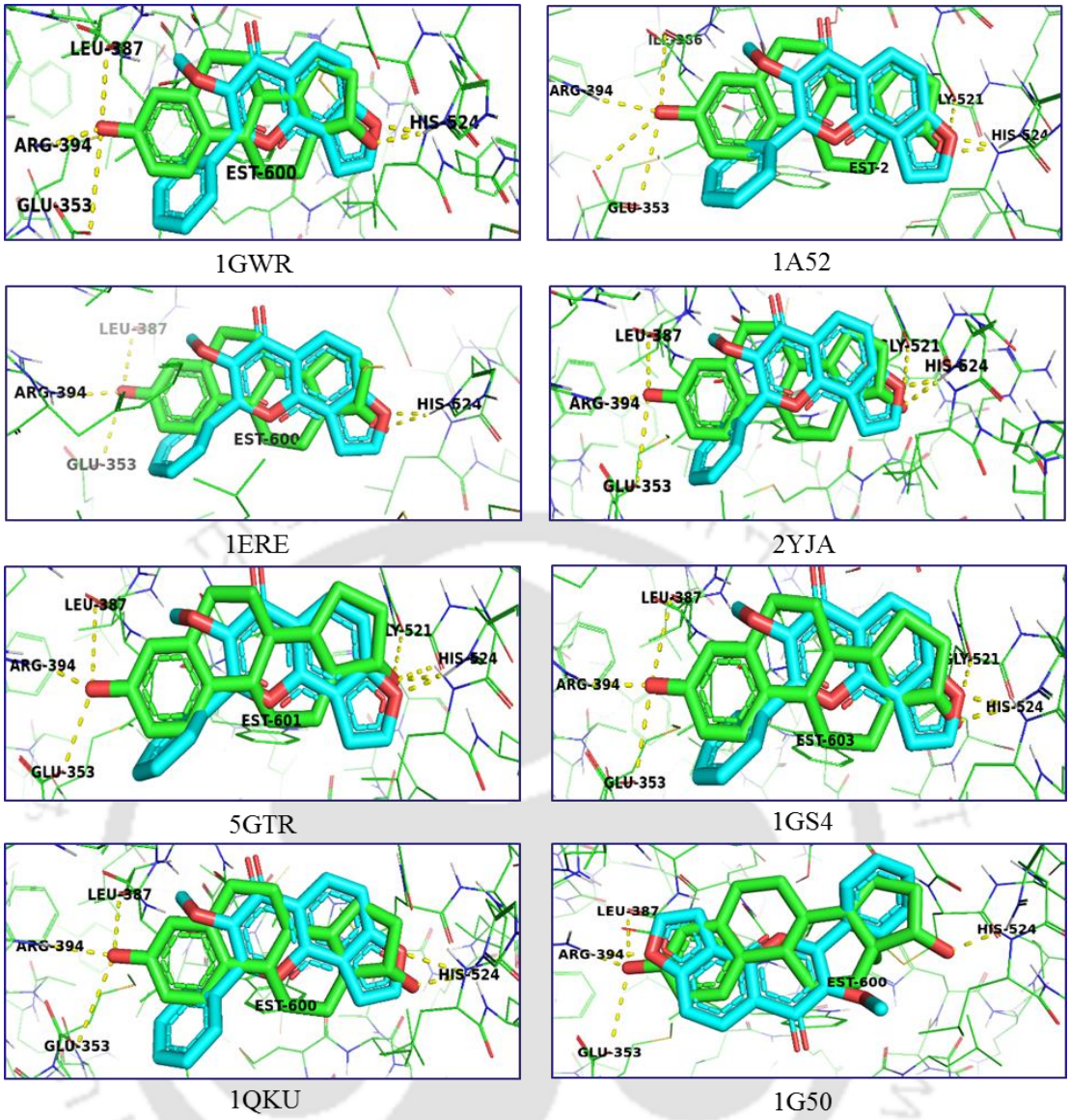


Fig. 6.4: Visualization of superimposed conformations of karanjin and E2 bound to ERα. PyMOL-derived depiction of karanjin(cyan) and E2(Green) conformation docked to eight ERα LBD structures.

In almost all the docked structures, the binding energy of karanjin was roughly 2 kcal/mol lower than that of E2. This energetic difference is electrostatic, attributed to the contribution of two additional H-bonds formed with Arg 394 and Glu 353 in the E2-bound structure. Table 6.2 describes the binding energy score and H-bonds of E2 and karanjin for eight agonist (E2)- and two antagonists (SERM)-bound crystal structures. Karanjin bound ERα LBD demonstrated binding energy and conformation in the ligand binding pocket equivalent to E2 and established an H-bond with the furan ring oxygen and His 524 (Table 6.2).

Table 6.2: List of docked structures with E2 and karanjin.

PDB	E2		Karanjin	
	Binding Energy(kcal/mol)	**H bonds**	**Binding Energy(kcal/mol)**	**H bonds**
1GWR	-10.45	H524, R394, E353	-8.35	H524
1A52	-10.22	H524, R394, E353	-7.84	H524
1ERE	-10.47	H524, R394, E353	-8.36	H524
1G50	-10.63	H524, R394, E353	-8.34	R394
1QKU	-9.82	H524, R394, E353	-7.94	H524
2YJA	-10.9	H524, R394, E353	-8.67	H524
5GS4	-10.59	H524, R394, E353	-8.17	H524
5GTR	-11.34	H524, R394, E353	-9.14	H524
1GWQ(RAL)	-9.03	R394	-8.61	None
3ERT(OHT)	-9.96	R394	-7.87	None

Karanjin, like E2, formed direct polar contacts (H-bonds) with His 524 and was tethered to the ligand binding pocket with similar contacting residues (Fig. 6.5A & C). Polar interactions with Arg 394, Glu 353, and His 524 were produced in E2 bound to 1GWR (Fig. 6.5B). The helix 12 conformations of the ERα LBD play a vital role in determining whether a ligand is an agonist or antagonist. There is a change in helix 12 conformation due to the absence of an H-bond with His 524 when antagonist 4-Hydroxytamoxifen (4OHT) is bound [179,180]. When docked with antagonists bound ERα LBD structures (4OHT, PDB: 3ERT (Fig. 6.5D) and Raloxifene, PDB: 1GWD), karanjin failed to form any H-bonds with His 524 (Table 6.2). Concomitantly, the present *in-silico* assessment complements the previously described partial E2-like (agonist) nature of karanjin.

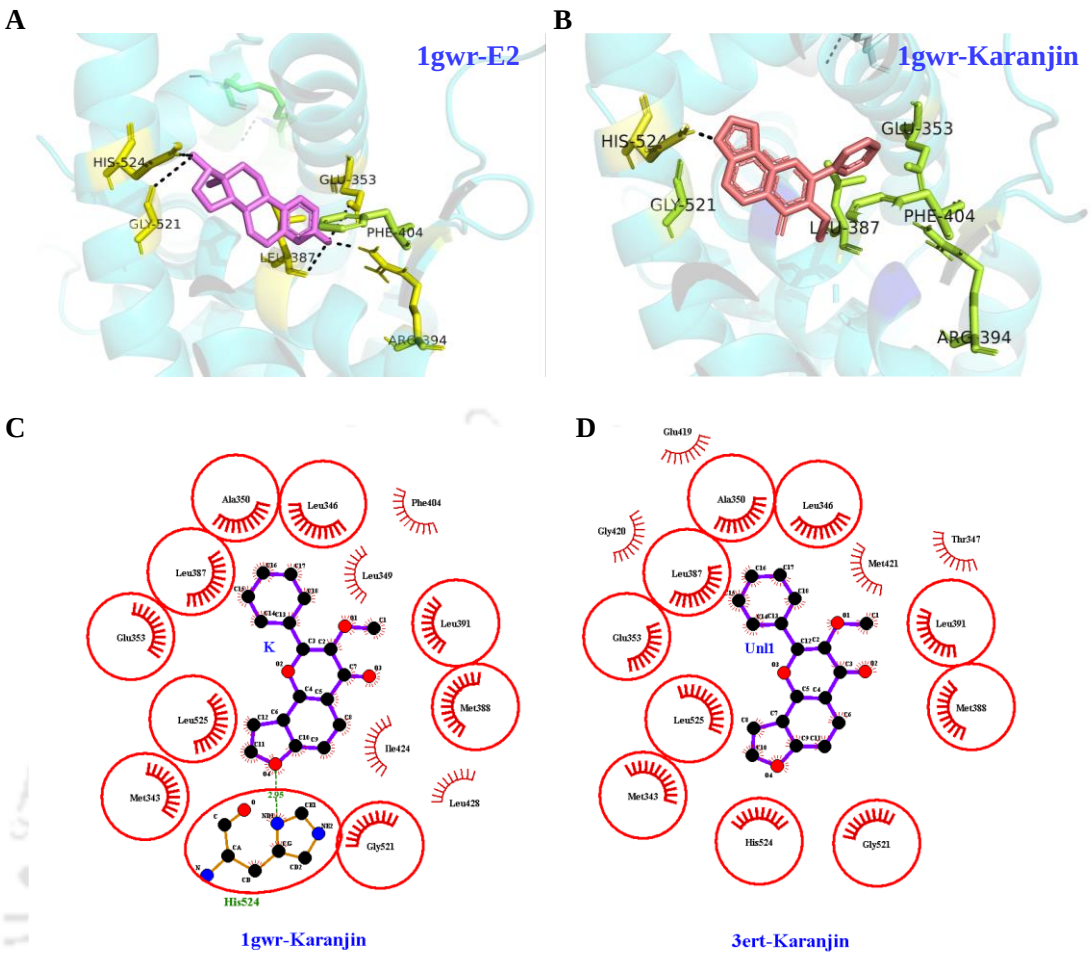


Fig. 6.5: Visualization of E2 and karanjin bound to ERα LBD. Autodock-derived depiction of conformation of A. E2 (pink) and B. Karanjin (red) bound to ERα LBD structure (PDB 1GWR). The black dotted line represents polar contacts. Ligplot-derived representation C. Karanjin bound to 1GWR (agonist conformation), the green dotted line represents H-bond with His 524. D. Karanjin bound to 3ERT (Antagonist conformation), no H-bond formed with ERα residues. The red rings with amino acid residues represent nearby residues.

6.2.2 *Karanjin modulates gene expression via ERα in a gene-dependent manner*

Modulation of estrogen-response-early genes led us to hypothesize the involvement of ERα, at least in part. Post E2-mediated activation, ERα is typically turned over by proteasomal degradation [181]. There was a significant decrease in ERα protein in MCF-7 cells after treatment with 10 μM karanjin for 24 h, which further decreased with time (Fig. 6.6A). Thus, post karanjin stimulation, the fate of ERα is similar to that brought about by estrogen.

To gather more evidence for the involvement of ERα, we examined the effect of ERα knockdown on karanjin-mediated modulation of gene expression. qRT-PCR was

applied to analyze the expression of selected karanjin-modulated genes in MCF-7 cells, which were treated with vehicle or 10 μM karanjin after prior treatment with scrambled or ERα-specific siRNA. As shown in Figure 6.6B, ERα protein was undetectable in cells treated with ERα-specific siRNA. Gene expression data were analyzed by two-way ANOVA to test the effect of karanjin treatment as a function of ERα status. We first analyzed the expression levels of *PR* and *TFF1* mRNAs. These genes are ERα-dependent classical estrogen-induced genes whose expression is significantly reduced upon ERα knockdown [144,168,169,182]. As expected, a significant main effect of ERα knockdown was observed on PR mRNA levels in MCF-7 cells (Fig. 6.6C, $p \approx 0$). The PR mRNA steady-state levels were significantly repressed upon ERα knockdown, which suggested the regulatory role of ERα in PR gene expression. However, there was no evidence of an interaction between karanjin and ERα ($p = 0.98$). Irrespective of the ERα status, karanjin did not modulate the levels of *PR* mRNA. This result is not surprising since our RNA-seq experiment did not reveal *PR* mRNA modulation by karanjin. A similar result was obtained for *TFF1*. There was no interaction between karanjin and ERα ($p = 0.485$), but there was a significant main effect of ERα (Fig. 6.6D, $p < 0.001$). Karanjin induced *TFF1* mRNA levels by 1.2-fold in cells treated with scrambled siRNA. However, this induction was not significant after multiple comparisons, although the induction is comparable to that significant induction inferred from the RNA-seq data (Fig. 6.6F). In the case of *CSTA* mRNA, there was no evidence of the interaction between karanjin and ERα (Fig. 6.6E, $p = 0.37$). However, there were significant main effects of karanjin treatment ($p = 0.007$) and ERα ($p \approx 0$). In contrast, analysis of *ADAMTS19* revealed that there was a significant interaction between karanjin and ERα (Fig. 6.6F, $p = 0.02$). Karanjin treatment significantly reduced *ADAMTS19* mRNA in scrambled siRNA-treated cells but not ERα siRNA-treated cells. ADAMTS19 is an ERα target gene, and upon E2 stimulus, there is downregulation in the steady-state level of ADAMTS19. This suppression is rescued and abrogated when the ERα is silenced or knocked down from the system via siRNA-mediated knockdown. Hence, signifying the karanjin-regulated ADAMTS19 mRNA level is under the regulatory control of ERα. We also analyzed the mRNA expression of other estrogen-response early genes, such as *STC2*, *SLC7A5*, and *TIPARP*. For *STC2* and *SLC7A5*, a significant interaction between karanjin and ERα (Fig. 6.6G, H; $p < 0.001$) was found; the induction in mRNA expression by karanjin being greater in cells treated with scrambled siRNA compared to ERα-specific siRNA. In the case of *TIPARP*, only the

main effect of karanjin was significant (Fig. 6.6I, $p < 0.001$). We also analyzed *CYP1A1* mRNA expression, which does not fall under the estrogen-response-early gene set. There was a modest but significant interaction between karanjin and ERα (Fig. 6.6J, $p = 0.042$), with significant main effects of karanjin ($p < 0.001$) and ERα ($p = 0.03$).

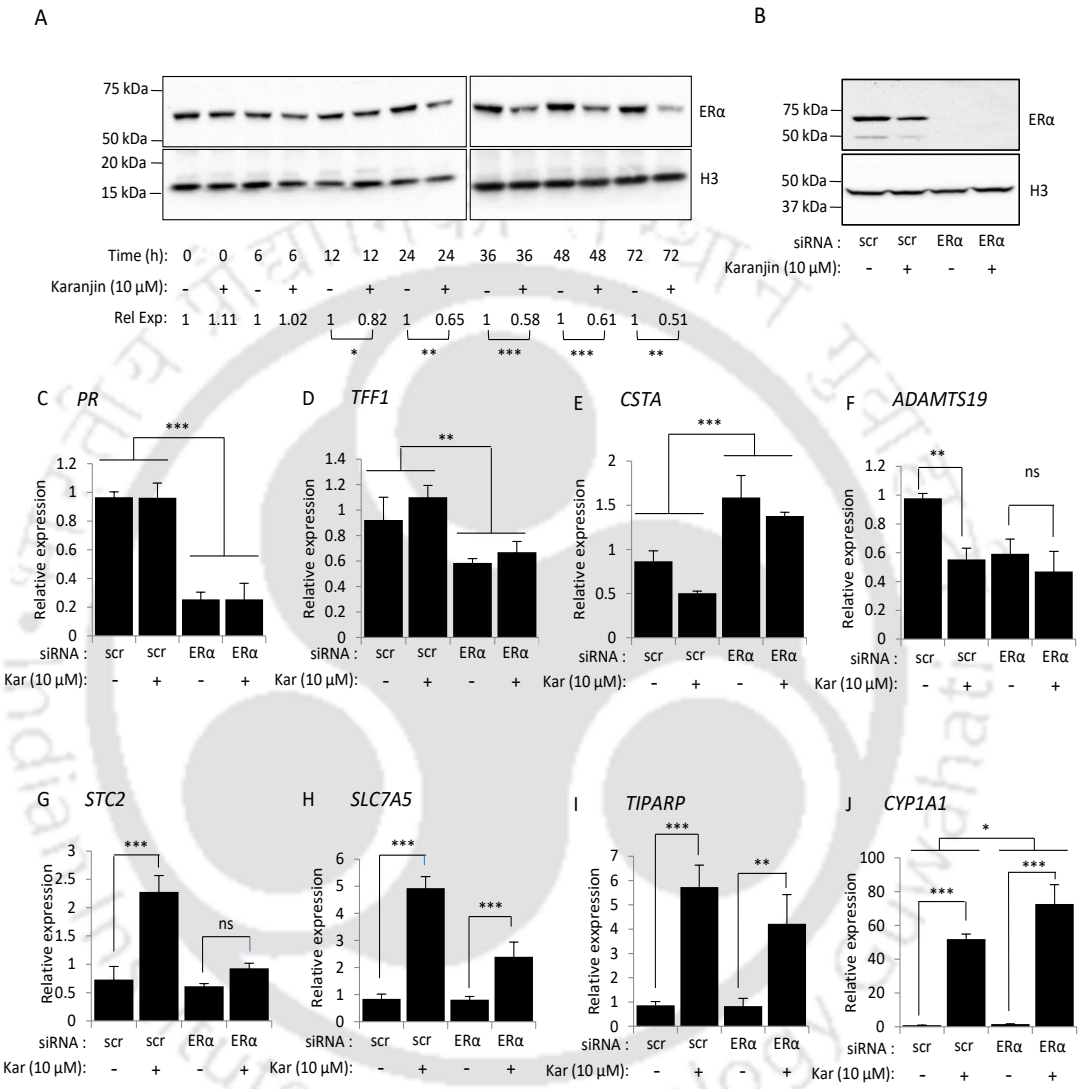


Fig. 6.6: Differential involvement of ERα in karanjin-mediated modulation of gene expression in MCF-7 cells. A. Karanjin-mediated reduction in ERα protein expression in MCF-7 cells. Cells were treated with 10 μM karanjin for the indicated periods of time. At each time point, the untreated cells (0 μM karanjin) served as control. Total protein extracted from Laemmli buffer was subjected to western blotting analysis using the ERα specific antibody. Histone H3 served as an internal control, which was probed with H3-specific antibody. Chemiluminescence signals were imaged in the ChemiDoc XRS+ system. The images were processed in ImageJ. The background-subtracted band intensity for ERα normalized against that obtained for Histone-H3 served as a measure of ERα protein expression. For each time point, ERα expression in the control sample was set to 1, and that obtained for the karanjin-treated sample was expressed relative to the control. The mean relative expression (Rel Exp) of ERα ±

sd (n = 5) for each time-point is shown below. For each time point, the data was analyzed by a one-tailed t-test. Significant results are indicated by asterisks. (*p < 0.05, **p < 0.01, ***p < 0.001). B. ERα knockdown. MCF-7 cells pre-treated with scrambled (scr) or ERα-specific siRNA were treated with vehicle or 10 µM karanjin (Kar) for a period of 24 h. Total protein was extracted from the phenolic fraction of RNA extraction reagent and subjected to western blot analysis using ERα and β-actin specific antibodies. β-actin served as an internal control. C-J. Total RNA was extracted from MCF-7 cells, which were treated as in B. The expression of the indicated genes was analyzed by qRT-PCR. The experiment was done with three replicate dishes for each experimental group. The expression in one control sample (scr + 0 µM karanjin) was set to 1, and those determined for the others were expressed relative to the control. Bars represent the mean relative expression ± sd (n = 3). For each gene, the data were analyzed by two-way ANOVA.

To gain further evidence if this ERα modulation is triggered in other ER-positive cell lines. We examined ERα protein levels and the effect of ERα knockdown in T47D cells. Like MCF-7, karanjin at 10 µM downregulated ERα protein in 24 h (Fig. 6.7A). For the knockdown study, T47D cells were treated with control (DMSO) or karanjin (10 µM) for 24h after pretreatment with scrambled or ERα -specific siRNA for 24 h. Interestingly, our RNA seq and qRT-PCR results in T47D cells revealed karanjin-modulated PR expression at the mRNA level. This induction of *PR* was abrogated upon ERα knockdown (Fig. 6.7B). The effect of ERα knockdown on karanjin-mediated modulation of PR is suggestive of the involvement of ERα in the regulation of PR, one of the hallmarks E2 target gene.

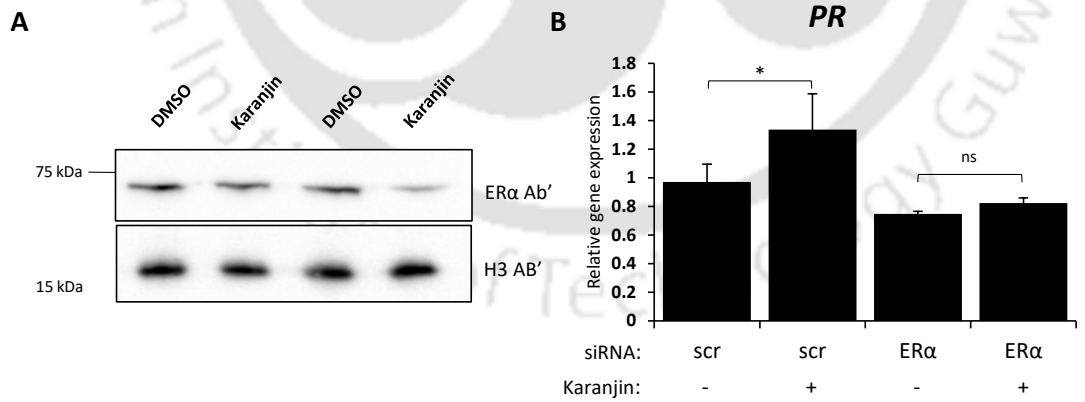


Fig. 6.7: Involvement of ERα in karanjin-mediated modulation of PR expression in T47D cells. A. Karanjin-mediated suppression in ERα protein expression in T47D cells. Cells were treated with DMSO or 10 µM karanjin for 24h. T47D cells pre-treated with scrambled (scr) or ERα-specific siRNA were treated with vehicle or karanjin (10 µM) for a period of 24 h. Total protein was extracted from the phenolic fraction of RNA extraction reagent and subjected to western blot analysis using ERα and H3-specific antibodies. H3 served as an internal control. Chemiluminescence signals were imaged in the

ChemiDoc XRS+ system. B. Total RNA was extracted from T47D cells, which were treated as in A. The expression of the indicated genes was analyzed by qRT-PCR. The experiment was done with three replicate dishes for each experimental group. The expression in one control sample (scr + 0 µM karanjin) was set to 1, and those determined for the others were expressed relative to the control. Bars represent the mean relative expression ± sd (n = 3). Asterisks indicate significant results. (*p < 0.05).

6.2.3 Karanjin treatment in MCF-7 cells increases ERα occupancy in the PS2, CSTA, HOXB2, CYP1A1, and TIPARP upstream region.

Research groups have demonstrated that estrogen regulates *PS2*, *CSTA*, *HOXB2*, *CYP1A1*, and *TIPARP* genes by means of engagement of ERα and interaction of ERα to the upstream regulatory region [144,168,183]. The examination of RNA sequencing data from MCF-7 cells, together with subsequent validation using qRT-PCR, revealed that the regulation of these genes is mediated by karanjin, and whether ERα is involved remains to be elucidated. We performed a ChIP assay to assess karanjin-mediated ERα binding to these E2 target genes independently. MCF-7 cells were treated with DMSO or 10 µM karanjin, and fixed chromatin fragments were immunoprecipitated with normal rabbit IgG or anti-ERα antibody. Immunoprecipitated DNA samples were subjected to PCR with primer pairs that amplify known sites of ERα occupancy in the *PS2*, *CSTA*, *HOXB2*, *CYP1A1*, and *TIPARP* locus.

Karanjin enhanced ERα occupancy in described E2 target genes. Thus, the ChIP experiments establish strong evidence that karanjin-mediated expression occurs via direct/ indirect recruitment of ERα to the upstream region of *PS2*, *CSTA*, *HOXB2*, *CYP1A1*, and *TIPARP* genes locus (Fig. 6.8).

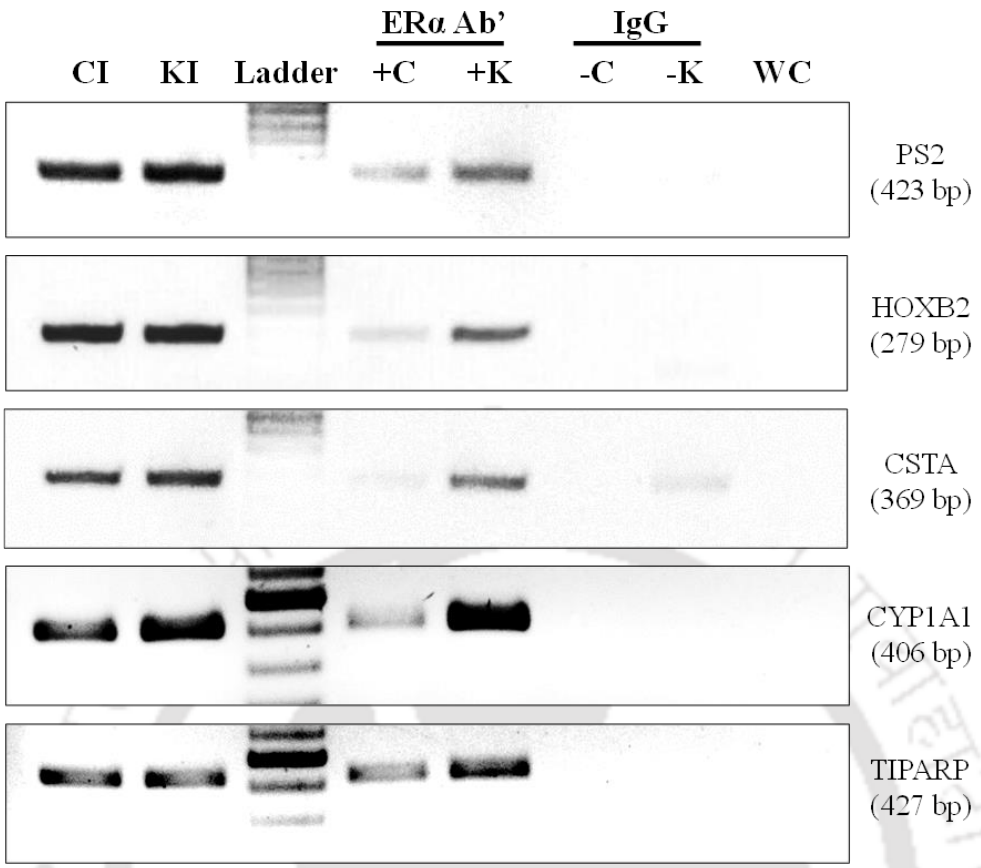


Fig. 6.8: Karanjin-induced binding of ERα in the upstream region *of PS2, HOXB2, CSTA, CYP1A1,* and *TIPARP.* ChIP analysis of MCF-7 cells with or without karanjin treatment. Sonicated chromatin samples were immunoprecipitated with ERα-specific or normal IgG antibody (negative control), and DNA was subjected to PCR using specific primers to amplify a specific region on the gene termed ChIP amplicon indicated as a black rectangle. CI and KI served as input samples, and WC as water control.

6.3 Discussion

Flavonoids have an intrinsic property to bind to nuclear receptors, elicit conformational changes, and recruit diverse coactivators or coregulators that modulate gene regulation. Hence, flavonoids are termed as natural selective estrogen receptor modulators (SERMs) [17,18]. Genistein, apigenin, quercetin, and soy isoflavones are some examples of flavonoids showing SERM-like behavior. SERMs are valuable therapeutic agents (e.g., tamoxifen, raloxifene) and have also served as tools to understand mechanisms of estrogen signaling and its effect on target cells and tissues. The actions of SERMs are known to be cell/tissue-dependent, or gene/promoter context-dependent, which makes them molecules of great utility [19]. An essential application of phytoestrogens is that they could be used as an alternative to the synthetic selective estrogen receptor modulators (SERMs), which exhibit estrogen agonist or

antagonist activity in a tissue-specific manner. Indeed, SERMs are used to treat estrogen-associated pathologies, such as breast cancer, brain diseases, osteoporosis, and menopausal symptoms. In other words, the challenge is to limit ER's mitogenic effects without compromising its protective properties, such as cell differentiation, neuroprotection, anti-osteoporosis effects, and antioxidant activities [184].

The current investigation was motivated by the paucity of evidence on the mechanistic understanding of partial estrogen-like effects instigated by karanjin. The findings call for an in-depth analysis of karanjin's effects on breast cancer cells and the mechanisms, particularly the involvement of ERα, before considering its usage in endocrine therapy [130]. Here, using an in-depth *in vitro* and *in silico* approach, we have predicted and explored the mechanistic evidence for the engagement of ERα in karanjin-mediated effects.

Docking helps rationalize the ligands' activity towards a target to perform structure-based drug design (SBDD). Docking assists in revealing novel compounds of therapeutic interest, forecasting ligand-protein interaction at a molecular basis, and delineating structure-activity relationships (SARs) [185,186]. Anti-estrogenic/estrogenic effect of flavonoids is attributed to their structural similarity with 17-β-estradiol (E2), due to which they can bind to the nuclear estrogen receptor (ER). On binding to ERs, they can either act as an agonist or antagonist. Estrogen-mimetic molecules show protective functions by interacting with ERα, such as lowering total serum cholesterol, reducing the incidence of cardiovascular diseases and osteoporosis, and neuroprotective effects. The helix 12 conformations of the ERα LBD play a vital role in determining whether a ligand is an agonist or antagonist. There is a change in helix 12 conformation due to the absence of an H-bond with His 524 when antagonist 4-Hydroxytamoxifen (4OHT) is bound [179,180]. The study of ligand-bound crystal structure plays a vital role in understanding the interaction mechanism through different structure and conformation changes induced by ligands leading to varied cellular response.

Molecular docking was used to investigate the atomic insight into the structure and dynamics of Estrogen Receptor (ERα) in complex with ligands: E2 and karanjin and quantify the binding energy. The Ligand binding pocket of ERα is hydrophobic. Thus, ligand binding is favorable by hydrophobic interaction between protein and

ligand but penalized by ligand desolvation. The desolvation penalty is higher for karanjin (more polar than E2), resulting in weaker binding energy. Non-polar substitutions in karanjin might result in better binding for ERα. Despite the structural similarity, E2 binding to ERα is favorable relative to karanjin. The furan ring oxygen of karanjin could form H-bond with His 524 compared to E2-bound His 524, Arg 394, and Glu 353 residues. Furan ring oxygen of karanjin imparts the hydrophilicity to the karanjin structure and thus could be involved in forming these polar contacts with the Ligand binding pocket of ERα LBD.

Additionally, when we analyze the agonist (E2) and antagonist (4-hydroxy Tamoxifen) structure, helix 12 conformation of ERα LBD plays a vital role in determining the agonist or antagonist activity of the ligand. This phenomenon occurred due to the absence of an H-bond with His 524 when 4OHT is bound [179,187]. Karanjin, when docked with 3ERT, antagonist 4OHT, and 1GWD antagonist Raloxifene bound structure, failed to form any H bonds with His 524, which indicates the agonist nature of karanjin.

In-silico results predict karanjin binds to ERα and induces the genomic changes incurred by karanjin. Further, to gain evidence on molecular insights into the karanjin-mediated regulation mechanism, ERα protein expression was probed. ERα protein expression in MCF-7 suggests a time-dependent inhibition of the protein. At later time periods, ERα protein levels were significantly suppressed, exhibiting a resemblance to the effects of estrogen. The protein expression study confirmed an enhanced ERα protein turnover following karanjin treatment in MCF-7 and T47D cells. Whether this protein degradation occurs due to proteasomal degradation or other transcriptional/post-transcriptional processes needs to be elucidated. The negative impact of ERα knockdown on karanjin-mediated alteration of gene expression provides enticing evidence in favor of ERα mediated actions of karanjin. In the absence of ERα, gene regulation of *PS2, CSTA, ADAMTS19, and STC2* was altered in MCF-7 and *PR* in T47D cells, suggesting the involvement of ERα in karanjin-mediated regulation. An increase in the occupancy of ERα and interaction of ERα with the upstream regulatory region of estrogen target genes *PS2, CSTA, HOXB2, TIPARP, and CYP1A1* further strengthens our proposition and corroborates with projecting karanjin as a partial estrogen-like molecule and SERM.

Türker and coworkers demonstrated that E2-activated ERα serves as a corepressor by directly interacting with the AhR-ARNT complex in the regulatory region of *CYP1A1* and *CYP1B1* [188]. In another research, estrogen receptor (ER) displayed TCDD- and time-dependent recruitment to the CYP1A1 promoter, boosted by E2 cotreatment. There is evidence that ER works as a coregulator of AhR-mediated transcriptional activation and that AhR recruitment of ER constitutes a new method of AhR-ER cross-talk [183]. Similarly, karanjin activated ERα might function as a coactivator by directly interacting with the AhR-ARNT complex at the regulatory region of *CYP1A1*.

In the previous chapters, we explored the genome-wide alteration at the mRNA level in ERα positive MCF-7 and T47D breast cancer cells using next-generation sequencing. The transcriptomic- or gene-modulatory footprint of karanjin encompasses a multitude of cellular processes. E2-response early and late pathway enrichment in T47D underscores the E2-like nature of karanjin. Not all genes are regulated in a similar pattern in both T47D and MCF-7. However, significant overlap yet differential modulation of genes in MCF-7 and T47D hint towards a SERM-like potential of karanjin. The partial estrogen-like or SERM-like activity of karanjin could underlie the concentration-dependent effect on MCF-7 cell proliferation. The compromised proliferation of MCF-7 and T47D treated with 10 μM karanjin in the face of induced expression of estrogen-response-early genes could be due to the additional effects of karanjin mediated via non-ERα targets. Karanjin is known to induce its effect partly via ERα; other nuclear receptors, such as AhR may also play a key role in karanjin-induced effects in breast cancer cells. Both ERα-positive and ERα-negative cells/tumors express the AhR. The AhR plays various functional roles in different breast cancer cell lines, exhibiting tumor promoter and tumor suppressor-like activities. Compelling evidence supports the existence of an inhibitory AhR-ERα interaction in which specific AhR ligands stimulate ERα degradation. At 10 μM and 50 μM karanjin, a decline in viable cell count might result from AhR-induced ERα degradation in ERα-positive cell lines but not in ERα-negative MDA-MB-231. Moreover, the induction of AhR downstream target genes *CYP1A1* and *CYP1B1* suggests the role of AhR in karanjin-mediated gene regulation. This aspect has been addressed in chapter 7 of the thesis.

Indian Institute of Technology Guwahati

CHAPTER 7

Karanjin modulates CYP1 family gene expression via AhR signaling axis and demonstrates AhR agonistic activity

Indian Institute of Technology Guwahati

7.1. Introduction

AhR belongs to the basic helix-loop-helix family of ligand activated transcription factors that contains a ligand-binding domain [AhR Per-AhR/Arnt-Sim (PAS)]. Upon activation, AhR mediates the transcription of downstream target genes in a cell type- and promoter context-dependent manner [189]. Among the many critical roles played by the AhR pathway, one of the most well-known is as a mediator of the toxicity of xenobiotics and environmental contaminants. In contrast, new research has uncovered additional AhR functions in the modulation of anti-viral immunity, homeostasis, and the inhibition of the adaptive immune response by several environmental toxins [190,191]. Recent research has further highlighted the AhR's physiological significance in controlling a wide range of cellular processes, including cell proliferation, the cell cycle, pluripotency, and stemness, as well as its functional relationships with numerous signaling pathways (including the ER) [102,108,192]. Several oncogenic cell lines have demonstrated that AhR agonists have a potent and selective growth inhibitory effect. Two AhR agonists, amino flavone (AF) and the benzothiazole prodrug Phortress®, have progressed through Phase I clinical trials and shown selectivity for the ER-positive MCF-7 and ineffectiveness against ER-negative cell lines [123]. Thus, novel agonists derived from natural sources may play a key role in cancer prevention.

Previously, we described that karanjin alters the expression of certain ERα target genes in a concentration- and cell-type-dependent manner, thereby establishing its role as a naturally derived estrogen-like regulator. The global transcriptome analysis of MCF-7 and T47D breast cancer cells treated with karanjin revealed that *CYP1A1* and *CYP1B1* were among the top significantly expressed genes, suggesting the involvement of AhR [130]. The present chapter describes the role of AhR in karanjin-mediated effects. We performed gene expression studies on AhR target genes in ER-positive cells to better understand the karanjin-mediated regulation of AhR. The three basic strategies were utilized: a) *in silico* approach to predict AhR karanjin binding, b) blocking AhR by means of AhR antagonist and AhR specific siRNA, and c) Indirect evidence of karanjin binding to the promoter region of AhR using ChIP assay. Interestingly, we concluded that karanjin regulated AhR downstream target genes via interaction with AhR and activation of the AhR signaling pathway.

7.2. Results

7.2.1 In silico molecular docking to predict the binding of karanjin to AhR Ligand binding domain

To computationally predict if karanjin behaves as an AhR ligand, rigid molecular docking was conducted through Autodock 4.2.6. The interaction between karanjin and AhR ligand binding domain (PAS B) was studied. Karanjin binds in the same ligand binding pocket as it is known ligand 2,3,7,8-Tetrachlorodibenzodioxin (TCDD). The binding energy score and H-bonds contribution of TCDD and karanjin are represented in Figure 7.1A & B. Karanjin forms H-bonds with Thr 289 and His 291 (Fig. 7.1A & 7.1C). Overall, ~100 ns of MD gives structural convergence (RMSD <2Å, Appendix 4) of the karanjin/TCDD-AhR complexes (Fig. 7.1C & D).

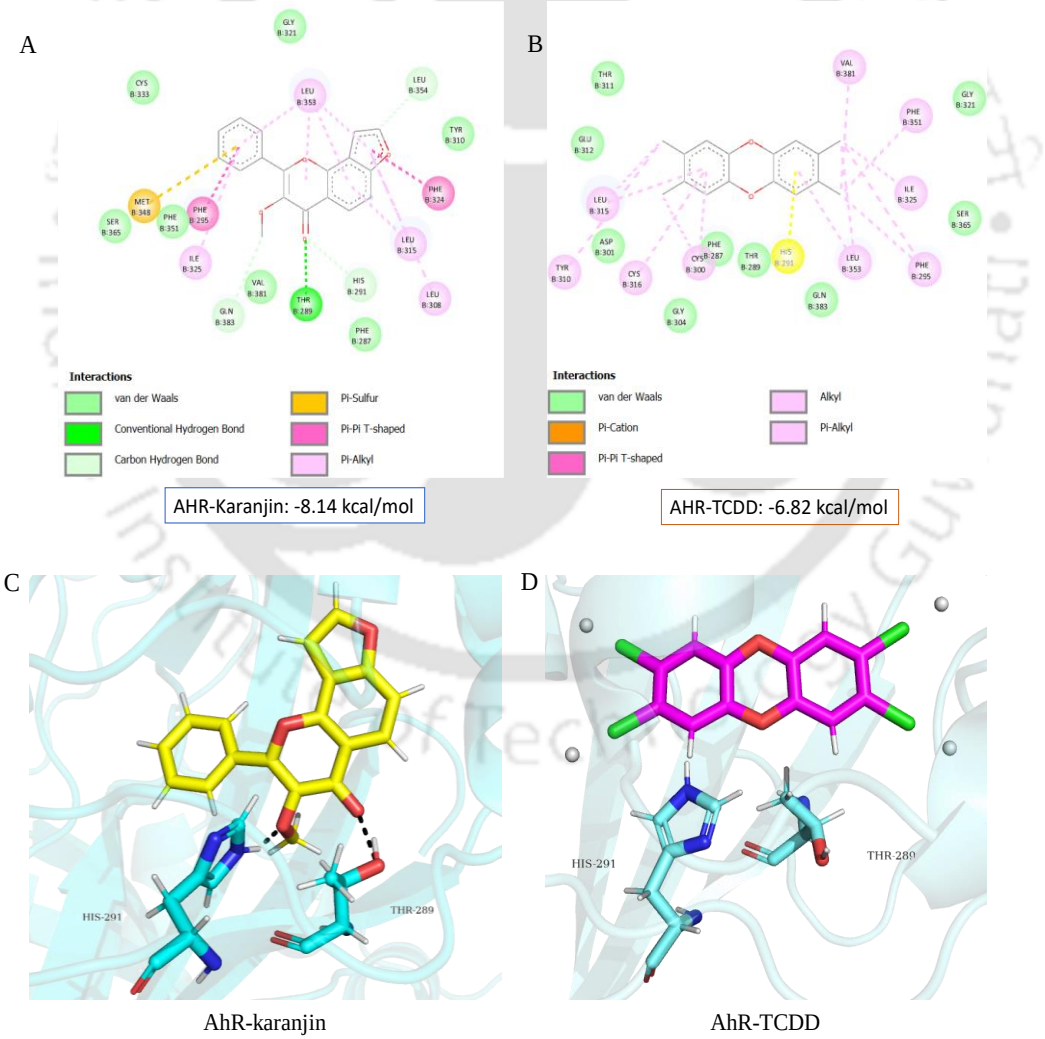


Fig. 7.1: Docking and Molecular Dynamics simulation of AhR Ligand binding domain with karanjin. A & B Discovery Studios derived a depiction of the conformation of karanjin and TCDD

bound to the AhR-LBD structure. C & D Molecular Dynamics simulation (100 ns) of AhR-karanjin and AhR-TCDD complex visualized in PyMOL.

The RMSD and RMSF were predicted for the AhR protein alone (black color) and bound to TCDD (red) or karanjin (blue). RMSF plot depicts the fluctuations that were reduced upon binding with TCDD or karanjin, compared to the protein alone. Similarly, the RMSD plot depicted TCDD or karanjin bound AhR demonstrated structural convergence reached within 20 ns, much earlier than the protein alone (Fig. 7.2). Molecular modeling analysis predicts karanjin interacts with AhR in a stable conformation, and binding energy was comparable to TCDD bound AhR complex, and we can predict karanjin to be an AhR ligand.

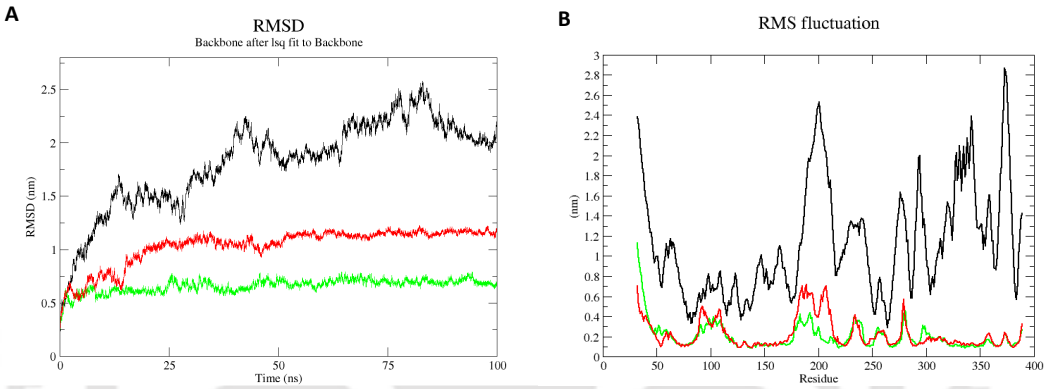


Fig. 7.2: RMSD and RMSF plot of AhR-karanjin and AhR-TCDD and AhR alone. RMSD and RMS fluctuation of MDS (100 ns) derived depiction of conformation of AhR (Black) or TCDD (Green) and karanjin (Red) bound to AhR.

7.2.2 AhR involvement in karanjin mediated gene regulation

Among the plethora of genes altered by karanjin in MCF-7 cells, global transcriptome profiling identified *CYP1A1* and *CYP1B1* as two of the most significantly up-regulated genes. Similar inductions in *CYP1A1* and *CYP1B1* gene expression were observed in T47D, yet another ERα positive breast cancer cell line. We also established that other AhR target genes, *CYP1A2* and *ABCG2*, were upregulated in both MCF-7 and T47D cells. *CYP1A1*, *CYP1B1*, *CYP1A2, and ABCG2* are the known downstream target genes for AhR signaling, and their regulation by karanjin suggests a possible AhR regulatory or agonist role for karanjin (Fig. 7.3). Moreover, our cell viability result and gene expression of AhR target genes follow a pattern where ER-positive cells have reduced viability and AhR target genes are induced.

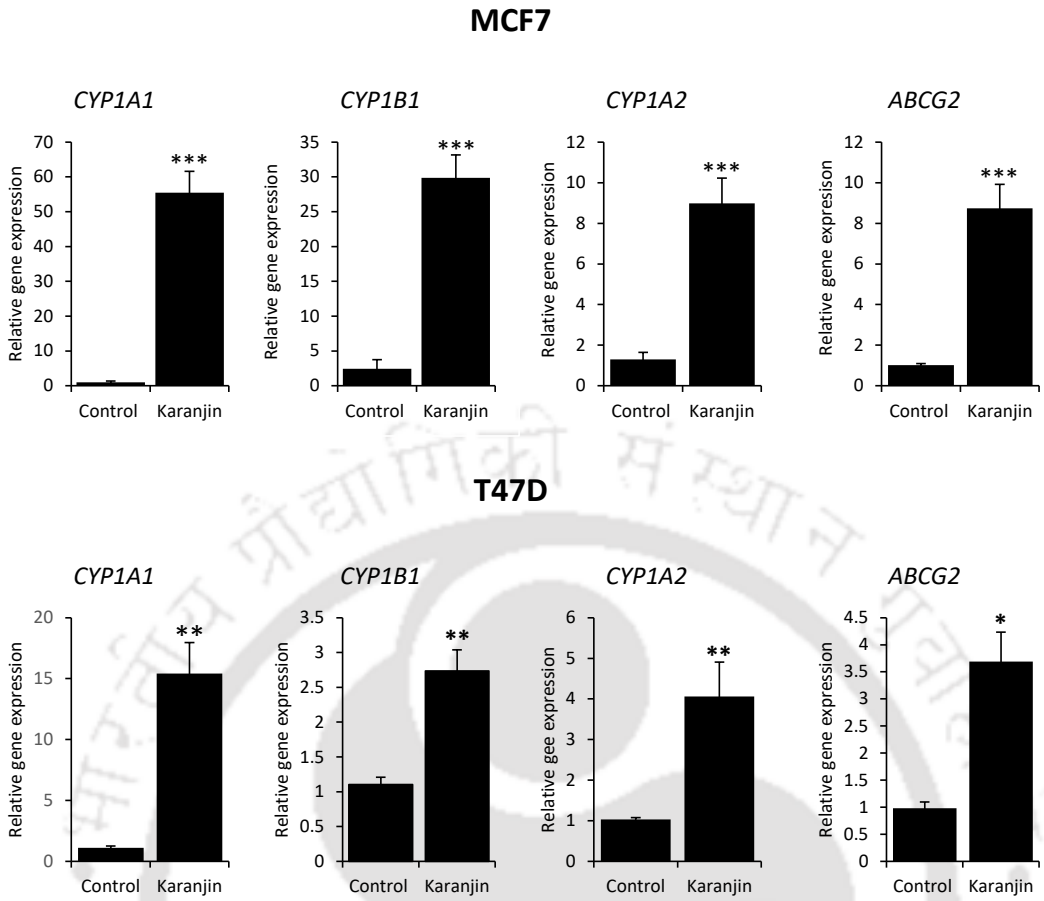


Fig. 7.3: Karanjin mediated regulation of *CYP1A1, CYP1B1, CYP1B1 and ABCG2*. MCF-7, and T47D cells were treated with vehicle (DMSO) or 10 µM karanjin for 24 h. Total RNA was isolated for analyzing gene expression in MCF-7, and T47D cells using qRT-PCR. Gene expression in vehicle-treated (control) cells and karanjin (10 µM) was expressed relative to control. Bars represent mean relative expression ± sd (n=3). The data were analysed by a one-tailed *t*-test. Asterisks (*p <0.05, ** p<0.01, ***p<0.001, ****p<0.0001) denote significant results.

7.2.3 Effect of karanjin on AhR protein expression.

AhR is a ligand-activated transcription factor that enables cells to adapt to changing conditions by sensing compounds from the environment, diet, microbiome, and cellular metabolism, and it plays important roles in development, immunity, and cancer. Xenobiotics can act as ligands: upon xenobiotic binding, they activate the expression of multiple phase I and II xenobiotic chemical metabolizing enzyme genes (such as the *CYP1A1* gene). Next to xenobiotics, natural ligands derived from plants, microbiota, and endogenous metabolism are potent AhR agonists.

In MCF-7 cells, AhR protein expression was repressed by 10 µM karanjin in 24 h. In T47D, there was a pronounced expression of the AhR protein when observed at a 24

h time point (Figure 7.4). This difference in AhR expression could be cell context- or promoter context-dependent due to the recruitment of differential coactivators, co-repressors, and binding partners. In addition, our examination of MCF-7 transcriptome data revealed an induction of the *AhRR* gene, which could represent a feedback mechanism suppressing the AhR protein in MCF-7 cells but not in T47D cells.

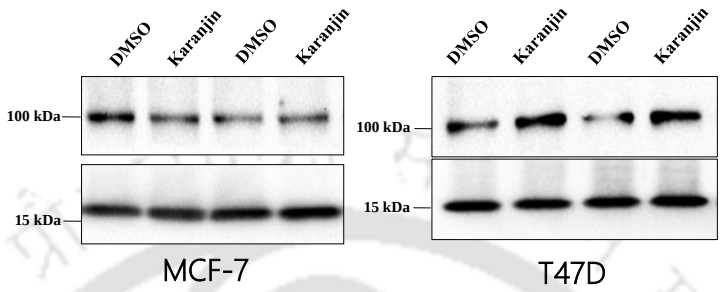


Fig. 7.4: Karanjin mediated differential expression of AhR protein in MCF-7 and T47D cells. Cells were treated with 10 µM karanjin for 24 h. The untreated cells DMSO (0 µM karanjin) served as control. Total protein extracted from Laemmli buffer was subjected to western blotting analysis using the AhR-specific antibody. Histone H3 served as an internal control. Chemiluminescence signals were imaged in ChemiDoc XRS+ system.

7.2.4 Effect of AhR antagonist on karanjin-mediated regulation AhR target genes.

In order to examine the influence of AhR on the regulation of AhR target genes, we employed CH-223191, an AhR antagonist, in MCF-7 and T47D cell lines. CH-223191 inhibits the translocation of AhR to the nucleus, thus inhibiting the regulation of downstream target genes. In the presence of CH-223191, the karanjin-modulated gene regulation was abrogated for AhR target gene sets (Fig. 7.5). The data were analyzed by two-way ANOVA to test the main effects of karanjin treatment, AhR antagonist, or their interaction. There were significant main effects of karanjin treatment, AhR knockdown, and interaction in all three AhR target genes in both cell lines. This suggests that karanjin-mediated *CYP1A1, CYP1A2, and CYP1B1* regulation is under the regulatory control of AhR.

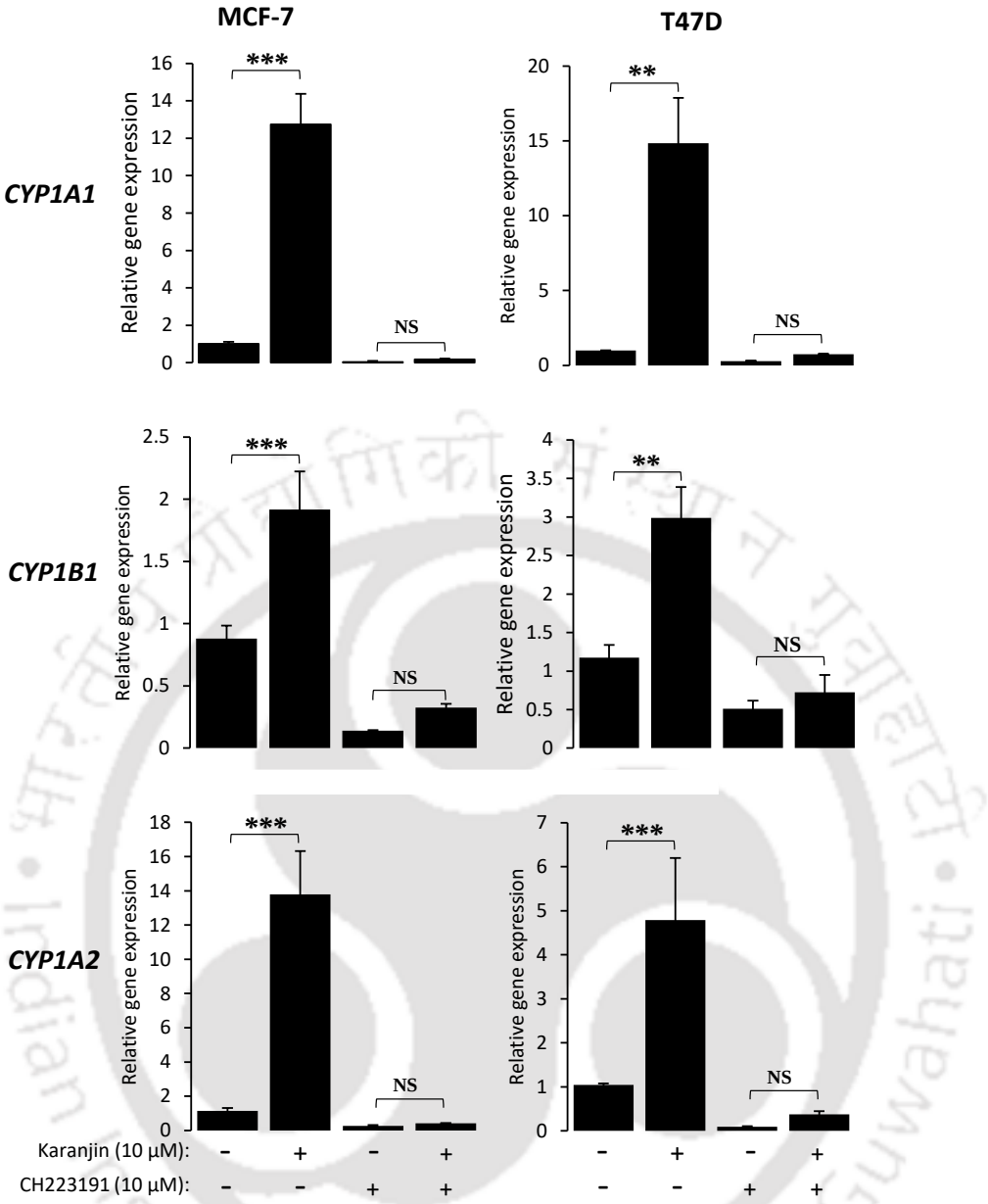


Fig. 7.5: AhR antagonist blocks karanjin-mediated regulation of *CYP1A1, CYP1B1, and CYP1A2*. MCF-7, and T47D cells were incubated in M2 media for 24 h followed by co-treatment of vehicle or 10 µM karanjin with CH-223191 for 24 h. Total RNA were isolated for analyzing *CYP1A1, CYP1B1, and CYP1A2* expression in MCF-7, and T47D cells expression using qRT-PCR. Bars represent the mean relative expression ± sd (n=3). The data were analyzed by two-way ANOVA followed by Tukey's HSD. Asterisks (*p <0.05, ** p<0.01, ***p<0.001, ****p<0.0001) denote significant results.

7.2.5 Effect of AhR knockdown on karanjin-mediated regulation of CYP1A1, CYP1B1, and CYP1A2 in MCF-7 cells

Blocking of AhR had an influence on the *CYP1A1*, *CYP1B1*, and *CYP1A2* gene expression. Karanjin-induced gene regulation was abrogated in the presence of AhR

antagonist, CH223191. To confirm the role of karanjin in modulating the AhR-CYP1A1 signaling axis, we depleted cells of AhR by knockdown; using specific siRNAs. AhR knockdown had an impact on the gene regulatability of *CYP1A1*, *CYP1B1,* and *CYP1A2* (Fig. 7.6). The data were analysed by two-way ANOVA to test the main effects of karanjin treatment, AhR knockdown, or their interaction. There were significant main effects of karanjin treatment, AhR knockdown, and interaction in all three AhR target genes. Based on our findings, we conclude that karanjin regulates AhR downstream target genes via activating the AhR signaling axis.

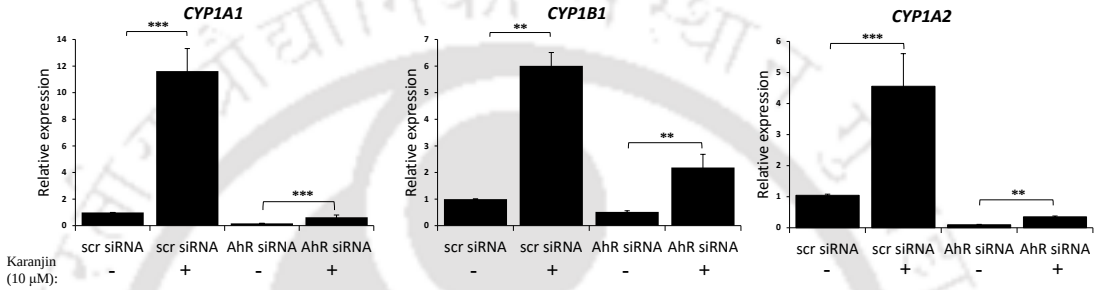


Fig. 7.6: Effect of AhR knockdown on karanjin-mediated regulation of *CYP1A1*, *CYP1B1,* and *CYP1A2* in MCF-7 cells. MCF-7 cells were transfected with scrambled siRNA (scr) or AhR-specific siRNA and incubated for 24 h followed by treatment with vehicle or 10 µM karanjin for 24 h. Total RNA were isolated for analyzing *CYP1A1* and *CYP1B1* expression using qRT-PCR. *CYP1A1*, *CYP1B1, and CYP1A2* expression in scr + vehicle-treated (control) and those in other treatment groups were expressed relative to the control. Bars represent mean relative expression ± sd (n=3). The data were analyzed by two-way ANOVA followed by Tukey's HSD. Asterisks (*p <0.05, ** p<0.01, ***p<0.001, ****p<0.0001) denote significant results.

7.2.6 Examination of the recruitment of AhR by karanjin to XRE-binding sites in AhR target gene CYP1A1

Liganded AhR modulates transcriptional regulation of gene expression by engaging with xenobiotic-response elements (XRE) in target gene promoters [15]. Previously, we described *CYP1A1* is highly induced by karanjin and blocking AhR utilising antagonist or siRNA abrogated its expression. These observations suggest AhR involvement in regulating *CYP1A1* gene expression. In this study, we examined the potential effects of karanjin therapy on the activation of the AhR and its subsequent engagement to the XRE-region within the CYP1A1 gene locus. Karanjin enhanced the recruitment of AhR to the XRE-binding site in the *CYP1A1* gene locus in MCF-7 cells when exposed to 10 µM karanjin for 24 h (Fig. 7.7). This observation provides strong

evidence that karanjin-mediated *CYP1A1* expression occurs via direct/indirect engagement of AhR to the XRE-region of the *CYP1A1* locus.

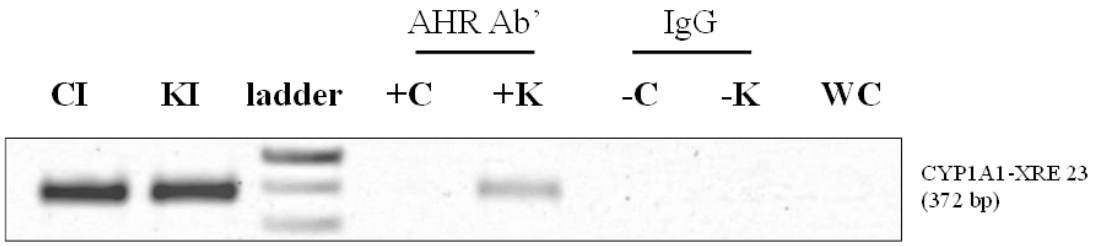


Fig. 7.7: Karanjin-induced binding of AhR in the upstream region of *CYP1A1* in MCF-7 cells. ChIP analysis of MCF-7 cells with or without karanjin treatment was performed. Sonicated chromatin samples were immunoprecipitated with AhR-specific or normal IgG antibody (negative control) as described in Materials and Methods. Immunoprecipitated DNA was subjected to PCR using primers to amplify the region shown as a black rectangle (ChIP amplicon). CI and KI are control and karanjin-treated input samples, respectively. WC stands for water control.

7.3 Discussion

Various *in vitro* and *in vivo* studies have reported karanjin pharmacological and biological effects. Karanjin is involved in glucose transport, drug metabolism, and drug efflux [74,79,193]. Recently, Joshi and co-workers reported that karanjin inhibited CYP1A1 and other CYP1 family enzymes when overexpressed in recombinant HEK-293 cells and in an EROD assay [74]. They demonstrated that karanjin is a competitive inhibitor of CYP1A1 and acts as a pan-inhibitor of CYP1 family enzymes. However, in ER-positive MCF-7 and T47D breast cancer cells, all these AhR-downstream target genes (which are under the regulatory control of AhR) were significantly upregulated when exposed to karanjin. Our cell viability results lead to an interesting observation that ER-positive and ER-negative cell lines behave differently upon karanjin treatment. Beta-naphthoflavone (βNF), an AhR agonist and a putative chemotherapeutic agent that has antitumor activity against mammary carcinomas, inhibits tumor growth in ER-positive (MCF-7) but not ER-negative (MDA-MB-231) breast cell lines through AhR-dependent regulation of PI3K/AKT and MAPK/ERK signaling [105]. They demonstrated G0/G1 phase arrest in the MCF-7 cells, which led to cellular senescence through AhR-dependent regulation of PI3K/AKT [11]. We reported similar observations while analyzing viable cell count using trypan blue assay, which hinted towards delayed cell cycle progression in ER-positive (MCF-7 and T47D cells) but not in ER-negative (MDA-MB-231) cells when administered karanjin at 10 µM and 50 µM.

Moreover, karanjin at 10 µM induced the CYP1 family of AhR target gene expression in MCF-7 and T47D cells via activating AhR. Hence, we postulate that this regulation of the PI3K/AKT signaling pathway by karanjin in breast cancer cells may result from AhR activation, which needs to be further elucidated. Additionally, specific AhR ligands promote ERα degradation, lending credence to the presence of an inhibitory AhR-ERα interaction. Viable cell count reductions at 10 µM and 50 µM karanjin may be attributed to AhR-induced ERα degradation in ER-positive but not in ER-negative cell lines. Our findings on ERα protein expression in MCF-7 and T47D further corroborate the hypothesis that ERα protein degradation is time-dependent, with significantly repressed ERα protein levels at later time points. Moreover, the chromatin immunoprecipitation result demonstrates enhanced occupancy of ERα and AhR in the *CYP1A1* promoter region in MCF-7. Collectively, our results indicate that the ERα-AhR-PI3K/AKT signaling axis plays a coordinated role in regulating the expression of genes and the cell-cycle fate of breast cancer cells.

More than 1.7 million new breast cancer cases are diagnosed yearly, making it the most frequent disease among women. AhR, often involved in the metabolism of xenobiotic ligands, is now considered a potential biological target for treating this fatal illness. Two important AhR agonists, amino flavone (AF) and the benzothiazole prodrug Phortress®, have progressed through Phase I clinical trials and shown selectivity for MCF-7 while being effective against ER-negative cell lines [123]. Karanjin-mediated differential regulation of AhR protein in MCF-7 and T47D might be attributed to the AhR agonism and feedback mechanism via the AhR repressor, AhRR. The active AhRR binds to ARNT (AhR nuclear translocator) competitively and inhibits the AHR's action creating a negative feedback loop [123]. In the case of MCF-7, when we analyzed the RNA-seq data, AhRR was upregulated (Fold Change-2) [130]. However, AhRR was unregulated in T47D, which could explain differential AhR protein expression in these cell lines apart from cell and promoter context-dependent changes.

A molecular modeling study was conducted to predict karanjin binding to AhR. Since the AhR crystal structure is missing, homology models of the AhR ligand-binding region have been created. The hypoxia-inducible factor 2 (HIF-2) PAS B domain has been used as the tertiary structure as its NMR and X-ray crystallography-derived three-

dimensional structure is available. Significant structural overlap exists between the AhR and HIF-2, and the structure of the ligand binding domain is highly conserved among members of this family [194]. Docking and molecular dynamics simulation of 100 ns gave structural convergence with TCDD-bound AhR, depicting the agonist nature of karanjin. Karanjin was bound in the same ligand binding pocket as TCDD, forming polar interactions with His 291 and Ser 365 residues are predicted to constitute part of the binding pocket and have been shown to be critical for agonist binding [195]. Hence, *in silico* data interpretation predicts AhR agonist nature of karanjin.

The involvement of AhR in the karanjin-mediated induction of xenobiotic metabolism genes (*CYP1A1, CYP1A2,* and *CYP1B1*) was evaluated using AhR antagonist CH223191 or AhR siRNA. The AhR antagonist CH223191 inhibits the ability of AhR agonists (2,3,7,8-tetrachlorodibenzo-p-dioxin and related HAHs), to bind to and/or activate the AhR, and AhR signal transduction [196]. Karanjin-dependent induction of *CYP1A1, CYP1A2,* and *CYP1B1* was abrogated in the presence of AhR antagonist. Similarly, AhR-specific siRNA significantly blocked karanjin-mediated up-regulation of *CYP1A1, CYP1A2, and CYP1B1* in MCF-7. Importantly, blocking AhR suppressed the *CYP1A1, CYP1A2, and CYP1B1* mRNA levels to the basal level, and in the presence of CH223191 or AhR-siRNA, karanjin mediated regulation was compromised. The enhanced AhR occupancy in the XRE-region following karanjin treatment in MCF-7 cells suggests that karanjin induces *CYP1A1* expression at the level of transcription and provides indirect evidence for karanjin-mediated AhR activation. ChIP data present enticing evidence to prove direct/indirect physical engagement of AhR to the *CYP1A1* promoter locus upon karanjin stimulus. The precise events post AhR binding that lead to transcriptional activation are worth addressing in future investigations.

Taken together, *in vitro* and *in silico* data indicate that karanjin may activate the AhR receptor and be positioned as a novel class of AhR agonist or selective AhR modulator. These findings identify karanjin as a novel AhR ligand that exerts an effect akin to dioxin but does not elicit an undesirable toxic effect. Further research targeting the potential AhR regulatory property of karanjin and associated gene modulatory effects may prove valuable for therapeutic intervention in targeted breast cancer therapy. Karanjin or rationally derivatized candidates from karanjin in combination with a known breast cancer drug can serve as an alternative to current therapy.

CHAPTER 8

Overall Conclusion and Future Scopes

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

8.1 Conclusion

Karanjin, the most abundant furanoflavone from *P. pinnata* seed oil, is highly revered in Ayurveda and traditional system of medicine to treat various ailments. Karanjin is known to modulate glucose transport, insulin signaling, PI3K-AKT, and NF-κB signaling, and is involved in drug metabolism and efflux. Recent investigations revealed anticancer and chemopreventive potential of karanjin in diverse cell lines with varying IC_{50} concentrations.

Despite the accumulation of knowledge regarding the diverse biological effects associated with karanjin, there is a scarcity of evidence pertaining to the genomic impacts induced by karanjin. The present study employed next-generation sequencing to unveil the genome-wide gene expression alterations at the mRNA level in ERα positive MCF-7 and T47D cells. It fosters molecular signatures and deeper insights into the knowledge of gene regulation exhibited by karanjin. The transcriptomic- or gene-modulatory footprint of karanjin encompasses a multitude of cellular processes, such as metabolism (glycolysis, fatty acid, and xenobiotic), signaling (TNFα, KRAS), ROS-scavenging, unfolded protein response, modulation of transcription factor targets (E2F and MYC), and hormonal response (androgen, and estrogen-response early). Thus, suggesting the role of karanjin in various biological processes and the prevention of various diseases and disorders.

Over time, our understanding of the anticancer role of karanjin has grown. However, the chemopreventive potential of karanjin is only briefly described in the literature. Various research groups have previously demonstrated *in vitro* effects of karanjin over a wide concentration range, but not in hormone-responsive breast cancer cells. Roy and co-workers have reported karanjin-mediated G2/M arrest, and apoptosis in HeLa cells in a concentration-dependent manner with an IC_{50} of 160 µM [9,10]. In colon cancer cells, karanjin treatment reduced BAX, p53, and PCNA levels and increased Bcl2 expression, thereby ameliorating colon carcinogenesis [197]. Karanjin exerted antitumor activity in MDA-MB-231 breast cancer cells by regulating the PI3K/Akt signaling pathway [11]. Moreover, the inhibitory effect that karanjin has on cell development is not surprising, considering the flavonoid nature of the compound. Surprisingly, we found no evidence for any short-term (24 h) effect of varying concentrations of karanjin on ER-positive and-negative breast cancer viable cell count.

However, the long-term (120 h) effect was significantly concentration-dependent, although the viable cell count increased at all concentrations of karanjin. The viable count was significantly lower at higher concentrations (10 and 50 μM), whereas it was, modestly, but significantly higher at lower concentrations (10 nM). These data suggest that the anti-proliferative effect of karanjin may not be universal, but rather be cell- and concentration-dependent. This presents a caveat to the anticancer potential of karanjin, and underscores the importance of dosage for various healthcare applications suggested in the literature. The concentration-dependent effect indicates more than one receptor for karanjin in breast cancer cells. The high-affinity receptors could be responsible for the proliferative actions of karanjin at lower concentrations. In contrast, low-affinity receptors could be responsible for cell cycle arrest at high concentrations.

This study unveils karanjin's previously unknown proliferative effect, characterized by its E2-like potential due to its flavone backbone. Enrichment of estrogen-response-early genes in both MCF-7 and T47D and overlap of karanjin-modulated genes with those regulated by 1 nM E2 or 1 μM tamoxifen suggest partial estrogen-like effects of karanjin. Modulation of estrogen target gene both in MCF-7 and T47D cells, enhanced ERα protein turnover following karanjin treatment, and the negative impact of ERα knockdown on karanjin-mediated alteration of gene expression provide enticing evidence in favor of ERα mediated actions of karanjin. An increase in the occupancy of ERα and interaction of ERα to the upstream regulatory region of estrogen target genes *PS2, CSTA, HOXB2, TIPARP,* and *CYP1A1* further strengthens our proposition and collaborates to projecting karanjin as a partial estrogen-like molecule and SERM-like potential. The partial estrogen-like or SERM-like activity of karanjin could underlie the concentration-dependent effect on MCF-7 and T47D cell proliferation. The compromised proliferation of MCF-7 and T47D cells treated with 10 μM karanjin in the face of induced expression of estrogen-response-early genes could be due to the additional effects of karanjin mediated via non-ERα targets.

Moreover, the enrichment of metabolism-related genes suggests involvement of AhR. Notably, *CYP1A1* and *CYP1B1* (classical downstream target genes for AhR) were among the top up-regulated genes in RNA seq analysis of MCF-7 and T47D cells. This study attempts to understand the possible regulation of AhR by karanjin using detailed *in vitro* and *in silico* approaches. For the first time, we identified karanjin as a novel AhR agonist that exerts an effect akin to dioxin but does not elicit an undesirable toxic

impact. Moreover, recently AhR targeting ligands have been proposed to be exciting prospects for breast cancer prevention, and one such AhR agonist, aminoflavone, has progressed to phase 1 clinical trial [123].

Thus, the present study provides valuable insight into the hitherto unexplored effects of karanjin on endocrine-responsive cells. The caveat exposed by the data concerning the anticancer potential of karanjin cannot be overlooked. We predict ERα-AhR-PI3K/Akt axis as a possible mechanism of delayed cell cycle progression in ERα positive MCF-7 and T47D cells but not in MDA-MB-231. Interestingly, ERα and AhR serve as important drug targets for cancer chemotherapy. Karanjin modulates both ERα and AhR and regulates its downstream target genes.

The present investigation pivots on the unexplored mechanistic understanding of ERα and AhR signaling in mediating the effects produced by karanjin. The current study identified the karanjin-induced engagement of ERα and AhR in the respective ERE and XRE regions of target genes. ChIP and knockdown results of AhR downstream target genes, *CYP1A1* and *TIPARP*, unveiled the possible interaction of the two pathways and suggested that their induction is under the regulatory control of both AhR and ERα. These findings suggest karanjin as a novel SERM and SAhRM, which holds potential as an essential therapeutic molecule. Furthermore, in-depth *in vivo* and clinical investigation awaits, considering karanjin itself or its modified derivatives that can serve as an alternative to the current estrogen replacement therapy or cancer chemoprevention.

8.2 Future scope

The present investigation sheds new light on the involvement of ERα and AhR in karanjin-mediated effects produced in breast cancer. Our contributions are noteworthy in elucidating the genomic correlates and signatures altered by karanjin. Utilizing next-generation sequencing and in-depth *in silico* and *in vitro* techniques, our study provides a possible mechanistic understanding of different molecular and physiological effects produced by karanjin. The ERα and AhR modulatory effect of karanjin holds promise in the development of new alternatives to endocrine therapy and chemopreventive drugs. Further studies are needed to carefully understand the mechanisms underlying these effects and evaluate their potential clinical applications. The following studies listed below are envisaged for utilizing karanjin in the development of new therapeutic approaches for cancer and other diseases.

- The RNA-seq results in MCF-7 and T47D cells will contribute to a better understanding of the effects of karanjin simulation in breast cancer cells. Investigators interested in karanjin and its biological effects can use, and independently analyze, the RNA-seq data for insights into the genomic effects.
- Karanjin, like other flavonoids, demonstrates dual nature in breast cancer cells. At lower concentrations have a proliferative whereas higher concentrations have an antiproliferative effect. The dose-dependent effect of karanjin needs further consideration, and karanjin action on other cell types and *in vivo* anticancer models at the physiological range must be rigorously tested to validate its anticancer potential.
- The observed E2-like property of karanjin caters it eligible as a putative alternative in Hormone Replacement Therapy (HRT). Moreover, whether karanjin influences improving bone density and mineralization needs further scrutiny. Karanjin's potential to yield avenues to develop alternatives to the current estrogen replacement therapy.
- It remains to be explored whether karanjin physically binds ERα, and modulates its transactivation function in a concentration-, gene-, or cell-type-dependent manner as exhibited by selective estrogen receptor modulators. A biophysical approach using ITC, Surface plasmon resonance (SPR), and Co-Crystallization approach can provide insights into the binding affinity, stoichiometry, and

thermodynamics of karanjin-ERα interaction, as well as the structural basis of their complex formation. Such information can inform the design of more potent and selective ER modulators for breast cancer therapy.

- In traditional Ayurvedic and Siddha medicine, Karanj oil has been used to treat various skin ailments such as Psoriasis, Acne, Eczema, Keloid tumors, and wound healing. Karanj oil is rich in essential fatty acids, including oleic acid, linoleic acid, palmitic acid, flavonoids, tannins, and other phytochemicals. The antibacterial and anti-inflammatory properties of Karanj oil make it an effective treatment for wounds. It can help to reduce inflammation, prevent infection, and promote wound healing. Hence, karanjin being the most abundant active metabolite from seed oil, hold the potential to be examined for its effect in skin related conditions.
- The study also revealed that karanjin treatment altered the expression of genes involved in drug metabolism and efflux, which may have important implications for its use in combination with other drugs. Karanjin sensitized neuroblastoma cells ABCG2 substrate Mitoxantrone, used to treat advanced-stage breast cancer. Karanjin-induces ABCG2 levels in MCF-7 and T47D breast cancer cells, and cotreatment with Mitoxantrone could hold promise in further breast cancer prevention, which needs to be investigated.
- Our results establish karanjin as a new AhR ligand or SAhRMs with an action similar to dioxin but without the latter's detrimental effects. Further investigation awaits into karanjin's putative AhR binding property and related gene modulatory effects that might lead to therapeutic intervention in targeted breast cancer treatment.
- Both ERα and AhR play crucial roles in determining the pronounced effect elicited by karanjin. Whether there lies a crosstalk or correlation between ERα and AhR still remains unaddressed and can open new avenues of research in breast cancer prevention.

Indian Institute of Technology Guwahati

References

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

[1] A. Singh, G. Bhatt, N. Gujre, S. Mitra, R. Swaminathan, A.M. Limaye, L. Rangan, Karanjin, Phytochemistry. 183 (2021) 112641. https://doi.org/10.1016/j.phytochem.2020.112641.

[2] S. Sangwan, D.V. Rao, R.A. Sharma, A review on *Pongamia pinnata* (L.) Pierre: a great versatile leguminous plant, Nat Sci. 8 (2010) 130–139.

[3] L.M.R. Al Muqarrabun, N. Ahmat, S.A.S. Ruzaina, N.H. Ismail, I. Sahidin, Medicinal uses, phytochemistry and pharmacology of *Pongamia pinnata* (L.) Pierre: A review, J Ethnopharmacol. 150 (2013) 395–420. https://doi.org/10.1016/j.jep.2013.08.041.

[4] A.A.M. Noor, S.N.N. Othman, P.T. Lum, S. Mani, M. Farooq Shaikh, M. Sekar, Molecules of interest - Karanjin - A review, Pharmacognosy Journal. 12 (2020) 938–945. https://doi.org/10.5530/pj.2020.12.133.

[5] A. Singh, I. Jahan, M. Sharma, L. Rangan, A. Khare, A. Panda, Structural characterization, *in silico* studies and *in vitro* antibacterial evaluation of a furanoflavonoid from karanj, Planta Medica Letters. 3 (2016) e91–e95. https://doi.org/10.1055/s-0042-105159.

[6] Akanksha, A.K. Srivastava, R. Maurya, Antihyperglycemic activity of compounds isolated from Indian medicinal plants, Indian J Exp Biol. 48 (2010) 294–298.

[7] Vismaya, S.M. Belagihally, S. Rajashekhar, V.B. Jayaram, S.M. Dharmesh, S.K.C. Thirumakudalu, Gastroprotective properties of karanjin from Karanja (*Pongamia pinnata*) seeds; Role as antioxidant and H+, K+-ATPase inhibitor, Evidence-Based Complementary and Alternative Medicine. 2011 (2011) 1–10. https://doi.org/10.1093/ecam/neq027.

[8] J.R. Guo, Q.Q. Chen, C. Wai-Kei Lam, W. Zhang, Effects of karanjin on cell cycle arrest and apoptosis in human A549, HepG2 and HL-60 cancer cells, Biol Res. 48 (2015) 1–7. https://doi.org/10.1186/s40659-015-0031-x.

[9] R. Roy, D. Pal, S. Sur, S. Mandal, P. Saha, C.K. Panda, Pongapin and Karanjin, furanoflavanoids of *Pongamia pinnata*, induce G2/M arrest and apoptosis in cervical cancer cells by differential reactive oxygen species modulation, DNA damage, and nuclear factor kappa-light-chain-enhancer of activated B cell signal, Phytotherapy Research. 33 (2019) 1084–1094. https://doi.org/10.1002/ptr.6302.

[10] R. Roy, S. Mandal, J. Chakrabarti, P. Saha, C.K. Panda, Downregulation of Hyaluronic acid-CD44 signaling pathway in cervical cancer cell by natural polyphenols Plumbagin, Pongapin and Karanjin, Mol Cell Biochem. 476 (2021) 3701–3709. https://doi.org/10.1007/s11010-021-04195-1.

[11] J. Yu, H. Yang, C. Lv, X. Dai, The cytotoxicity of karanjin toward breast cancer cells is involved in the PI3K/Akt signaling pathway, Drug Dev Res. 83 (2022) 1673–1682. https://doi.org/10.1002/ddr.21986.

[12] D. Raghav, S. Mahanty, K. Rathinasamy, Biochemical and toxicological investigation of karanjin, a bio-pesticide isolated from *Pongamia* seed oil, Pestic Biochem Physiol. 157 (2019) 108–121. https://doi.org/10.1016/j.pestbp.2019.03.011.

[13] T.T.Y. Wang, N. Sathyamoorthy, J.M. Phang, Molecular effects of genistein on estrogen receptor mediated pathways, Carcinogenesis. 17 (1996) 271–275. https://doi.org/10.1093/carcin/17.2.271.

[14] R. Mukherjee, P. Pandya, D. Baxi, A. V. Ramachandran, Endocrine Disruptors–'Food' for Thought, Proc Zool Soc. 74 (2021) 432–442. https://doi.org/10.1007/s12595-021-00414-1.

[15] P. Gong, Z. Madak-Erdogan, J.A. Flaws, D.J. Shapiro, J.A. Katzenellenbogen, B.S. Katzenellenbogen, Estrogen receptor-α and aryl hydrocarbon receptor involvement in the actions of botanical estrogens in target cells, Mol Cell Endocrinol. 437 (2016) 190–200. https://doi.org/10.1016/j.mce.2016.08.025.

[16] E. Goya-Jorge, M.E.J. Rodríguez, M.S.I. Veitía, R.M. Giner, Plant occurring flavonoids as modulators of the aryl hydrocarbon receptor, Molecules. 26 (2021). https://doi.org/10.3390/molecules26082315.

[17] T. Oseni, R. Patel, J. Pyle, V.C. Jordan, Selective estrogen receptor modulators and phytoestrogens, Planta Med. 74 (2008) 1656–1665. https://doi.org/10.1055/s-0028-1088304.

[18] P. Diel, S. Olff, S. Schmidt, H. Michna, Molecular identification of potential selective estrogen receptor modulator (SERM) like properties of phytoestrogens in the human breast cancer cell line MCF-7, Planta Med. 67 (2001) 510–514. https://doi.org/10.1055/s-2001-16474.

[19] S. Martinkovich, D. Shah, S.L. Planey, J.A. Arnott, Selective estrogen receptor modulators: Tissue specificity and clinical utility, Clin Interv Aging. 9 (2014) 1437–1452. https://doi.org/10.2147/CIA.S66690.

[20] K.C. An, Selective estrogen receptor modulators, Asian Spine J. 10 (2016) 787–791. https://doi.org/10.4184/asj.2016.10.4.787.

[21] W.C. Park, V.C. Jordan, Selective estrogen receptor modulators (SERMS) and their roles in breast cancer prevention, Trends Mol Med. 8 (2002) 82–88. https://doi.org/10.1016/S1471-4914(02)02282-7.

[22] M.A. Callero, A.I. Loaiza-Pérez, The role of aryl hydrocarbon receptor and crosstalk with estrogen receptor in response of breast cancer cells to the novel

antitumor agents benzothiazoles and aminoflavone, Int J Breast Cancer. 2011 (2011) 1–9. https://doi.org/10.4061/2011/923250.

[23] E. Swedenborg, I. Pongratz, AhR and ARNT modulate ER signaling, Toxicology. 268 (2010) 132–138. https://doi.org/10.1016/j.tox.2009.09.007.

[24] J. Russo, I.H. Russo, The role of estrogen in the initiation of breast cancer, Journal of Steroid Biochemistry and Molecular Biology. 102 (2006) 89–96. https://doi.org/10.1016/j.jsbmb.2006.09.004.

[25] J.A. Cauley, F.L. Lucas, L.H. Kuller, K. Stone, W. Browner, S.R. Cummings, Elevated serum estradiol and testosterone concentrations are associated with a high risk for breast cancer, Ann Intern Med. 130 (1999) 270–277. https://doi.org/10.7326/0003-4819-130-4_part_1-199902160-00004.

[26] L. Castagnetta, O.M. Granata, L. Cocciadiferro, A. Saetta, L. Polito, G. Bronte, S. Rizzo, I. Campisi, B. Agostara, G. Carruba, Sex steroids, carcinogenesis, and cancer progression, Ann N Y Acad Sci. 1028 (2004) 233–246. https://doi.org/10.1196/annals.1321.028.

[27] V. Jayagopal, P. Albertazzi, E.S. Kilpatrick, E.M. Howarth, P.E. Jennings, D.A. Hepburn, S.L. Atkin, Beneficial effects of soy phytoestrogen intake in postmenopausal women with type 2 diabetes, Diabetes Care. 25 (2002) 1709–1714. https://doi.org/10.2337/diacare.25.10.1709.

[28] P. Gong, Z. Madak-Erdogan, J. Li, J. Cheng, C.M. Greenlief, W. Helferich, J.A. Katzenellenbogen, B.S. Katzenellenbogen, Transcriptomic analysis identifies gene networks regulated by estrogen receptor α (ERα) and ERβ that control distinct effects of different botanical estrogens, Nucl Recept Signal. 12 (2014) e001. https://doi.org/10.1621/nrs.12001.

[29] M. Marino, P. Galluzzo, Is There an Answer? Are flavonoids agonists or antagonists of the natural hormone 17β-estradiol? IUBMB Life. 60 (2008) 241–244. https://doi.org/10.1002/iub.34.

[30] P. Galluzzo, M. Marino, Nutritional flavonoids impact on nuclear and extranuclear estrogen receptor activities., Genes Nutr. 1 (2006) 161–76. https://doi.org/10.1007/bf02829966.

[31] J. Clardy, C. Walsh, Lessons from natural molecules, Nature. 432 (2004) 829–837. https://doi.org/10.1038/nature03194.

[32] D.J. Newman, G.M. Cragg, Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019, J Nat Prod. 83 (2020) 770–803. https://doi.org/10.1021/acs.jnatprod.9b01285.

[33] WHO global centre for traditional medicine, Who. (2022). https://www.who.int/initiatives/who-global-centre-for-traditional-

medicine%0Ahttps://www.who.int/news/item/25-03-2022-who-establishes-the-global-centre-for-traditional-medicine-in-India.

[34] V. Tyler, L. Brady, J. Robbers, Pharmacognosy, 1988. http://182.160.97.198:8080/xmlui/handle/123456789/767.

[35] S. Parasuraman, G.S. Thing, S.A. Dhanaraj, Polyherbal formulation: Concept of ayurveda, Pharmacogn Rev. 8 (2014) 73–80. https://doi.org/10.4103/0973-7847.134229.

[36] D. Raghav, S. Mahanty, K. Rathinasamy, Biochemical and toxicological investigation of karanjin, a bio-pesticide isolated from *Pongamia* seed oil, Pestic Biochem Physiol. 157 (2019) 108–121. https://doi.org/10.1016/j.pestbp.2019.03.011.

[37] M.K. Gupta, S. Senthilkumar, A.K. Chiranjivi, K. Banik, S. Girisa, A.B. Kunnumakkara, V.K. Dubey, L. Rangan, Antioxidant, anti-tyrosinase and anti-inflammatory activities of 3,5-dihydroxy-4′,7-dimethoxyflavone isolated from the leaves of Alpinia nigra, Phytomedicine Plus. 1 (2021) 100097. https://doi.org/10.1016/j.phyplu.2021.100097.

[38] A.K. Pandey, P. Kumar, P. Singh, N.N. Tripathi, V.K. Bajpai, Essential oils: Sources of antimicrobials and food preservatives, Front Microbiol. 7 (2017) 2161. https://doi.org/10.3389/fmicb.2016.02161.

[39] S. Kumar, A.K. Pandey, Chemistry, and biological activities of flavonoids: An overview, The Scientific World Journal. 2013 (2013) 70–78. https://doi.org/10.1155/2013/162750.

[40] A. Ullah, S. Munir, S.L. Badshah, N. Khan, L. Ghani, B.G. Poulson, A.H. Emwas, M. Jaremko, Important flavonoids and their role as a therapeutic agent, Molecules. 25 (2020) 5243. https://doi.org/10.3390/molecules25225243.

[41] L. Chen, H. Cao, Q. Huang, J. Xiao, H. Teng, Absorption, metabolism, and bioavailability of flavonoids: a review, Crit Rev Food Sci Nutr. 62 (2022) 7730–7742. https://doi.org/10.1080/10408398.2021.1917508.

[42] J.E. Vargas, R. Puga, J.D.F. Poloni, L.F. Saraiva Macedo Timmers, B.N. Porto, O. Norberto De Souza, D. Bonatto, P.M. Condessa Pitrez, R. Tetelbom Stein, A network flow approach to predict protein targets and flavonoid backbones to treat respiratory syncytial virus infection, Biomed Res Int. 2015 (2015). https://doi.org/10.1155/2015/301635.

[43] L.N. Misra, N.A.V. Wouatsa, S. Kumar, R. Venkatesh Kumar, F. Tchoumbougnang, Antibacterial, cytotoxic activities and chemical composition of fruits of two Cameroonian Zanthoxylum species, J Ethnopharmacol. 148 (2013) 74–80. https://doi.org/10.1016/j.jep.2013.03.069.

[44] W. Zhu, Q. Jia, Y. Wang, Y. Zhang, M. Xia, The anthocyanin cyanidin-3-O-β-glucoside, a flavonoid, increases hepatic glutathione synthesis and protects hepatocytes against reactive oxygen species during hyperglycemia: Involvement of a cAMP-PKA-dependent signaling pathway, Free Radic Biol Med. 52 (2012) 314–327. https://doi.org/10.1016/j.freeradbiomed.2011.10.483.

[45] L. Duthie, R.M. Reynolds, Changes in the maternal hypothalamic-pituitary-adrenal axis in pregnancy and postpartum: Influences on maternal and fetal outcomes, Neuroendocrinology. 98 (2013) 106–115. https://doi.org/10.1159/000354702.

[46] H.S. Gehlot, D. Panwar, N. Tak, A. Tak, I.S. Sankhla, N. Poonar, R. Parihar, N.S. Shekhawat, M. Kumar, R. Tiwari, J. Ardley, E.K. James, J.I. Sprent, Nodulation of legumes from the Thar desert of India and molecular characterization of their rhizobia, Plant Soil. 357 (2012) 227–243. https://doi.org/10.1007/s11104-012-1143-5.

[47] L. Fan, Y. Zhang, R. Huang, S. Qin, T. Yi, F. Xu, Y. Tang, X. Qu, H. Chen, J. Miao, Determination of five flavonoids in different parts of *Fordia cauliflora* by ultra performance liquid chromatography/triple-quadrupole mass spectrometry and chemical comparison with the root of *Millettia pulchra* var. laxior, Chem Cent J. 7 (2013) 1–9. https://doi.org/10.1186/1752-153X-7-126.

[48] A.F. Magalhães, A.M.A. Tozzi, E.G. Magalhães, M.A. Nogueira, S.C.N. Queiroz, Flavonoids from *Lonchocarpus latifolius* roots, Phytochemistry. 55 (2000) 787–792. https://doi.org/10.1016/S0031-9422(00)00300-9.

[49] W. Shao, Y. Zhu, S. Guang, S. Zhang, F. Chen, Study on chemical constituents of thickfruit millettia root, Tianran Chanwu Yanjiu Yu Kaifa. 13 (2001) 1–4.

[50] C.M.R. Gomes, O.R. Gottlieb, G.B. Marini Bettolo, F. Delle Monache, R.M. Polhill, Systematic significance of flavonoids in Derris and Lonchocarpus, Biochem Syst Ecol. 9 (1981) 129–147. https://doi.org/10.1016/0305-1978(81)90031-4.

[51] V.P. Pathak, T.R. Saini, R.N. Khanna, Glabrachalcone, a chromenochalcone from *Pongamia glabra* seeds, Phytochemistry. 22 (1983) 1303–1304. https://doi.org/10.1016/0031-9422(83)80254-4.

[52] U. Sriphana, C. Yenjai, S. Tungnoi, J. Srirapa, A. Junsongduang, Flavonoids from *Milletia leucantha* and their cytotoxicity, Nat Prod Commun. 13 (2018) 961–962. https://doi.org/10.1177/1934578x1801300810.

[53] W.A. Peer, A.S. Murphy, Flavonoids as signal molecules: Targets of flavonoid action, in: The Science of Flavonoids, Springer New York, New York, NY, 2006: pp. 239–268. https://doi.org/10.1007/978-0-387-28822-2_9.

[54] V. Kesari, Characterization of candidate plus trees (CPTs) of *Pongamia pinnata* (L.), a versatile legume from North Guwahati, Assam (2010).

[55] F. Zsila, Z. Bikádi, M. Simonyi, Probing the binding of the flavonoid, quercetin to human serum albumin by circular dichroism, electronic absorption spectroscopy and molecular modelling methods, Biochem Pharmacol. 65 (2003) 447–456. https://doi.org/10.1016/S0006-2952(02)01521-6.

[56] N.V. Bringi, Non-Traditional Oilseed and Oils in India, (1987) 109–117. https://doi.org/10.3/jquery-ui.js.

[57] Rastogi BN and Mehrotra RP, Compendium of Indian Medicinal Plants., Central Drug Research Institute and Publications & Information Directorate, New Delhi, (1993).

[58] J.B. Harborne, Indian Medicinal Plants. A Compendium of 500 Species. Vol.1; Edited by P. K. Warrier, V. P. K. Nambiar and C. Ramankutty, Journal of Pharmacy and Pharmacology. 46 (2011) 935–935. https://doi.org/10.1111/j.2042-7158.1994.tb05722.x.

[59] D.B. Limaye, Karanjin part I: a crystalline constituent of the oil from *Pongamia* glabra., Proceedings of the 12th Indian Academy and Science Congress, India. 118 (1925) 118–125.

[60] S.K. Talapatra, A.K. Mallik, B. Talapatra, Pongaglabol, a new hydroxyfuranoflavone, and aurantiamide acetate, a dipeptide from the flowers of *Pongamia* glabra, Phytochemistry. 19 (1980) 1199–1202. https://doi.org/10.1016/0031-9422(80)83083-4.

[61] M.M. Saha, U.K. Mallik, A.K. Mallik, A chromenoflavanone and two caffeic esters from *Pongamia* glabra, Phytochemistry. 30 (1991) 3834–3836. https://doi.org/10.1016/0031-9422(91)80130-S.

[62] T. Tanaka, M. Iinuma, K. Yuki, Y. Fujii, M. Mizuno, Flavonoids in root bark of *Pongamia pinnata*, Phytochemistry. 31 (1992) 993–998. https://doi.org/10.1016/0031-9422(92)80055-J.

[63] N.V.S. Rao, J. Veerabhadrarao, T.R. Seshadri, A note on the preparation and reactions of Karanjin, Proceedings of the Indian Academy of Sciences - Section A. 10 (1939) 65–70. https://doi.org/10.1007/bf03170990.

[64] L. Ramachandra Row, New flavones from *Pongamia pinnata* (L.) Merr, Aust J Chem. 5 (1952) 754–759. https://doi.org/10.1071/ch9520754.

[65] V.K. Gore, P. Satyamoorthy, Determination of pongamol and karanjin in karanja oil by reverse phase HPLC, Anal Lett. 33 (2000) 337–346. https://doi.org/10.1080/00032710008543056.

[66] S. Kumar Dhanmane, F. A. Salih, Isolation of karanjin from *Pongamia pinnata* and its identification by difference analytical techniques, Kurdistan Journal of Applied Research. (2018) 156–160. https://doi.org/10.24017/science.2018.2.26.

[67] S. Panpraneecharoen, Variability of oil content, fatty acid composition and karanjin content in *Pongamia pinnata* and its relationship with biodiesel quality, Annu Res Rev Biol. 4 (2014) 2283–2294. https://doi.org/10.9734/arrb/2014/9218.

[68] Vismaya, W. Sapna Eipeson, J.R. Manjunatha, P. Srinivas, T.C. Sindhu Kanya, Extraction, and recovery of karanjin: A value addition to karanja (*Pongamia pinnata*) seed oil, Ind Crops Prod. 32 (2010) 118–122. https://doi.org/10.1016/j.indcrop.2010.03.011.

[69] M. Verma, S. Pradhan, S. Sharma, S.N. Naik, R. Prasad, Efficacy of karanjin and phorbol ester fraction against termites (Odontotermes obesus), Int Biodeterior Biodegradation. 65 (2011) 877–882. https://doi.org/10.1016/j.ibiod.2011.05.007.

[70] S.D. Katekhaye, M.S. Kale, K.S. Laddha, A simple and improved method for isolation of karanjin from *Pongamia pinnata* Linn. seed oil, 2012.

[71] A. Ghosh Mitra, S. Mandal, P. Das, S. Dasgupta, S. Mukhopadhyay, A. Mukhopadhyay, J. Banerjee, M. Kar, Antioxidant and free radical scavenging properties of seed components of *Pongamia pinnata*-a comparative study, Organic & Medicinal Chem IJ. 7 (2018) 1–5. https://doi.org/10.19080/omcij.2018.07.555723.

[72] R.S. Susarla, M.M. Murthy, B.V.S.K. Rao, P.P. Chakrabarti, R.B.N. Prasad, S. Kanjilal, A method for isolation of karanjin from the expelled cake of *Pongamia* glabra, European Journal of Lipid Science and Technology. 114 (2012) 1097–1101. https://doi.org/10.1002/ejlt.201200016.

[73] A. kumar Pandey, A.K. Bajpai, A. Kumar, M. Pal, V. Baboo, A. Dwivedi, Isolation, identification, molecular and electronic structure, vibrational spectroscopic investigation, and anti-HIV-1 activity of karanjin using density functional theory, Journal of Theoretical Chemistry. 2014 (2014) 1–13. https://doi.org/10.1155/2014/680987.

[74] P. Joshi, V.R. Sonawane, I.S. Williams, G.J.P. McCann, L. Gatchie, R. Sharma, N. Satti, B. Chaudhuri, S.B. Bharate, Identification of karanjin isolated from the Indian beech tree as a potent CYP1 enzyme inhibitor with cellular efficacy: Via screening of a natural product repository, Medchemcomm. 9 (2018) 371–382. https://doi.org/10.1039/c7md00388a.

[75] N. Shejawal, S. Menon, S. Shailajan, Bioavailability of karanjin from *Pongamia pinnata* L. In Sprague dawley rats using validated RP-HPLC method, J Appl Pharm Sci. 4 (2014) 10–14. https://doi.org/10.7324/japs.2014.40303.

[76] N. Jaiswal, P.P. Yadav, R. Maurya, A.K. Srivastava, A.K. Tamrakar, Karanjin from *Pongamia pinnata* induces GLUT4 translocation in skeletal muscle cells in a phosphatidylinositol-3-kinase-independent manner, Eur J Pharmacol. 670 (2011) 22–28. https://doi.org/10.1016/j.ejphar.2011.08.049.

[77] P.P. Patel, N.D. Trivedi, Effect of karanjin on 2,4,6-trinitrobenzenesulfonic acid-induced colitis in Balb/c mice, Indian J Pharmacol. 49 (2017) 161–167. https://doi.org/10.4103/ijp.ijp_234_15.

[78] P. Saini, L. Lakshmayya, V. Bisht, Anti-Alzheimer activity of isolated karanjin from *Pongamia pinnata* (L.) pierre and embelin from *Embelia ribes* Burm.f., AYU (An International Quarterly Journal of Research in Ayurveda). 38 (2017) 76. https://doi.org/10.4103/ayu.ayu_174_16.

[79] M. Michaelis, F. Rothweiler, T. Nerreter, M. Sharifi, T. Ghafourian, J. Cinatl, Karanjin interferes with ABCB1, ABCC1, and ABCG2, Journal of Pharmacy and Pharmaceutical Sciences. 17 (2014) 92–105. https://doi.org/10.18433/j3bw2s.

[80] C. Arulvasu, S. Vasantha, G. Babu, Anticancer activity of *Pongamia glabra* V. seed oil extract against selected human cancer cell lines, Irjp. 2012 (2012) 3. www.irjponline.com.

[81] M. Bose, M. Chakraborty, S. Bhattacharya, D. Mukherjee, S. Mandal, R. Mishra, Prevention of arthritis markers in experimental animal and inflammation signalling in macrophage by karanjin isolated from *Pongamia pinnata* seed extract, Phytotherapy Research. 28 (2014) 1188–1195. https://doi.org/10.1002/ptr.5113.

[82] A.K. Tamrakar, P.P. Yadav, P. Tiwari, R. Maurya, A.K. Srivastava, Identification of pongamol and karanjin as lead compounds with antihyperglycemic activity from *Pongamia pinnata* fruits, J Ethnopharmacol. 118 (2008) 435–439. https://doi.org/10.1016/j.jep.2008.05.008.

[83] M. Bose, M. Chakraborty, S. Bhattacharya, P. Bhattacharjee, S. Mandal, M. Kar, R. Mishra, Suppression of NF-κB p65 nuclear translocation and tumor necrosis factor-α by *Pongamia pinnata* seed extract in adjuvant-induced arthritis, J Immunotoxicol. 11 (2014) 222–230. https://doi.org/10.3109/1547691x.2013.824931.

[84] N. Arshad, N. Rashid, S. Absar, M.S.A. Abbasi, S. Saleem, B. Mirza, UV-absorption studies of interaction of karanjin and karanjachromene with ds. DNA:

Evaluation of binding and antioxidant activity, Central European Journal of Chemistry. 11 (2013) 2040–2047. https://doi.org/10.2478/s11532-013-0327-z.

[85] M. Bose, M. Chakraborty, S. Bhattacharya, D. Mukherjee, S. Mandal, R. Mishra, Prevention of arthritis markers in experimental animal and inflammation signalling in macrophage by karanjin isolated from *Pongamia pinnata* seed extract, Phytotherapy Research. 28 (2014) 1188–1195. https://doi.org/10.1002/ptr.5113.

[86] P.P. Patel, N.D. Trivedi, Simple, efficient, and economic method forisolation and analysis of karanjin and pongamol from karanja seed oil and screening of antimicrobial potential, Int J Pharm Pharm Sci. 7 (2015) 248–252.

[87] C. Magnusson, J.A. Baron, N. Correia, R. Bergström, H.O. Adami, I. Persson, Breast-cancer risk following long-term oestrogen-and oestrogen- progestin-replacement therapy, Int J Cancer. 81 (1999) 339–344. https://doi.org/10.1002/(sici)1097-0215(19990505)81:3<339::aid-ijc5>3.0.co;2-6.

[88] G.G.J.M. Kuiper, J.G. Lemmen, B. Carlsson, J.C. Corton, S.H. Safe, P.T. Van Der Saag, B. Van Der Burg, J.Å. Gustafsson, Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor β, Endocrinology. 139 (1998) 4252–4263. https://doi.org/10.1210/endo.139.10.6216.

[89] R.C.Y. Choi, J.T.T. Zhu, A.W.Y. Yung, P.S.C. Lee, S.L. Xu, A.J.Y. Guo, K.Y. Zhu, T.T.X. Dong, K.W.K. Tsim, Synergistic action of flavonoids, baicalein, and daidzein in estrogenic and neuroprotective effects: A development of potential health products and therapeutic drugs against Alzheimer's disease, Evidence-Based Complementary and Alternative Medicine. 2013 (2013) 635694. https://doi.org/10.1155/2013/635694.

[90] P. Ascenzi, A. Bocedi, M. Marino, Structure-function relationship of estrogen receptor α and β: Impact on human health, Mol Aspects Med. 27 (2006) 299–402. https://doi.org/10.1016/j.mam.2006.07.001.

[91] S. Choi, T. Ha, J. Ahn, S. Kim, K. Kang, I. Hwang, S. Kim, Estrogenic activities of isoflavones and flavones and their structure-activity relationships, Planta Med. 74 (2008) 25–32. https://doi.org/10.1055/s-2007-993760.

[92] G.G.J.M. Kuiper, B. Carlsson, K. Grandien, E. Enmark, J. Häggblad, S. Nilsson, J.Å. Gustafsson, Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors and α and β, Endocrinology. 138 (1997) 863–870. https://doi.org/10.1210/endo.138.3.4979.

[93] C.J. Gruber, W. Tschugguel, C. Schneeberger, J.C. Huber, Production and actions of estrogens, New England Journal of Medicine. 346 (2002) 340–352. https://doi.org/10.1056/nejmra000471.

[94] S.L. Planey, R. Kumar, J.A. Arnott, Estrogen receptors (ERα versus ERβ): Friends or foes in human biology?, Journal of Receptors and Signal Transduction. 34 (2014) 1–5. https://doi.org/10.3109/10799893.2013.853188.

[95] S. Nilsson, S. Mäkelä, E. Treuter, M. Tujague, J. Thomsen, G. Andersson, E. Enmark, K. Pettersson, M. Warner, J.Å. Gustafsson, Mechanisms of estrogen action, Physiol Rev. 81 (2001) 1535–1565. https://doi.org/10.1152/physrev.2001.81.4.1535.

[96] K.W. Nettles, J.B. Bruning, G. Gil, E.E. O'Neill, J. Nowak, A. Hughs, Y. Kim, E.R. DeSombre, R. Dilis, R.N. Hanson, A. Joachimiak, G.L. Greene, Structural plasticity in the oestrogen receptor ligand-binding domain, EMBO Rep. 8 (2007) 563–568. https://doi.org/10.1038/sj.embor.7400963.

[97] M. Ruff, M. Gangloff, J.M. Wurtz, D. Moras, Structure-function relationships in DNA- and ligand-binding domains of estrogen receptors, Breast Cancer Research. 2 (2000) 353–359. https://doi.org/10.1186/bcr80.

[98] D.M. Tanenbaum, Y. Wang, S.P. Williams, P.B. Sigler, Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains, Proc Natl Acad Sci U S A. 95 (1998) 5998–6003. https://doi.org/10.1073/pnas.95.11.5998.

[99] R. Kumar, M.N. Zakharov, S.H. Khan, R. Miki, H. Jang, G. Toraldo, R. Singh, S. Bhasin, R. Jasuja, The Dynamic Structure of the Estrogen Receptor, J Amino Acids. 2011 (2011) 1–7. https://doi.org/10.4061/2011/812540.

[100] A.M. Brzozowski, A.C.W. Pike, Z. Dauter, R.E. Hubbard, T. Bonn, O. Engström, L. Öhman, G.L. Greene, J.Å. Gustafsson, M. Carlquist, Molecular basis of agonism and antagonism in the oestrogen receptor, Nature. 389 (1997) 753–758. https://doi.org/10.1038/39645.

[101] A.C.W. Pike, A.M. Brzozowski, R.E. Hubbard, A structural biologist's view of the oestrogen receptor, Journal of Steroid Biochemistry and Molecular Biology. 74 (2000) 261–268. https://doi.org/10.1016/S0960-0760(00)00102-3.

[102] L. Larigot, L. Juricek, J. Dairou, X. Coumoul, AhR signaling pathways and regulatory functions, Biochim Open. 7 (2018) 1–9. https://doi.org/10.1016/j.biopen.2018.05.001.

[103] V. Rothhammer, F.J. Quintana, The aryl hydrocarbon receptor: an environmental sensor integrating immune responses in health and disease, Nat Rev Immunol. 19 (2019) 184–197. https://doi.org/10.1038/s41577-019-0125-8.

[104] I.A. Murray, A.D. Patterson, G.H. Perdew, Aryl hydrocarbon receptor ligands in cancer: Friend and foe, Nat Rev Cancer. 14 (2014) 801–814. https://doi.org/10.1038/nrc3846.

[105] C. Wang, C.X. Xu, Y. Bu, K.M. Bottum, S.A. Tischkau, Beta-naphthoflavone (DB06732) mediates estrogen receptor-positive breast cancer cell cycle arrest through AhR-dependent regulation of PI3K/AKT and MAPK/ERK signaling, Carcinogenesis. 35 (2014) 703–713. https://doi.org/10.1093/carcin/bgt356.

[106] A. Paris, N. Tardif, M.D. Galibert, S. Corre, AhR and cancer: From gene profiling to targeted therapy, Int J Mol Sci. 22 (2021) 1–22. https://doi.org/10.3390/ijms22020752.

[107] J. Matthews, S. Ahmed, AHR- and ER-mediated toxicology and chemoprevention, Advances in Molecular Toxicology. 7 (2013) 1–38. https://doi.org/10.1016/b978-0-444-62645-5.00001-8.

[108] S. Safe, U.H. Jin, H. Park, R.S. Chapkin, A. Jayaraman, Aryl hydrocarbon receptor (AhR) ligands as selective AhR modulators (SAHRMS), Int J Mol Sci. 21 (2020) 1–16. https://doi.org/10.3390/ijms21186654.

[109] M. Rodriguez, D.A. Potter, CYP1A1 regulates breast cancer proliferation and survival, Molecular Cancer Research. 11 (2013) 780–792. https://doi.org/10.1158/1541-7786.mcr-12-0675.

[110] G.I. Murray, M.C. Taylor, M.C.E. McFadyen, J.A. McKay, W.F. Greenlee, M.D. Burke, W.T. Melvin, Tumor-specific expression of cytochrome P450 CYP1B1, Cancer Res. 57 (1997) 3026–3031. http://www.ncbi.nlm.nih.gov/pubmed/9230218.

[111] D.C. Spink, H. Pietro Eugster, D.W. Lincoln, J.D. Schuetz, E.G. Schuetz, J.A. Johnson, L.S. Kaminsky, J.F. Gierthy, 17β-Estradiol hydroxylation catalyzed by human cytochrome P450 1A1: A comparison of the activities induced by 2,3,7,8-tetrachlorodibenzo-p-dioxin in MCF-7 cells with those from heterologous expression of the cDNA, Arch Biochem Biophys. 293 (1992) 342–348. https://doi.org/10.1016/0003-9861(92)90404-K.

[112] M. Kittaneh, A.J. Montero, S. Glück, Molecular profiling for breast cancer: a comprehensive review, Biomark Cancer. 5 (2013) BIC.S9455. https://doi.org/10.4137/bic.s9455.

[113] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J Clin. 68 (2018) 394–424. https://doi.org/10.3322/caac.21492.

[114] A. Jemal, F. Bray, J. Ferlay, Global Cancer Statistics: 2011, CA Cancer J Clin. 49 (1999) 1,33-64. https://doi.org/10.3322/caac.20107.

[115] J. Ferlay, M. Colombet, I. Soerjomataram, C. Mathers, D.M. Parkin, M. Piñeros, A. Znaor, F. Bray, Estimating the global cancer incidence and mortality in 2018:

GLOBOCAN sources and methods, Int J Cancer. 144 (2019) 1941–1953. https://doi.org/10.1002/ijc.31937.

[116] R.O. Roblin, Medicinal chemistry, Chemical and Engineering News. 31 (1953) 48–49. https://doi.org/10.1021/cen-v031n001.p048.

[117] R. García-Becerra, N. Santos, L. Díaz, J. Camacho, Mechanisms of resistance to endocrine therapy in breast cancer: Focus on signaling pathways, miRNAs and genetically based resistance, Int J Mol Sci. 14 (2013) 108–145. https://doi.org/10.3390/ijms14010108.

[118] C.K. Osborne, A. Wakeling, R.I. Nicholson, Fulvestrant: An oestrogen receptor antagonist with a novel mechanism of action, Br J Cancer. 90 (2004) S2–S6. https://doi.org/10.1038/sj.bjc.6601629.

[119] R. Cheng, L. Qi, X. Kong, Z. Wang, Y. Fang, J. Wang, Identification of the significant genes regulated by estrogen receptor in estrogen receptor-positive breast cancer and their expression pattern changes when tamoxifen or fulvestrant resistance occurs, Front Genet. 11 (2020) 538734. https://doi.org/10.3389/fgene.2020.538734.

[120] D. Zardavas, M. Piccart, Neoadjuvant therapy for breast cancer, Annu Rev Med. 66 (2015) 31–48. https://doi.org/10.1146/annurev-med-051413-024741.

[121] H.K. Chew, Medicine cabinet: Adjuvant therapy for breast cancer: Who should get what?, Western Journal of Medicine. 174 (2001) 284–287. https://doi.org/10.1136/ewjm.174.4.284.

[122] M.S. Sheikh, M. Garcia, P. Pujol, J.A. Fontana, H. Rochefort, Why are estrogen-receptor-negative breast cancers more aggressive than the estrogen-receptor-positive breast cancers?, Invasion Metastasis. 14 (1994) 329–336.

[123] J.R. Baker, J.A. Sakoff, A. McCluskey, The aryl hydrocarbon receptor (AhR) as a breast cancer drug target, Med Res Rev. 40 (2020) 972–1001. https://doi.org/10.1002/med.21645.

[124] J.A. Katzenellenbogen, B.W. O'Malley, B.S. Katzenellenbogen, Tripartite steroid hormone receptor pharmacology: Interaction with multiple effector sites as a basis for the cell- and promoter-specific action of these hormones, Molecular Endocrinology. 10 (1996) 119–131. https://doi.org/10.1210/me.10.2.119.

[125] S. Safe, H. Han, J. Goldsby, K. Mohankumar, R.S. Chapkin, Aryl hydrocarbon receptor (AhR) ligands as selective AhR modulators: Genomic studies, Curr Opin Toxicol. 11–12 (2018) 10–20. https://doi.org/10.1016/j.cotox.2018.11.005.

[126] L.L. Legette, W.H. Lee, B.R. Martin, J.A. Story, A. Arabshahi, S. Barnes, C.M. Weaver, Genistein, a phytoestrogen, improves total cholesterol, and Synergy, a prebiotic, improves calcium utilization, but there were no synergistic effects,

Menopause. 18 (2011) 923–931. https://doi.org/10.1097/gme.0b013e3182116e81.

[127] D. Dolciami, M. Ballarotto, M. Gargaro, L.C. López-Cara, F. Fallarino, A. Macchiarulo, Targeting Aryl hydrocarbon receptor for next-generation immunotherapies: Selective modulators (SAhRMs) versus rapidly metabolized ligands (RMAhRLs), Eur J Med Chem. 185 (2020) 111842. https://doi.org/10.1016/j.ejmech.2019.111842.

[128] J. Gilbert, G.N. De Iuliis, A. McCluskey, J.A. Sakoff, A novel naphthalimide that selectively targets breast cancer via the aryl hydrocarbon receptor pathway, Sci Rep. 10 (2020) 1–12. https://doi.org/10.1038/s41598-020-70597-8.

[129] M. Kalia, Biomarkers for personalized oncology: Recent advances and future challenges, Metabolism. 64 (2015) S16–S21. https://doi.org/10.1016/j.metabol.2014.10.027.

[130] G. Bhatt, A. Gupta, L. Rangan, A. Mukund Limaye, Global transcriptome analysis reveals partial estrogen-like effects of karanjin in MCF-7 breast cancer cells, Gene. 830 (2022) 146507. https://doi.org/10.1016/j.gene.2022.146507.

[131] W. Strober, Trypan blue exclusion test of cell viability., Current Protocols in Immunology. Appendix 3 (2001) Appendix 3B. https://doi.org/10.1002/0471142735.ima03bs21.

[132] S. Anders, Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data, Soil. 5 (2010) https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

[133] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data, Bioinformatics. 30 (2014) 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

[134] A. Dobin, C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T.R. Gingeras, STAR: Ultrafast universal RNA-seq aligner, Bioinformatics. 29 (2013) 15–21. https://doi.org/10.1093/bioinformatics/bts635.

[135] Y. Liao, G.K. Smyth, W. Shi, FeatureCounts: An efficient general-purpose program for assigning sequence reads to genomic features, Bioinformatics. 30 (2014) 923–930. https://doi.org/10.1093/bioinformatics/btt656.

[136] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol. 15 (2014) 1–21. https://doi.org/10.1186/s13059-014-0550-8.

[137] A.A. Sergushichev, an algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation, BioRxiv. (2016) 060012.

[138] J. Guan, W. Zhou, M. Hafner, R.A. Blake, C. Chalouni, I.P. Chen, T. De Bruyn, J.M. Giltnane, S.J. Hartman, A. Heidersbach, R. Houtman, E. Ingalla, L. Kategaya, T. Kleinheinz, J. Li, S.E. Martin, Z. Modrusan, M. Nannini, J. Oeh, S. Ubhayakar, X. Wang, I.E. Wertz, A. Young, M. Yu, D. Sampath, J.H. Hager, L.S. Friedman, A. Daemen, C. Metcalfe, Therapeutic Ligands Antagonize Estrogen Receptor Function by Impairing Its Mobility, Cell. 178 (2019) 949-963.e18. https://doi.org/10.1016/j.cell.2019.06.026.

[139] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method, Methods. 25 (2001) 402–408. https://doi.org/10.1006/meth.2001.1262.

[140] O.H. Lowry, N.J. Rosebrough, A.L. Farr, R.J. Randall, Protein measurement with the Folin phenol reagent., J Biol Chem. 193 (1951) 265–275. https://doi.org/10.1016/s0021-9258(19)52451-6.

[141] W.S. Choi, K.J. Chung, M.S. Chang, J.K. Chun, H.W. Lee, S.Y. Hong, A turbidimetric determination of protein by trichloroacetic acid, Arch Pharm Res. 16 (1993) 57–61. https://doi.org/10.1007/BF02974129.

[142] S. Shivaswamy, V.R. Iyer, Genome-wide analysis of chromatin status using tiling microarrays, Methods. 41 (2007) 304–311. https://doi.org/10.1016/j.ymeth.2006.11.002.

[143] D.J.S. John Mary, G. Sikarwar, A. Kumar, A.M. Limaye, Interplay of ERα binding and DNA methylation in the intron-2 determines the expression and estrogen regulation of cystatin A in breast cancer cells, Mol Cell Endocrinol. 504 (2020) 679043. https://doi.org/10.1016/j.mce.2020.110701.

[144] A. Kumar, A. Dhillon, M.C. Manjegowda, N. Singh, D.J.S.J. Mary, S. Kumar, D. Modi, A.M. Limaye, Estrogen suppresses HOXB2 expression via ERα in breast cancer cells, Gene. 794 (2021) 145746. https://doi.org/10.1016/j.gene.2021.145746.

[145] Y. Shang, X. Hu, J. DiRenzo, M.A. Lazar, M. Brown, Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription, Cell. 103 (2000) 843–852. https://doi.org/10.1016/S0092-8674(00)00188-4.

[146] C. Colovos, T.O. Yeates, Verification of protein structures: Patterns of nonbonded atomic interactions, Protein Science. 2 (1993) 1511–1519. https://doi.org/10.1002/pro.5560020916.

[147] S. Forli, R. Huey, M.E. Pique, M.F. Sanner, D.S. Goodsell, A.J. Olson, Computational protein-ligand docking and virtual drug screening with the AutoDock suite, Nat Protoc. 11 (2016) 905–919. https://doi.org/10.1038/nprot.2016.051.

[148] A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, Journal of Physical Chemistry B. 102 (1998) 3586–3616. https://doi.org/10.1021/jp973084f.

[149] W.L. Jorgensen, J. Tirado-Rives, Potential energy functions for atomic-level simulations of water and organic and biomolecular systems, Proc Natl Acad Sci U S A. 102 (2005) 6665–6670. https://doi.org/10.1073/PNAS.0408037102.

[150] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: Fast, flexible, and free, J Comput Chem. 26 (2005) 1701–1718. https://doi.org/10.1002/jcc.20291.

[151] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: A Linear Constraint Solver for molecular simulations, J Comput Chem. 18 (1997) 1463–1472. https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

[152] A. Kumar, M.C. Manjegowda, D.J.S. John Mary, U. Pal, S. Kumar, A.M. Limaye, Estrogen receptor-α is a determinant of protocadherin-8 expression in breast cancer cells, Gene Rep. 14 (2019) 6–11. https://doi.org/10.1016/j.genrep.2018.10.014.

[153] S. Terasaka, Y. Aita, A. Inoue, S. Hayashi, M. Nishigaki, K. Aoyagi, H. Sasaki, Y. Wada-Kiyama, Y. Sakuma, S. Akaba, J. Tanaka, H. Sone, J. Yonemoto, M. Tanji, R. Kiyama, Using a customized DNA microarray for expression profiling of the estrogen-responsive genes to evaluate estrogen activity among natural estrogens and industrial chemicals, Environ Health Perspect. 112 (2004) 773–781. https://doi.org/10.1289/ehp.6753.

[154] R. El Ansari, M.L. Craze, I. Miligy, M. Diez-Rodriguez, C.C. Nolan, I.O. Ellis, E.A. Rakha, A.R. Green, The amino acid transporter SLC7A5 confers a poor prognosis in the highly proliferative breast cancer subtypes and is a key therapeutic target in luminal B tumours, Breast Cancer Research. 20 (2018) 21. https://doi.org/10.1186/s13058-018-0946-6.

[155] D.M. Miller, S.D. Thomas, A. Islam, D. Muench, K. Sedoris, c-Myc and cancer metabolism, Clinical Cancer Research. 18 (2012) 5546–5553. https://doi.org/10.1158/1078-0432.ccr-12-0977.

[156] M. Zörnig, G.I. Evan, Cell cycle: On target with Myc, Current Biology. 6 (1996) 1553–1556. https://doi.org/10.1016/S0960-9822(02)70769-0.

[157] C. Lammer, S. Wagerer, R. Saffrich, D. Mertens, W. Ansorge, I. Hoffmann, The cdc25B phosphatase is essential for the G2/M phase transition in human cells, J Cell Sci. 111 (1998) 2445–2453. https://doi.org/10.1242/jcs.111.16.2445.

[158] R. Métivier, G. Penot, M.R. Hübner, G. Reid, H. Brand, M. Koš, F. Gannon, Estrogen receptor-α directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter, Cell. 115 (2003) 751–763. https://doi.org/10.1016/S0092-8674(03)00934-6.

[159] J.A. Lavigne, Y. Takahashi, G.V.R. Chandramouli, H. Liu, S.N. Perkins, S.D. Hursting, T.T.Y. Wang, Concentration-dependent effects of genistein on global gene expression in MCF-7 breast cancer cells: An oligo microarray study, Breast Cancer Res Treat. 110 (2008) 85–98. https://doi.org/10.1007/s10549-007-9705-6.

[160] S.T. Pearce, V.C. Jordan, The biological role of estrogen receptors α and β in cancer, Crit Rev Oncol Hematol. 50 (2004) 3–22. https://doi.org/10.1016/j.critrevonc.2003.09.003.

[161] H. Hong, W.S. Branham, H.W. Ng, C.L. Moland, S.L. Dial, H. Fang, R. Perkins, D. Sheehan, W. Tong, Human sex hormone-binding globulin binding affinities of 125 structurally diverse chemicals and comparison with their binding to androgen receptor, estrogen receptor, and α-Fetoprotein, Toxicological Sciences. 143 (2015) 333–348. https://doi.org/10.1093/toxsci/kfu231.

[162] A.I. Constantinou, A.E. Krygier, R.R. Mehta, Genistein induces maturation of cultured human breast cancer cells and prevents tumor growth in nude mice, American Journal of Clinical Nutrition. 68 (1998) 1426S-1430S. https://doi.org/10.1093/ajcn/68.6.1426S.

[163] L. Fioravanti, V. Cappelletti, P. Miodini, E. Ronchi, M. Brivio, G. Di Fronzo, Genistein in the control of breast cancer cell growth: Insights into the mechanism of action in vitro, Cancer Lett. 130 (1998) 143–152. https://doi.org/10.1016/S0304-3835(98)00130-X.

[164] J.T. Hsu, H.C. Hung, C.J. Chen, W.L. Hsu, C. Ying, Effects of the dietary phytoestrogen biochanin A on cell growth in the mammary carcinoma cell line MCF-7, Journal of Nutritional Biochemistry. 10 (1999) 510–517. https://doi.org/10.1016/S0955-2863(99)00037-6.

[165] S. Yang, Q. Zhou, X. Yang, Caspase-3 status is a determinant of the differential responses to genistein between MDA-MB-231 and MCF-7 breast cancer cells, Biochim Biophys Acta Mol Cell Res. 1773 (2007) 903–911. https://doi.org/10.1016/j.bbamcr.2007.03.021.

[166] S. Ranganathan, D. Halagowder, N.D. Sivasithambaram, Quercetin suppresses twist to induce apoptosis in MCF-7 breast cancer cells, PLoS One. 10 (2015). https://doi.org/10.1371/journal.pone.0141370.

[167] H. Heberle, V.G. Meirelles, F.R. da Silva, G.P. Telles, R. Minghim, InteractiVenn: A web-based tool for the analysis of sets through Venn diagrams, BMC Bioinformatics. 16 (2015) 1–7. https://doi.org/10.1186/s12859-015-0611-3.

[168] D.J.S. John Mary, M.C. Manjegowda, A. Kumar, S. Dutta, A.M. Limaye, The role of cystatin A in breast cancer and its functional link with ERα, Journal of Genetics and Genomics. 44 (2017) 593–597. https://doi.org/10.1016/j.jgg.2017.10.001.

[169] a M. Brown, J.M. Jeltsch, M. Roberts, P. Chambon, Activation of pS2 gene transcription is a primary response to estrogen in the human breast cancer cell line MCF-7., Proc Natl Acad Sci U S A. 81 (1984) 6344–6348. https://doi.org/10.1073/pnas.81.20.6344.

[170] M. Lippman, G. Bolan, K. Huff, The effects of estrogens and antiestrogens on hormone responsive human breast cancer in long-term tissue culture, Cancer Res. 36 (1976) 4595–4601. http://www.ncbi.nlm.nih.gov/pubmed/1000504.

[171] D.T. Zava, G. Duwe, Estrogenic and antiproliferative properties of genistein and other flavonoids in human breast cancer cells in vitro, Nutr Cancer. 27 (1997) 31–40. https://doi.org/10.1080/01635589709514498.

[172] B. Ren, H. Cam, Y. Takahashi, T. Volkert, J. Terragni, R.A. Young, B.D. Dynlacht, E2F integrates cell cycle progression with DNA repair, replication, and G2/M checkpoints, Genes Dev. 16 (2002) 245–256. https://doi.org/10.1101/gad.949802.

[173] W.F. Anderson, N. Chatterjee, W.B. Ershler, O.W. Brawley, Estrogen receptor breast cancer phenotypes in the Surveillance, Epidemiology, and End Results database, Breast Cancer Res Treat. 76 (2002) 27–36. https://doi.org/10.1023/a:1020299707510.

[174] P. Thangavel, A. Puga-Olguín, J.F. Rodríguez-Landa, R.C. Zepeda, Genistein as potential therapeutic candidate for menopausal symptoms and other related diseases, Molecules. 24 (2019). https://doi.org/10.3390/molecules24213892.

[175] L. Khaodhiar, H.A. Ricciotti, L. Li, W. Pan, M. Schickel, J. Zhou, G.L. Blackburn, Daidzein-rich isoflavone aglycones are potentially effective in reducing hot flashes in menopausal women, Menopause. 15 (2008) 125–132. https://doi.org/10.1097/gme.0b013e31805c035b.

[176] D.M. Tit, S. Bungau, C. Iovan, D.C.N. Cseppento, L. Endres, C. Sava, A.M. Sabau, G. Furau, C. Furau, Effects of the hormone replacement therapy and of soy isoflavones on bone resorption in postmenopause, J Clin Med. 7 (2018). https://doi.org/10.3390/jcm7100297.

[177] K.D. Dykstra, L. Guo, E.T. Birzin, W. Chan, Y.T. Yang, E.C. Hayes, C.A. DaSilva, L.Y. Pai, R.T. Mosley, B. Kraker, P.M.D. Fitzgerald, F. DiNinno, S.P. Rohrer, J.M. Schaeffer, M.L. Hammond, Estrogen receptor ligands. Part 16: 2-Aryl indoles as highly subtype selective ligands for ERα, Bioorg Med Chem Lett. 17 (2007) 2322–2328. https://doi.org/10.1016/j.bmcl.2007.01.054.

[178] S. Eiler, M. Gangloff, S. Duclaud, D. Moras, M. Ruff, Overexpression, purification, and crystal structure of native erα lbd, Protein Expr Purif. 22 (2001) 165–173. https://doi.org/10.1006/prep.2001.1409.

[179] M. Muchtaridi, H.N. Syahidah, A. Subarnas, M. Yusuf, S.D. Bryant, T. Langer, Molecular docking and 3D-pharmacophore modeling to study the interactions of chalcone derivatives with estrogen receptor alpha, Pharmaceuticals. 10 (2017) 1–12. https://doi.org/10.3390/ph10040081.

[180] N. V. Puranik, P. Srivastava, G. Bhatt, D.J.S. John Mary, A.M. Limaye, J. Sivaraman, Determination and analysis of agonist and antagonist potential of naturally occurring flavonoids for estrogen receptor (ERα) by various parameters and molecular modelling approach, Sci Rep. 9 (2019) 1–11. https://doi.org/10.1038/s41598-019-43768-5.

[181] G. Reid, M.R. Hübner, R. Métivier, H. Brand, S. Denger, D. Manu, J. Beaudouin, J. Ellenberg, F. Gannon, Cyclic, proteasome-mediated turnover of unliganded and liganded ERα on responsive promoters is an integral feature of estrogen signaling, Mol Cell. 11 (2003) 695–707. https://doi.org/10.1016/S1097-2765(03)00090-X.

[182] A.M. Nardulli, G.L. Greene, B.W. O'Malley, B.S. Katzenellenbogen, Regulation of progesterone receptor messenger ribonucleic acid and protein levels in MCF-7 cells by estradiol: Analysis of estrogen's effect on progesterone receptor synthesis and degradation, Endocrinology. 122 (1988) 935–944. https://doi.org/10.1210/endo-122-3-935.

[183] J. Matthews, B. Wihlén, J. Thomsen, J.-Å. Gustafsson, Aryl hydrocarbon receptor-mediated transcription: ligand-dependent recruitment of estrogen receptor α to 2,3,7,8-tetrachlorodibenzo- p -dioxin-responsive promoters , Mol Cell Biol. 25 (2005) 5317–5328. https://doi.org/10.1128/mcb.25.13.5317-5328.2005.

[184] S. Lecomte, F. Demay, F. Ferrière, F. Pakdel, Phytochemicals targeting estrogen receptors: Beneficial rather than adverse effects?, Int J Mol Sci. 18 (2017) 1–19. https://doi.org/10.3390/ijms18071381.

[185] L. Pinzi, G. Rastelli, Molecular docking: Shifting paradigms in drug discovery, Int J Mol Sci. 20 (2019). https://doi.org/10.3390/ijms20184331.

[186] K. Danao, D. Nandurkar, V. Rokde, R. Shivhare, U. Mahajan, Molecular Docking: Metamorphosis in Drug Discovery, (2023). https://doi.org/10.5772/intechopen.105972.

[187] M. Muchtaridi, M. Yusuf, A. Diantini, S.B. Choi, B.O. Al-Najjar, J. V. Manurung, A. Subarnas, T.H. Achmad, S.R. Wardhani, H.A. Wahab, Potential activity of fevicordin-a from Phaleria macrocarpa (Scheff) Boerl. seeds as estrogen receptor antagonist based on cytotoxicity and molecular modelling studies, Int J Mol Sci. 15 (2014) 7225–7249. https://doi.org/10.3390/ijms15057225.

[188] G. Türker, The effect of heavy metals on preterm mortality and morbidity, handbook of fertility: nutrition, diet, lifestyle, and reproductive health. (2015) 45–59. https://doi.org/10.1016/B978-0-12-800872-0.00005-6.

[189] V. Rothhammer, F.J. Quintana, The Aryl hydrocarbon receptor: an environmental sensor integrating immune responses in health and disease, Nat Rev Immunol. 19 (2019) 184–197. https://doi.org/10.1038/s41577-019-0125-8.

[190] M.F. Torti, F. Giovannoni, F.J. Quintana, C.C. García, The aryl hydrocarbon receptor as a modulator of anti-viral immunity, Front Immunol. 12 (2021) 624293. https://doi.org/10.3389/fimmu.2021.624293.

[191] C. Gutiérrez-Vázquez, F.J. Quintana, Regulation of the immune response by the aryl hydrocarbon receptor, Immunity. 48 (2018) 19–33. https://doi.org/10.1016/j.immuni.2017.12.012.

[192] C. Rejano-Gordillo, A. Ordiales-Talavero, A. Nacarino-Palma, J.M. Merino, F.J. González-Rico, P.M. Fernández-Salguero, Aryl hydrocarbon receptor: from homeostasis to tumor progression, Front Cell Dev Biol. 10 (2022). https://doi.org/10.3389/fcell.2022.884004.

[193] A.K. Tamrakar, N. Jaiswal, P.P. Yadav, R. Maurya, A.K. Srivastava, Pongamol from *Pongamia pinnata* stimulates glucose uptake by increasing surface GLUT4 level in skeletal muscle cells, Mol Cell Endocrinol. 339 (2011) 98–104. https://doi.org/10.1016/j.mce.2011.03.023.

[194] S. Giani Tagliabue, S.C. Faber, S. Motta, M.S. Denison, L. Bonati, Modeling the binding of diverse ligands within the Aryl hydrocarbon receptor ligand binding domain, Sci Rep. 9 (2019) 10693. https://doi.org/10.1038/s41598-019-47138-z.

[195] A. Perkins, J.L. Phillips, N.I. Kerkvliet, R.L. Tanguay, G.H. Perdew, S.K. Kolluri, W.H. Bisson, A structural switch between agonist and antagonist bound conformations for a ligand-optimized model of the human aryl hydrocarbon

receptor ligand binding domain, Biology (Basel). 3 (2014) 645–669. https://doi.org/10.3390/biology3040645.

[196] B. Zhao, D.E. DeGroot, A. Hayashi, G. He, M.S. Denison, Ch223191 is a ligand-selective antagonist of the Ah (dioxin) receptor, Toxicological Sciences. 117 (2010) 393–403. https://doi.org/10.1093/toxsci/kfq217.

[197] J. Zhang, Y. Xie, Q. Fan, C. Wang, Effects of karanjin on dimethylhydrazine induced colon carcinoma and aberrant crypt foci are facilitated by alteration of the p53/Bcl2/BAX pathway for apoptosis, Biotechnic and Histochemistry. 96 (2021) 202–212. https://doi.org/10.1080/10520295.2020.1781258.

Appendix

Appendix I

Isolation and purification of bioactive compounds:

To isolate compounds, crude extracts (2g) were subjected to column chromatography on silica gel (80g, 60-120 mesh, 3×60cm glass column) and compounds were eluted gradually with increasing polarity of hexane-ethyl acetate two phase eluent system. The purity of eluted fractions was subsequently checked by thin layer chromatography (TLC) with hexane-ethyl acetate mobile system. Further through series of crystallization and recrystallization with solvents Karanjin crystals were obtained. These crystals were subjected to analytical HPLC, NMR, HRMS for evaluating the purity of Karanjin.

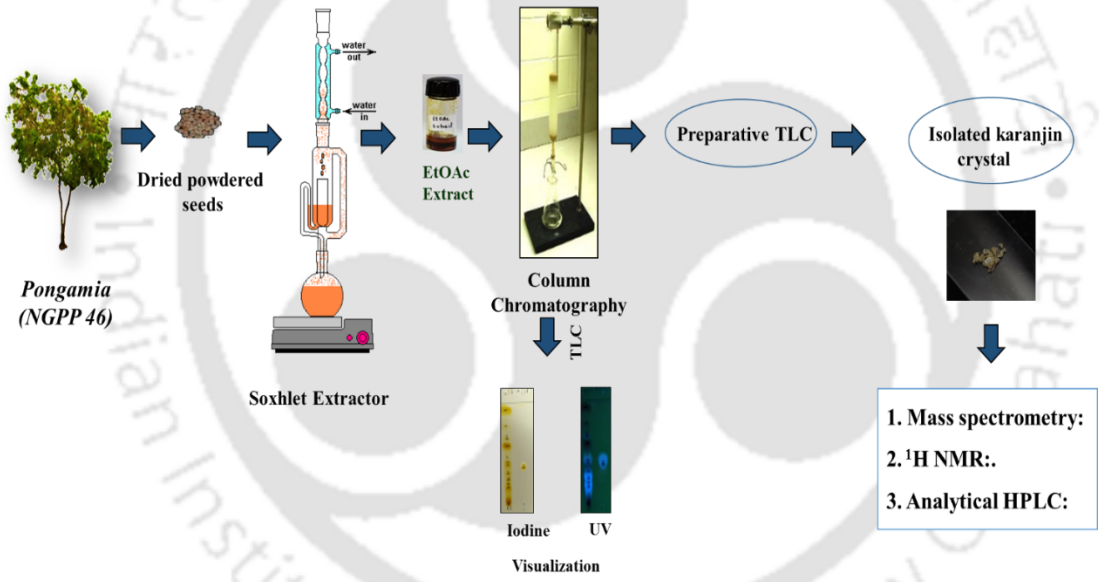


Fig. 1A1: Schematic diagram for isolation and characterization of karanjin.

Certificate of analysis for Karanjin from Yucca Enterprises

YUCCA ENTERPRISES
A-246, Antop Hill warehousing Co.,
Barkat Ali Naka, VIT Marg,
WADALA (E), MUMBAI 400 037.
Phone : 9892435011,
Email : yuccaenterprises@yahoo.com

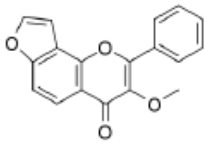
Certificate of Analysis

KARANGIN

Batch : Yucca/KG/2019/04/21 **Date** : 22.04.2019

Chemical Name : 3-Methoxy-2-phenyl-4H-furo(2,3-h)(1)benzopyran-4-one; Karanjin
Chemical Formula : $C_{18}H_{12}O_4$
Chemical Family : Flavonoid
Molecular Weight : 292.29
CAS Number : 521-88-0
Description : White powder with a faint yellow cast

Molecular structure :



Intended use : For laboratory use only
Solubility : Soluble in ethyl and methyl alcohol, benzene and petroleum ether.
Storage condition : +2°C to +8°C, Protect from light

Analytical test :

S. No.	Test	Result
1	Test for identity (by TLC)	Complies
2	Purity test (by HPLC)	≥ 90 %

For Yucca Enterprises

Characterization of karanjin

Karanjin isolated earlier from *P. pinnata* seed oil, and standard purchased from Yucca Enterprises was subjected to HRMS, HPLC and NMR for validation.

High-resolution mass spectrometry (HRMS)

About 1 mg of karanjin was dissolved in 1 mL of HPLC grade acetonitrile and filtered. The samples were injected via auto-sampler to record the mass spectra using standard protocol. Mass spectrum was recorded on an Agilent Technologies 6520 Accurate Mass Q-TOF LC/MS (Agilent, CA, USA) in ESI +ive ion mode with a flow rate of 0.2 mL/min over a total run time of 30 min. HRMS data for karanjin isolated in-house and that procured from Yucca Enterprises are shown in Fig. 1A2 and Fig. 1A3.

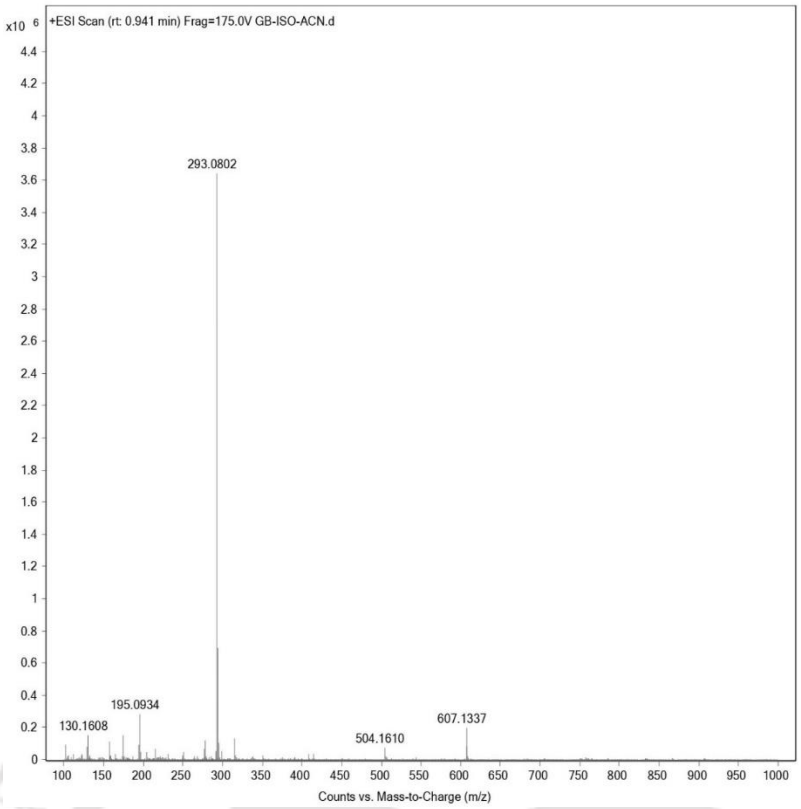


Fig. 1A2. Mass spectrum of isolated karanjin. The *m/z* of the compound in the positive mode $[M + H]^+$ is 293.0802.

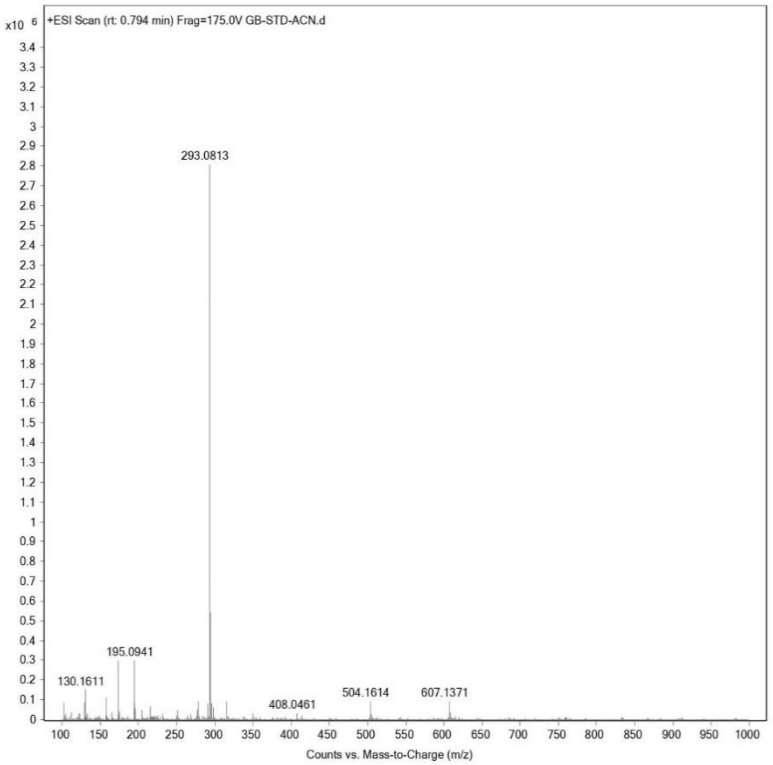


Fig. 1A3. Mass spectrum of karanjin from Yucca enterprise. The *m/z* of the compound in the positive mode $[M + H]^+$ is 293.0813.

Purification and detection by HPLC

To assess the purity of karanjin, high performance liquid chromatography analysis was performed using analytical HPLC system (Shimdazu, Japan) with a degassing unit (DGU 20ASR) and autosampler (SIL 20AHT) and liquid chromatogram (LC 20AD) system. The chromatographic separation was performed using reverse-phase C-18 column (Luna ®, 5µM, 100 Å, 250×4.6 mm). Karanjin isolated in-house or standard (Yucca enterprises, India) stock of 1 mg/mL was prepared in methanol: water: acetic acid (85:13.5:1.5) filtered through a 0.45 µm syringe filter. 10 µl of filtered samples were then injected and analyzed in isocratic mode methanol: water: acetic acid (85:13.5:1.5) for 20 mins at a flow rate of 1.0 mL/min. Karanjin was detected using a UV/Vis detector (SPD-20A) at a set wavelength of 264 nm. HPLC data for karanjin isolated in-house and standard procured from Yucca Enterprises are shown in Fig. 1A4 and Fig. 1A5.

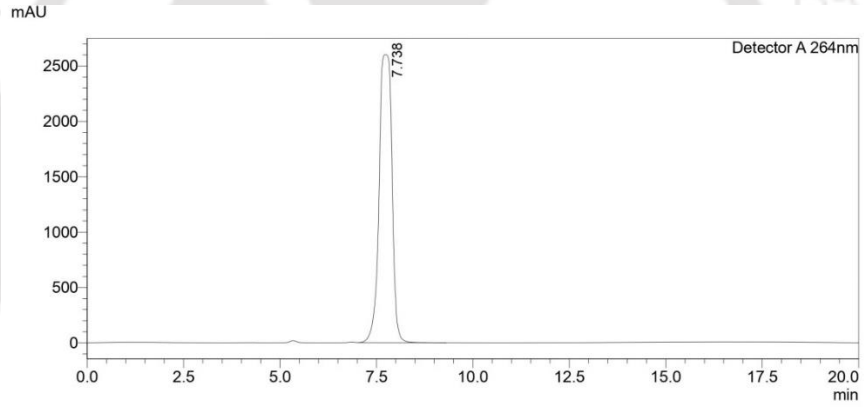


Fig. 1A4. HPLC chromatogram of karanjin purified from *P. pinnata*. The compound showed a retention time of 7.738 min and greater than 98% purity.

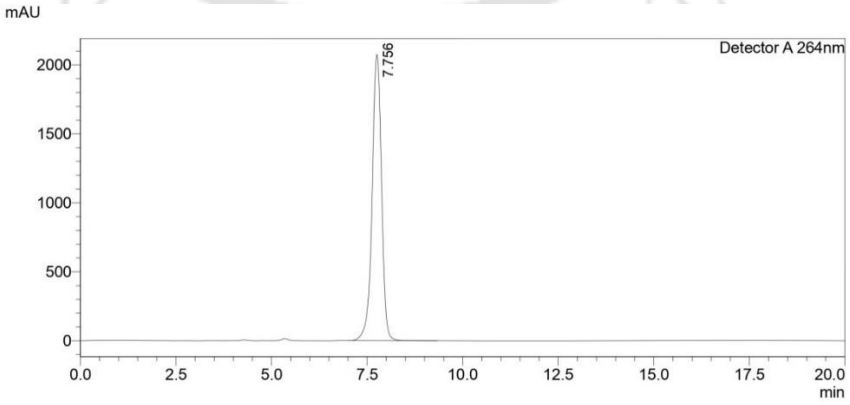
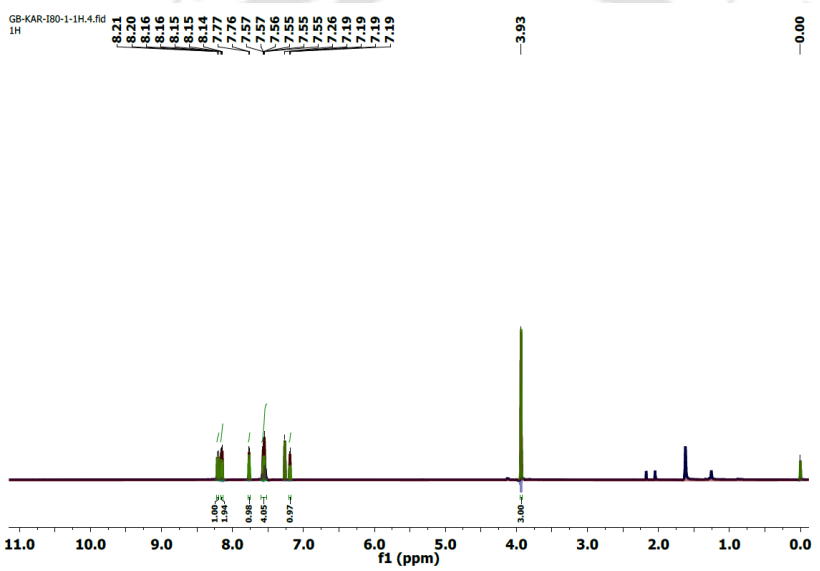


Fig. 1A5. HPLC chromatogram of karanjin standard from Yucca Enterprises. The compound showed a retention time of 7.756 min, and greater than 98% purity.

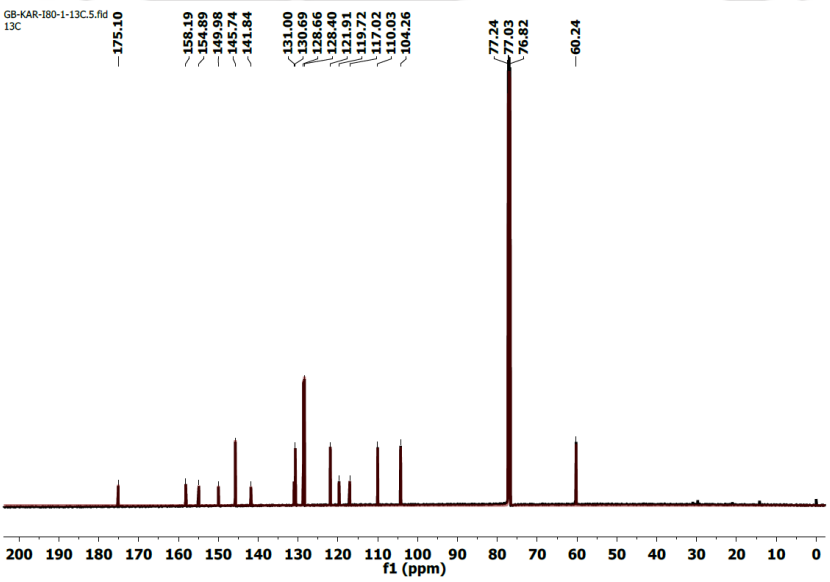
Nuclear magnetic resonance

One-dimension spectral analysis was carried out by dissolving 10 mg of isolated and standard karanjin from Yucca Enterprises in 600 µl $CDCl_3$. ^{1}H NMR spectra at 600 MHz and ^{13}C spectra at 150 MHz were acquired on a Bruker Ascend ^{TM} 600 (Bruker BioSpin AG, Switzerland) with a central peak of the $CDCl_3$ triplet (δ 77.04 ppm) as the internal standard along with tetramethyl silane (TMS) as an internal reference. ^{1}H and ^{13}C spectral chemical shifts and coupling constants are expressed in δ and Hz, respectively. NMR was controlled by the software TopSpin 2.1. NMR data for karanjin isolated in-house and standard procured from Yucca Enterprises are shown in Fig. 1A6A, B, C and D.

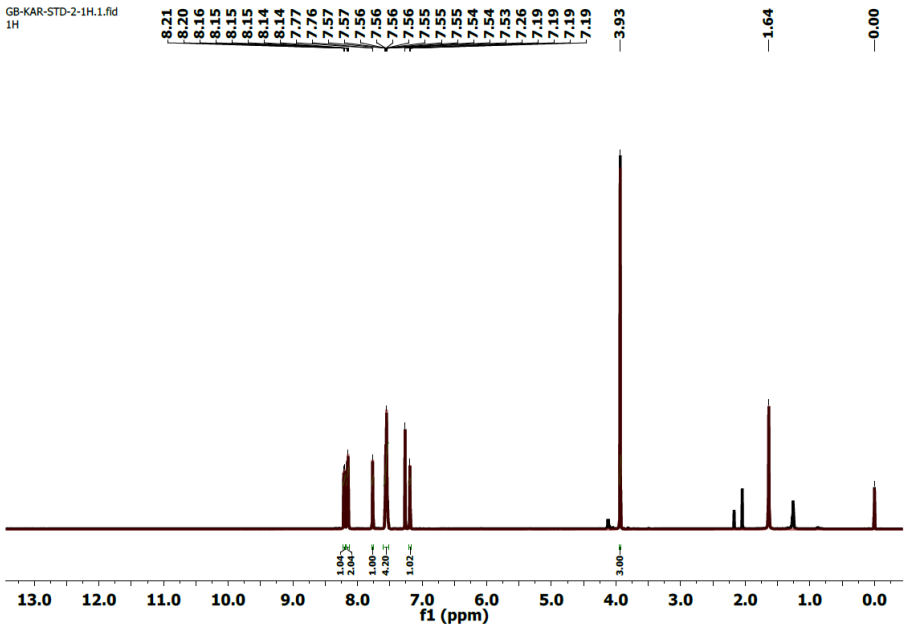
A.



B.



C.



D.

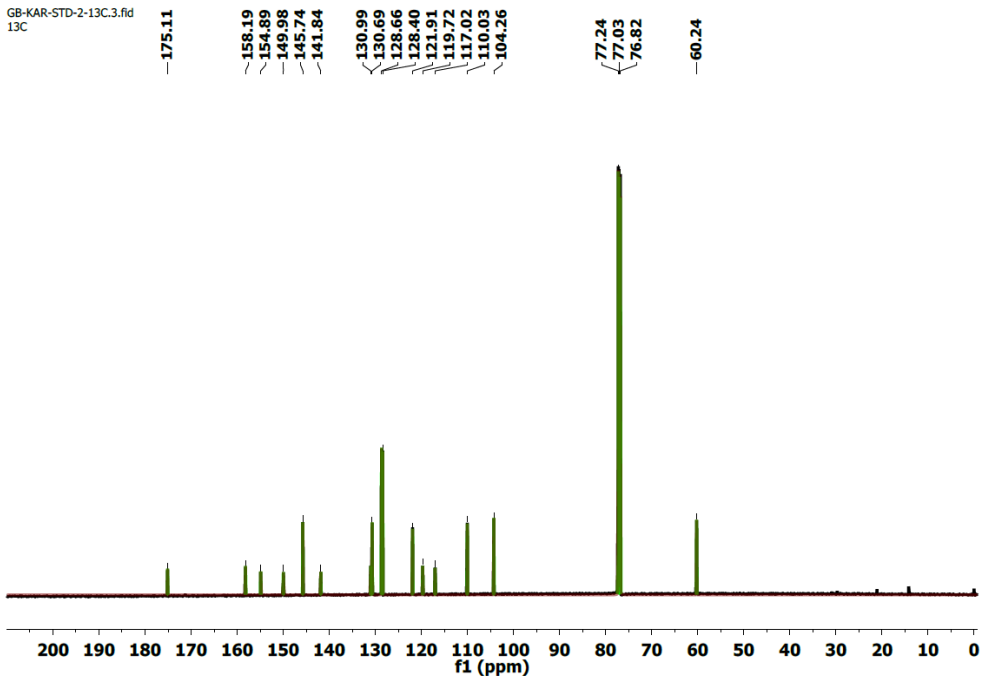


Fig. 1A6. ^{1}H NMR (A, C), ^{13}C NMR spectra (B, D) for karanjin isolated in-house (A, B) and standard purchased from Yucca Enterprises (C, D).

Appendix II

Table 2A1: Quality measures of RNA samples used for RNA-seq

Sr. No	Sample Name	Conc (ng/µl)	Final Volume (µl)	Total Amount (µg)	RIN Value	rRNA Ratio
1	G1	381.066	19	7.24	9.6	1.9
2	G2	311.152	21	6.53	9.9	2.2
3	G3	511.198	19	9.71	9.8	2.4
4	K1	438.895	19	8.33	9.9	2.1
5	K2	383.067	21	8.04	9.8	2.2
6	K3	421.955	21	8.86	10	2.1

Note: G1, G2, G3- three biological replicates of total RNA from MCF-7 cells treated with vehicle (DMSO)

K1, K2, K3- three biological replicates of total RNA from MCF-7 cells treated with 10 µM karanjin.

Table 2A2: Sequencing read statistics of raw data using Illumina Novaseq 6000.

Sr. No	Sample Name	Total read bases[@]	Total reads[@]	GC (%)	AT (%)	Q20(%)	Q30(%)
1	G1	3845557830	38074830	50.26	49.74	98.39	95.34
2	G2	3741880522	37048322	49.97	50.03	98.3	95.12
3	G3	4551822752	45067552	50.35	49.65	98.44	95.45
4	K1	3896588484	38580084	51.16	48.84	98.5	95.64
5	K2	4301862094	42592694	49.84	50.16	98.51	95.55
6	K3	4175723396	41343796	50.19	49.81	98.37	95.29

Note: [@]Paired-end reads

G1, G2, G3- correspond to samples from MCF-7 cells treated with vehicle (DMSO)

K1, K2, K3- correspond to samples from MCF-7 cells treated with 10 µM karanjin.

Table 2A3: Read statistics before and after trimming.

Sr. No	Sample	File Names (Paired-end)	Total Reads		% Data lost
			Before trimming	After trimming	
1	G1	G1_1	19037415	18876043	0.8477
2		G1_2	19037415	18876043	
3	G2	G2_1	18524161	18384899	0.7518
4		G2_2	18524161	18384899	
5	G3	G3_1	22533776	22366336	0.7431
6		G3_2	22533776	22366336	
7	K1	K1_1	19290042	19149767	0.7272
8		K1_2	19290042	19149767	
9	K2	K2_1	21296347	21128011	0.7904
10		K2_2	21296347	21128011	

11	K3	K3_1	20671898	20501108	0.8262
12		K3_2	20671898	20501108	

Note: G1, G2, G3- correspond to samples from MCF-7 cells treated with vehicle (DMSO)

K1, K2, K3- correspond to samples from MCF-7 cells treated with 10 µM karanjin.

Table 2A4: Quality measures of RNA samples used for RNA-seq of T47D

Sr. No	Sample Name	Conc. (ng/µl)	Final Volume (µl)	Total Amount (µg)	RIN Value	rRNA Ratio
Vehicle (0.1% DMSO)	TD1	980.9	15	14.714	9.9	1.98
	TD2	738.4	15	11.076	10	1.91
	TD3	782	15	11.73	9.9	1.98
	TD4	812.3	15	12.185	9.9	1.99
10µM karanjin	TK1	920.7	15	13.811	9.8	1.95
	TK2	924.7	15	13.871	9.8	1.95
	TK3	923.3	15	13.85	10	1.92
	TK4	1217.3	15	18.26	10	1.99

Note: TD1, TD2, TD3 and TD4- four biological replicates of total RNA from T47D cells treated with vehicle (DMSO)
TK1, TK2, TK3, and TK4- four biological replicates of total RNA from T47D cells treated with 10 µM karanjin.

Table 2A5: Read statistics before and after trimming.

Sr. No	Sample	File Names (Paired-end)	Total Reads		% Data lost
			Before trimming	After trimming	
Vehicle (0.1% DMSO)	TD1	TD1_S43_R1_001	19148360	12349879	3.25
		TD1_S43_R2_001	19148360	12349879	
	TD2	TD2_S44_R1_001	31351842	21684783	1.61
		TD2_S44_R2_001	31351842	21684783	
	TD3	TD3_S45_R1_001	23039100	16349790	1.48
		TD3_S45_R2_001	23039100	16349790	
	TD4	TD4_S46_R1_001	23406290	15525887	2.14
		TD4_S46_R2_001	23406290	15525887	
10µM karanjin	TK1	TK1_S47_R1_001	24469569	16750232	2.87
		TK1_S47_R2_001	24469569	16750232	
	TK2	TK2_S48_R1_001	18587752	13946182	0.99
		TK2_S48_R2_001	18587752	13946182	
	TK3	TK3_S49_R1_001	23065861	16546712	1.82
		TK3_S49_R2_001	23065861	16546712	
	TK4	TK4_S49_R1_001	21284080	16348445	0.99
		TK4_S49_R2_001	21284080	16348445	

Note: TD1, TD2, TD3, TD4- correspond to samples from T47D cells treated with vehicle (DMSO)
TK1, TK2, TK3, TK4- correspond to samples from T47D cells treated with 10 µM karanjin.

Table 2A6: List of common genes regulated by tamoxifen and karanjin. Red cells represent negative log_2FC values implying downregulation of the gene. Green cells represent positive log_2FC implying upregulation of the gene.

Gene	Tamoxifen	Karanjin
GREB1	-2.55735	0.49515
FOSL2	0.93781	0.90447
IER3	0.84203	0.77485
TMEM120B	0.92590	0.98744
ESR1	-0.47007	-0.79839
RET	0.78069	-1.48716
PHLDA1	0.83261	1.29955
ANXA9	0.72957	-0.83170
SLC9A3R1	0.42184	-0.37965
THBS1	-0.42953	-0.44266
SLC7A5	0.52527	2.88241
IGFBP4	-0.56627	-0.35077
DCDC2	0.55065	-0.49863
STARD10	0.39895	-0.24456
IGFBP5	0.61801	-1.01068
TMEM64	0.70083	-0.36397
TFF1	-0.41660	0.69688
KRT18	-0.22320	-0.42981
IGF1R	0.27579	0.55780
WISP2	0.41743	0.72403
GPER1	-0.40173	-1.10271
AGR3	1.42952	-0.58730
GFRA1	0.14759	-0.44755
PCSK6	-0.31629	-0.52841
SOX4	-0.21216	0.22894
RUNX1	-0.33836	0.62229
AR	0.34493	-0.48380
CHRD	-0.74749	-1.14409
CAMK2N1	-0.33133	-0.31221
DEGS2	0.28426	-0.76916
LAMC2	0.70334	0.62629
BAG3	0.15985	-0.37680

Gene	Tamoxifen	Karanjin
ACSS1	0.84822	-1.11254
ABCG1	-0.84740	-0.51581
SLC39A6	-0.21267	-0.45094
SYTL5	0.47828	-0.50104
HR	0.70226	0.62160
TXNIP	0.35485	-0.30567
CLIC3	0.58364	-0.44674
PREX1	0.22564	-0.53848
PLCD3	-0.51035	-0.24346
LOXL1	0.46926	0.49477
SLC24A3	-0.56067	-0.56601
KRT8	-0.16156	-0.46869
STC1	0.41809	0.59671
TBC1D9	-0.22039	-0.58733
PCDH7	-0.46997	-1.00301
SP1	0.27299	-0.30004
TXNRD1	0.33188	0.82848
SCNN1A	0.49363	-0.62562
SPTBN1	-0.26363	0.31114
FLNB	0.23159	-0.31630
KCNC3	-0.36726	-0.72312
CD9	-0.15108	-0.35416
CA2	0.26254	-0.69021
ACHE	-0.34861	-0.85178
TINCR	0.25000	0.42035
NT5E	-0.16356	-1.14029
LACTB	-0.30758	0.49780
RBL2	0.21375	-0.40112
TMEM150C	0.20269	-0.45499
GDPD3	-0.36800	-0.65505
TUBB3	-0.17962	0.77720

Gene	Tamoxifen	Karanjin
LTBP3	-0.29467	-0.45979
CERS2	0.16688	-0.48687
SH3BGRL	0.26594	-0.36174
SEMA4C	-0.20818	-0.43081
PITPNM2	0.73282	1.34756
KRT19	0.14678	-0.44848
DDIT4	0.34856	0.32507
ST3GAL1	0.23705	0.85668
DSP	-0.14198	0.37818
GPRC5A	-0.22756	-0.29944
SYNE2	-0.28570	0.45358
GREB1L	0.55469	-0.57718
MT2A	-0.64036	-0.59522
ALDH3B2	-0.43826	0.46571
CYP1B1	-0.23035	2.72335
ABHD2	-0.23062	-0.34981
SPEG	0.63428	-0.67022
CA12	0.17531	0.40489
CLSTN2	-0.63517	0.66979
KITLG	-0.54133	-0.77083
TPCN1	0.28573	0.39156
IDH1	0.25064	-0.26904
MAGED2	0.17328	-0.38337
LMCD1	0.44012	1.31182
ADCY5	0.35298	-0.63689
SIM1	1.26923	-1.44562
LDHA	0.15873	0.41152
EPS8L1	-0.25663	-0.43861
DIO1	1.98040	-1.40182
HNRNPAB	-0.12155	0.26342
PMP22	-0.23010	-0.92279

Table 2A6: List of common genes regulated by estrogen and karanjin. Red cells represent negative log_2FC values implying downregulation of the gene. Green cells represent positive log_2FC implying upregulation of the gene.

Gene	Estrogen	Karanjin
CYP1A1	-1.75158	6.17733
SLC7A5	1.688633	2.88241
STC2	0.760613	1.1071
LAMA3	1.378145	1.84031
TIPARP	1.127496	2.05882
SLC3A2	0.572105	0.97688
PHLDA1	0.434351	1.29955
TSKU	0.45057	1.10389
ALDH1A3	-1.50793	2.46633
GDF15	-0.60991	1.31592
LMCD1	-0.69094	1.31182
RND1	-0.86829	1.56001
FOSL2	0.741212	0.90447
PMP22	-1.06477	-0.9228
RUNX2	0.477098	1.43623
RET	1.430755	-1.4872
IL1R1	-2.17744	-1.1697
ANXA9	-0.39428	-0.8317
ADAMTS19	-0.62557	-1.1133
SLC35C1	-0.79732	-1.1612
SLC7A11	0.983892	2.43785
TXNRD1	0.277647	0.82848
TMEM45B	-1.0134	1.18506
AREG	3.124479	1.48712
PSCA	-2.22258	-0.9852
CRY2	-0.81375	-0.8571
CYSRT1	-0.80867	0.96535
ATOH8	-2.11823	1.42633
ST3GAL1	-0.58697	0.85668
BMF	-2.44599	0.94572
GPER1	-0.97146	-1.1027
CD44	1.116591	0.88285
TMEM120B	0.873404	0.98744
KCNJ8	-1.05771	-1.2645
SLC12A2	-0.6015	-0.6495
NPNT	-0.79012	-0.6665
ESR1	-0.47548	-0.7984

Gene	Estrogen	Karanjin
ABCC3	-0.83471	0.32744
MRPL3	0.339822	0.32502
SLC2A11	-0.46149	0.55357
FLOT1	-0.25824	-0.3573
FGF13	-1.08767	-0.6125
RUVBL1	0.578492	0.33547
ATP2C2	-0.69536	-0.4168
BMP4	-0.80989	-0.557
FAM198B	0.859587	0.95393
PLEKHA7	-0.45517	-0.3606
ECE1	0.873938	-0.4499
SCUBE2	-0.26608	0.57498
APP	-0.22446	-0.2646
KIAA0513	-1.1095	-0.3224
LARS	0.299862	0.34716
LAMC2	0.941714	0.62629
WWOX	-0.42486	-0.3957
IFT140	-0.56394	-0.456
MARCKS	-0.55908	-0.3821
PCDH7	-0.77329	-1.003
CRIM1	-0.51498	-0.4369
EPB41L2	-0.53443	-0.4114
HNRNPH3	0.417431	0.30371
ARHGEF37	-1.06151	-0.6431
DBP	-0.95665	-0.5677
MEST	0.392134	0.30831
TTK	0.342633	0.90583
RNF207	-1.41223	-0.5321
GFRA1	0.422923	-0.4475
RAB30	-0.43571	-0.4508
TSPAN9	-0.57789	-0.3994
ABCA1	-0.89092	-0.4276
ABCE1	0.668304	0.4938
GSTM3	-0.28895	0.26429
COL12A1	0.599568	-1.1044
CEP55	0.608627	0.56906
S100A16	-0.50954	-0.28

Gene	Estrogen	Karanjin
SPATS2L	0.273344	0.40452
CRISPLD2	-0.52641	0.43058
CDCA8	0.521254	0.42469
BOD1	0.277616	0.31512
ZG16B	-0.69532	-0.4268
NBEAL2	-0.44377	-0.3247
AIF1L	-0.30348	-0.3125
LYAR	0.873111	0.5951
CKB	-0.62952	-0.434
PNO1	0.457455	0.3697
NOL4L	-0.32054	0.36801
TFRC	0.434736	0.29958
NELL2	-2.06886	-1.4076
TCF7L1	-1.28343	-0.7051
PACSIN1	-0.68216	-0.4272
HSPE1	0.578903	0.40782
SPTBN1	0.29014	0.31114
SPINT1	-0.41111	-0.2975
KLHL5	-0.50183	0.53935
EPHA7	-0.57402	-0.479
HDAC5	-0.8136	-0.3866
IGSF1	3.329681	0.50214
ALDH4A1	-0.81534	-0.3309
MEX3A	-0.36707	0.41394
GABBR2	-1.82556	-1.0885
LIF	0.702269	1.24899
GPR157	-0.65938	0.33505
DNTTIP2	0.357915	0.89487
ABHD11	-0.40919	-0.3057
WDR43	0.545842	0.30751
HSP90AB1	0.467331	0.22645
KIF2C	0.376307	0.37724
RNF223	0.745848	0.36729
CERS4	-0.43523	-0.3561
GPRC5C	-0.60541	0.27944
KSR1	0.597607	0.65579
TXN	0.483713	0.2561

ACSS1	-0.88194	-1.1125
PPP1R3C	-0.91188	-0.6368
NCAM2	-0.45638	-0.6753
NDRG1	-0.59345	-1.0034
BMP7	-0.76204	-0.6021
LDLRAD4	-0.76891	-0.729
SYNJ2BP	-0.44679	-0.765
SPTSSA	0.352374	0.6814
CABLES1	-0.27484	0.60923
LRP5	-0.47271	0.52774
DNMBP	0.727415	0.66431
WISP2	1.077031	0.72403
TRERF1	-0.95702	-0.8596
TFF1	2.298919	0.69688
TUBA1C	0.421166	0.58071
VTCN1	-0.79609	1.69128
CERS2	-0.51071	-0.4869
PP14571	-0.95377	-0.9123
TSC22D3	-0.98198	-0.4916
RAB3D	-1.00866	-0.6995
NRBP1	-0.54008	-0.4779
FN1	0.46273	-0.7339
CHRD	-2.97177	-1.1441
YPEL2	-1.31399	-0.5608
DKK1	-0.69777	-0.8622
ISOC1	0.34044	-0.5913
INPP4A	-0.47411	0.55204
SMIM5	-0.93201	0.90516
SIX4	-0.52164	-0.5818
PREX1	-0.17437	-0.5385
SPDEF	-0.54518	-0.5206
ACKR3	-0.45018	-0.6248
KRT18	-0.36464	-0.4298
UBE2C	0.487766	0.61816
PRPS1	0.510245	0.58894
STC1	0.452835	0.59671
PFKFB3	-0.31107	-0.5263
HMCN1	-0.41572	-0.4652
KIAA1324	0.366045	-0.6869
GOLGB1	-0.38362	0.90165
RAP1GAP	-0.96451	0.63543

ATP2A3	-0.88584	-0.3134
PLLP	-1.05811	-0.5734
TLE3	0.36105	-0.3248
DLC1	0.471079	-0.7455
HCAR1	-1.54447	-0.456
SH3BGRL	-0.51166	-0.3617
CTDSP1	-0.37386	-0.3132
NUP107	0.524756	0.40631
NAA20	0.239287	0.3007
SLC9A3R2	-0.49305	0.36781
MTHFR	-0.91668	-0.4645
RASEF	-0.51392	-0.3296
BIK	-1.02725	-0.5036
COMT	-0.59362	-0.2784
GPC4	-0.35121	-0.4878
DKC1	0.668535	0.33343
GRN	-0.4755	-0.2529
HSPA8	0.551927	0.25122
NATD1	-1.12318	-0.4546
OSGIN1	0.810572	0.38904
BUB1B	0.528073	0.57544
SCN1B	-0.36402	-0.3802
NCAPG	0.460454	0.85167
DDIT4	-1.08202	0.32507
AQP3	-0.61538	0.32515
GCLM	0.599053	0.43509
KIF14	0.262833	1.17006
AEN	0.742537	0.32527
TBC1D9	-0.72279	-0.5873
TUBB4B	0.386359	0.46056
TMTC2	-1.02662	0.53154
NT5E	-1.66106	-1.1403
SLIT2	-1.18843	-1.1439
KIAA1161	-0.42748	-0.4465
CORO2A	-0.87547	-0.474
RAB26	-1.20989	-0.7085
CD9	-0.25947	-0.3542
CKS2	0.448754	0.49821
LDHA	0.910461	0.41152
SNHG7	-0.25145	0.38731
CA2	0.701731	-0.6902

FAM83B	-0.70966	-0.4364
ZBTB42	-0.5043	-0.3179
ARHGAP32	-0.41477	-0.3462
MRTO4	0.611522	0.34243
G6PD	0.266311	0.29508
SREBF1	-0.2273	-0.3913
MID2	-0.51096	-0.3384
MAPK3	-0.28646	-0.2511
SFXN2	2.033102	-0.36
CAV2	0.441529	0.39542
CAMK2N1	-1.07375	-0.3122
WDYHV1	0.386264	0.3312
NOLC1	0.591656	0.29979
SNED1	-0.41338	-0.4658
IDH1	-0.28869	-0.269
KIFC3	-1.406	-0.434
IGFBP4	1.494262	-0.3508
TMC4	-0.93905	-0.4205
SIX1	-0.60364	-0.5262
TACC3	0.445793	0.31884
PIM1	0.687684	0.49835
STRBP	0.281979	0.26954
HPRT1	0.871493	0.33406
RABGGTB	0.500559	0.36448
FLNB	0.582385	-0.3163
ALDH3B2	-0.36838	0.46571
SOX4	-0.60067	0.22894
HMBS	0.514845	0.28517
ABCA12	0.70634	-0.6741
RRS1	0.521749	0.35234
FARSB	0.620703	0.32879
PVT1	0.38793	0.40796
BTG2	-1.52902	-0.4611
RHOC	-0.39684	-0.3705
KRT81	-0.94708	0.71896
TM7SF2	-0.473	-0.3633
SAMD4A	-0.47248	0.50019
JAG1	-0.40645	-0.9247
ZFP36L2	-0.19128	-0.6121
TPX2	0.392047	0.33314
NHP2	0.283017	0.40866

IGFBP5	0.583799	-1.0107
TSPAN1	-1.07738	-0.7133
KRT8	-0.28676	-0.4687
FZD8	-1.31306	-0.8005
ARNT2	-0.9015	-0.8448
TGFB2	-1.23017	-0.9265
RNF224	-1.3355	1.1133
PCDH10	-0.98857	-0.7052
MYC	0.66237	1.00247
FAM161B	-0.64168	-0.8302
LXN	-0.93553	0.70516
CXCR4	-0.48187	-1.5061
RARA	0.509096	-0.5299
TP53INP1	-1.81276	-0.4546
CENPE	0.288627	1.96979
C1orf21	-0.43756	-0.5387
PFKP	0.422214	0.40237
NFRKB	-0.38334	0.49256
ADCY5	-0.39958	-0.6369
KITLG	-1.08387	-0.7708
KPNA2	0.497823	0.49408
AURKA	0.479986	0.44109
SELENBP1	-1.10957	-0.5376
GDPD3	-0.86798	-0.655
CTSH	-0.75853	-0.4964
SLC34A3	-1.01086	1.03868
PNPLA7	-1.05749	-0.8336
CCNB1	0.457928	0.4067
CYB5R1	-0.31418	-0.3572
VLDLR	-0.51264	-0.6551
PCSK6	-0.93851	-0.5284
WWP1	-0.2666	-0.4303
CA12	0.692127	0.40489
SYTL5	1.778343	-0.501
TOM1L2	-1.02068	-0.4386
DCDC2	-0.48971	-0.4986
TUBB	0.56829	0.39704
NTN4	-1.11339	0.49717
TGFB3	-1.71313	-0.8282
ARPIN	-1.50722	-0.7616
TMEM150C	-0.65778	-0.455

C2orf54	-2.45978	-0.9029
INPP5J	-1.4691	-0.9262
NR3C1	-0.47164	0.56554
SCNN1A	-1.36197	-0.6256
SASH1	-0.79742	-0.49
CALML3-AS1	-0.91918	-1.0236
GHR	-1.40799	-1.5709
PLEKHF2	-0.49242	-0.3993
SPEG	-0.51744	-0.6702
ATRX	-0.25797	0.94589
PRR11	0.244951	0.52062
ME1	0.463535	0.4709
COL9A2	-1.27725	-0.7937
ERMP1	-0.57793	-0.3316
TRIM16	-0.5793	0.51942
ACHE	-3.43826	-0.8518
LFNG	-0.54835	-0.5927
HSH2D	-0.31324	-0.7153
IGF1R	0.217154	0.5578
ABCG1	-1.06329	-0.5158
DOPEY2	-0.37147	-0.396
RAB9A	-0.41178	0.41698
HR	0.724799	0.6216
CHST1	-1.33087	-1.8138
ZNF217	-1.11263	-0.3428
S1PR3	-0.82007	0.46718
SNHG3	0.494196	0.41826
FOXP4	-0.33415	-0.4206
PBK	0.515338	0.40147
TMEM184A	-0.4161	-0.2744
SMPDL3B	0.919566	0.88933
UCK2	0.57607	0.33102
EPOR	-0.96045	-0.626
NCS1	0.387845	0.28764
CDKN3	0.507873	0.38505
TP53INP2	-1.20973	-0.4489
CLSTN1	-0.31378	-0.2665
UGT1A6	-1.49938	0.72815
JUP	-0.2376	0.21067
PIK3IP1	-1.44754	-0.4035
GALNT10	-0.68066	0.30741

AHSA1	0.642493	0.36126
GNL3	0.423313	0.44538
CDCA3	0.573291	0.52049
MT2A	0.481078	-0.5952
DMRTA1	-1.66979	-0.609
APOD	-0.6323	-0.4663
DLL1	-0.51114	0.4403
PARP9	-0.48827	-0.5402
SECTM1	-1.37958	1.14013
NOP58	0.489763	0.55523
OSR2	-0.37233	-0.4471
RGS22	1.910792	1.46079
C1QBP	0.445925	0.32405
KCNC3	-0.89787	-0.7231
AGR3	1.785798	-0.5873
RTN4	0.289675	0.29716
RBL2	-0.53701	-0.4011
HILPDA	-0.64705	-0.382
BCL2L1	-0.47191	-0.3421
CCNB2	0.36822	0.48113
BAG3	-0.26922	-0.3768
CA8	1.706083	0.80736
AKR1C2	0.495678	0.47337
NLN	0.609003	0.41071
CLU	-0.60212	-0.5054
SLC39A6	0.319382	-0.4509
EPS8L1	-0.39634	-0.4386
PLK1	0.353433	0.52835
HMMR	0.269244	0.97119
GREB1	2.916463	0.49515
LRRC75A-AS1	0.288952	0.32079
MB	-1.19225	-0.4908
SNX33	-0.5394	-0.4439
CAPN13	-1.34403	-0.3603
SLC27A2	0.578858	0.58952
PTTG1	0.48712	0.41733
LTBP3	-0.89805	-0.4598
TUFT1	-0.52707	0.35093
BUB1	0.365493	0.58692
ADAT2	1.068878	0.76357
TINCR	0.356238	0.42035

HSPD1	0.590409	0.4198	SIM1	-1.78121	-1.4456	FAM174B	-0.70405	0.43097
RNF144B	-1.42335	-0.5905	TMEM205	-0.43402	-0.3981	IFRD1	0.47622	0.35657
TUBA1B	0.730043	0.46353	TBC1D2	-0.77788	-0.3326	PRRT2	-0.51352	-0.8444
SLC9A3R1	0.861735	-0.3796	SSR3	0.296486	0.2699	ANKRD50	-0.7715	-0.4137
TMEM64	0.956628	-0.364	SH3GLB2	-0.39772	-0.2581	SEMA4C	-0.52421	-0.4308
ATP5G1	0.664534	0.34983	TRAFD1	-0.57143	0.22855	P2RX2	-2.15942	-0.6054
CTNND2	-0.96865	-0.5184	IFFO2	-0.45036	0.39528	CTPS1	0.878573	0.33756
NCOA3	-0.56461	-0.3471	DARS2	0.603367	0.28356	AHRR	-1.00846	0.99727
NME1	0.861508	0.48023	HNRNPAB	0.517167	0.26342	YPEL3	-0.91846	-0.3846
RUNX1	0.424036	0.62229	TXNIP	-0.7109	-0.3057	PPT1	-0.33853	-0.3064
LPCAT1	0.387774	0.32008	PLCD3	0.312313	-0.2435	FAM83D	0.368989	0.51746
CENPA	0.432966	0.69643	NDRG4	-0.44602	0.35286	MMP16	-1.59719	-0.8704
DDX21	0.584913	0.43964	AOX1	-1.28534	-0.8208	BTG1	-0.55682	-0.3642
ABHD2	0.691698	-0.3498	CSTA	-2.58213	-1.7107	DLGAP5	0.360403	0.79869
SUSD2	-0.55383	-1.0217	TP53I11	-0.44489	-0.3558	CDH1	0.378598	-0.2792
IRAK1	0.369123	0.33223	RASD1	-0.78809	0.99337	TGFBI	-0.46841	0.49453
CDKN2B	-1.59539	-0.4207	UBE2S	0.568225	0.41852	ALOXE3	0.667374	1.13734
KCNJ3	-0.91568	-0.473	NEDD9	-0.67199	0.54357	DOCK8	-0.87052	-0.4047
HAR1B	-0.80863	1.24839	CCT5	0.549373	0.35278	C11orf24	0.598047	0.44118
VAT1L	-1.87241	-0.9528	CCNA2	0.438331	0.4945	TRAK1	-0.52093	-0.4085
SIPA1L2	0.279529	0.49162	ATP8B2	-0.96893	-0.5039			

Table 2A9: List of common genes regulated by karanjin in MCF-7 and T47D breast cancer cells. Red cells represent negative log_2FC values implying downregulation of the gene. Green cells represent positive log_2FC implying upregulation of the gene.

Gene	MCF-7	T47D	Gene	MCF-7	T47D
TP53INP2	-0.44889	-0.67426	**IGSF3**	-0.24576	-0.37477
SREBF1	-0.39126	-0.6126	**NDRG1**	-1.00342	-0.35113
ZG16B	-0.4268	-0.57974	**NBEAL2**	-0.32467	-0.34717
VPS9D1	-0.395	-0.55911	**FAM83D**	0.517457	-0.31225
PLPPR2	-0.40241	-0.55456	**ARHGAP32**	-0.3462	0.294584
PLCD3	-0.24346	-0.53431	**GAS5**	0.323995	0.366372
UPK3B	-0.396	-0.53248	**PPP1R3B**	-0.48495	0.446681
LIMK1	0.361057	-0.47941	**CD44**	0.882848	0.530303
GPRC5A	-0.29944	-0.45983	**AGR3**	-0.5873	0.532298
TM7SF2	-0.36332	-0.43756	**NEDD9**	0.543568	0.538306
PFKFB3	-0.52631	-0.43635	**TIPARP**	2.058824	0.679224
HILPDA	-0.38202	-0.42582	**GPNMB**	-0.89328	0.696212
SPINT1	-0.29754	-0.41182	**GNL3**	0.445381	0.716173
RAB3D	-0.69954	-0.39693	**CYP1B1**	2.723354	1.034115
FLNA	0.234818	-0.39066	**CENPE**	1.969793	1.563819
TLE3	-0.32482	-0.39011	**CYP1A1**	6.17733	2.607662
NCS1	0.287636	-0.38456			

Table 2A10: List of common genes regulated by karanjin and tamoxifen. Red cells represent negative log_2FC values implying downregulation of the gene. Green cells represent positive log_2FC implying upregulation of the gene.

No.	Gene	E2	Karanjin	No.	Gene	E2	Karanjin
1	L1CAM	-0.73895249	-0.85323	62	PXN	-0.336557037	-0.39292
2	ANGPTL4	-0.814197874	-0.78695	63	PTPRU	-0.483986961	-0.39271
3	SLC52A2	0.361570215	-0.75202	64	SRF	0.301846414	-0.39253
4	EGLN3	-1.09325703	-0.71921	65	TLE3	0.361049597	-0.39011
5	SBK1	-0.447292811	-0.70253	66	DAG1	-0.234865372	-0.38909
6	OPLAH	-0.612526759	-0.68099	67	EXT1	-0.474362422	-0.38569
7	MPG	-0.28973729	-0.67838	68	CPAMD8	-1.016875423	-0.38528
8	TP53INP2	-1.209733548	-0.67426	69	NCS1	0.387845031	-0.38456
9	TNS4	1.02045621	-0.65163	70	SLC44A2	-0.392743349	-0.37541
10	CAPN5	-0.775205746	-0.63697	71	KRT7	-0.447215677	-0.36913
11	E2F2	0.598674235	-0.63097	72	HNRNPL	0.269278942	-0.36115
12	SRM	0.669935994	-0.62875	73	INSIG1	0.355029527	-0.35693
13	UNC13D	-0.860950918	-0.6206	74	MOV10	-0.250410269	-0.35518
14	E2F1	0.712539033	-0.61298	75	NOTCH1	-0.270565343	-0.35343
15	SREBF1	-0.22730371	-0.6126	76	NDRG1	-0.593453104	-0.35113
16	CCND3	0.323715096	-0.60974	77	RRM2	0.644190794	-0.35107
17	TCF19	0.43134917	-0.60041	78	TNS3	-1.494395401	-0.34752
18	LIG1	0.451744249	-0.59346	79	NBEAL2	-0.443765854	-0.34717
19	CDT1	0.403374456	-0.58965	80	VEGFA	0.296526169	-0.33051
20	SIPA1	-0.621262803	-0.58705	81	SCD	0.367440531	-0.32909
21	MVD	0.428040208	-0.5853	82	POLE	0.316149367	-0.32856
22	ZG16B	-0.695315316	-0.57974	83	ATP1A1	0.49090023	-0.3277
23	TK1	0.603763027	-0.57849	84	MCM3	0.591206578	-0.32482
24	JUNB	-0.708115161	-0.57093	85	FAM83D	0.368989387	-0.31225
25	PLXNA1	-0.313674067	-0.54569	86	ACLY	0.241935476	-0.2934
26	PLCD3	0.312312957	-0.53431	87	UBA1	-0.187354249	-0.2899
27	SCARB1	0.599714521	-0.53206	88	CD24	-1.139904809	-0.28306
28	NACC2	-0.495331887	-0.52811	89	ARHGAP32	-0.414772806	0.294584
29	ASF1B	0.455452868	-0.52429	90	CDV3	1.008603047	0.361278
30	ELF3	-0.448059907	-0.52017	91	SLC2A13	-0.643434171	0.5199
31	SLC12A4	-0.425066914	-0.5063	92	CD44	1.116590815	0.530303
32	CNP	-0.349343972	-0.50562	93	AGR3	1.78579776	0.532298
33	VWA1	-0.455592841	-0.50034	94	NEDD9	-0.6719876	0.538306
34	MCM2	0.598279834	-0.49491	95	CXCL12	2.606555532	0.560171
35	ESPL1	0.23350743	-0.49484	96	ARHGAP11A	0.483848773	0.60344
36	TMPRSS2	-1.44763785	-0.48982	97	FMN1	1.656216062	0.645363
37	ERBB2	-0.930176559	-0.48774	98	GAB1	-0.581580002	0.668079
38	NLGN2	-0.334362347	-0.47838	99	DSCAM-AS1	0.520179965	0.670973
39	IL4R	0.67492132	-0.47613	100	TIPARP	1.127495804	0.679224
40	SERINC2	-0.483375051	-0.47289	101	RRP15	0.485213718	0.680157
41	UHRF1	0.597298437	-0.47053	102	VPS13D	-0.360089556	0.688875
42	GLB1L2	0.611232858	-0.46833	103	SGK3	2.061613452	0.699716
43	PGP	0.423779095	-0.46683	104	GNL3	0.42331334	0.716173

44	SNRNP25	0.421999034	-0.46295	105	HSPA4	0.397196592	0.753814
45	MISP	-0.535993597	-0.45955	106	TMEM106B	-0.380150019	0.765503
46	SOX13	-0.340041645	-0.45212	107	PAK1IP1	0.634367714	0.78387
47	SEMA3F	-0.417162533	-0.449	108	SULF1	0.544637363	0.793788
48	PLXNB2	-0.581774464	-0.44762	109	CADM1	0.524040287	0.798591
49	TM7SF2	-0.472996433	-0.43756	110	PGR	2.901930007	0.815701
50	PFKFB3	-0.311065644	-0.43635	111	ITGB3BP	0.648397288	0.877211
51	TNFRSF21	-0.280672277	-0.43541	112	SYF2	-0.399655468	0.879748
52	MYBL2	0.867406946	-0.4354	113	MYB	1.390885859	0.931898
53	HILPDA	-0.647050634	-0.42582	114	SNX10	0.417576299	0.980616
54	NT5DC2	0.239927254	-0.42477	115	PDZK1	3.311357936	1.014441
55	WFS1	0.724708923	-0.42066	116	NUF2	0.43497266	1.02368
56	SLC9A1	-0.473048689	-0.41705	117	CTNNAL1	0.36838445	1.073807
57	SPINT1	-0.411112311	-0.41182	118	ZNF33B	-0.319617109	1.097307
58	BCAM	-0.911522342	-0.40685	119	LRRCC1	0.423696801	1.127445
59	BCL9L	-0.550505455	-0.40597	120	CALCR	1.28943057	1.294358
60	CAPN1	-0.267229177	-0.40222	121	CENPE	0.288627013	1.563819
61	RAB3D	-1.00866275	-0.39693	122	CYP1A1	-1.751582358	2.607662

Table 2A11: List of common genes regulated by karanjin and tamoxifen. Red cells represent negative log_2FC values implying downregulation of the gene. Green cells represent positive log_2FC implying upregulation of the gene.

No.	Gene	Karanjin	Tamoxifen	No.	Gene	Karanjin	Tamoxifen
1	**CXCL12**	0.560171	-0.68028	12	**CD24**	-0.28306	-0.15871
2	**SULF1**	0.793788	-0.58247	13	**LSS**	-0.47694	0.148716
3	**PLCD3**	-0.53431	-0.51035	14	**CDV3**	0.361278	0.158632
4	**L1CAM**	-0.85323	-0.26709	15	**DHCR24**	-0.50293	0.184866
5	**MCM2**	-0.49491	-0.25855	16	**MVD**	-0.5853	0.200237
6	**CYP1B1**	1.034115	-0.23035	17	**DSCAM-AS1**	0.670973	0.26282
7	**LIG1**	-0.59346	-0.22797	18	**INSIG1**	-0.35693	0.273184
8	**GPRC5A**	-0.45983	-0.22756	19	**ELF3**	-0.52017	0.28006
9	**CDT1**	-0.58965	-0.20387	20	**SCD**	-0.32909	0.290853
10	**RRM2**	-0.35107	-0.19696	21	**SCARB1**	-0.53206	0.301801
11	**TK1**	-0.57849	-0.18141	22	**AGR3**	0.532298	1.42952

Appendix III

Table 3A1: List of primers used in qRT-PCR

Primer name	Sequence	Amplicon length (bp)	Annealing temp. (°C)	Primer name	Sequence	Amplicon length (bp)	Annealing temp. (°C)
CSTA-F	ATCTGAGGCCAAA CCCGCC	275	60	TIPARP-F	TCATTGGCAGATCA AAAGGACAAC	160	60
CSTA-R	AGCCCGTCAGCTCG TCATC			TIPARP-R	CACGTTCATGGCAT TCAAATCTGC		
CycA-F	GGGCCGCGTCTCCT TTGAGC	158	60	STC2-F	ATGCTACCTCAAGC ACGACC	129	60
CycA-R	GGCGTGTGAAGTC ACCACCC			STC2-R	CAGGTCAGCAGCA AGTTCAC		
BRINP2-F	GACTGGCTGCTCAC AGACC	126	60	CD44-F	CAAGTTTTGGTGGC ACGCAG	135	60
BRINP2-R	CACCTTCCAACGGG CAAAC			CD44-R	GTCCGAGAGATGC TGTAGCG		
B3GALT5-F	CCTCTTGGCATTTA CACTGTGG	115	60	CYP1A1-F	ACCTTTGAGAAGG GCCACATCCG	154	60
B3GALT5-R	CCCCCAGAACCAG AAGGC			CYP1A1-R	TGACTGTGTCAAAC CCAGCTCCAAAG		
TFF1-F	GGGTCCCCTGGTGC TTCTAT	140	60	CDC25B- F	CTGTAGCCTGGACA AGAGAGTC	112	60
TFF1-R	AGCCGAGCTCTGG GACTAA			CDC25B- R	GTAGTCGTTGACAG CACGGT		
CHST1-F	GCCCTTTCGACCTG GAGG	134	60	CENPF-F	GGCTGCACAGAAG TTAGCG	143	60
CHST1-R	CAAGGGGTGAGGT CAAAGAGG			CENPF-R	GGAGGATGGTGCC TGAATCTAC		
SLC7A5-F	GTGGACTTCGGGA ACTATCACC	126	60	MYC-F	CCGTCCTCGGATTC TCTGCT	231	60
SLC7A5-R	GGACCCCACGAAG AAGAGC			MYC-R	TGGGCTGTGAGGA GGTTTGC		
CYP1B1-F	GCCACTATCACTGA CATCTTCG	129	60	CREG2-F	CAGATGATCGCAG TGTCTCCA	153	60
CYP1B1-R	CACGACCTGATCCA ATTCTGC			CREG2-R	GCCTCCATACCATT TCTGAAGC		
AHR-F	TAGCCTGCTGCCTT TCCCACAAG	168	60				
AHR-R	TGCTGCTCTACAGT TATCCTGGCCT						

Note: F and R indicate sense and antisense primers, respectively.

Table 3A2: List of primers used in Chromatin immunoprecipitation assay

Primer name	Sequence	Amplicon length (bp)	Annealing temperature (°C)
PS2-F	CATTGCCTCCTCTCTGCTCC	423	58
PS2-R	ACTGTTGTCACGGCCAAGCC		
CSTA F	TCCTACTGGATCTCAGCCAC	369	60
CSTA-R	GCCCTGTTCTTAGAATAGTGC		
HOXB2-F	GACTGGCTGCTCACAGACC	279	60

HOX2-R	CACCTTCCAACGGGCAAAC		
CYP1A1-F	AGTCCCAATTCCAAGGCGTC	406	60
CYP1A1-R	CCTTCGCCATCCATTCCGAT		
TIPARP-F	GCCAAAACCTCACAGCTTCC	427	60
TIPARP-R	TGTGCGGTGGACTTATGCTC		

Note: F and R indicate sense and antisense primers, respectively.

Appendix IV

AhR Panel

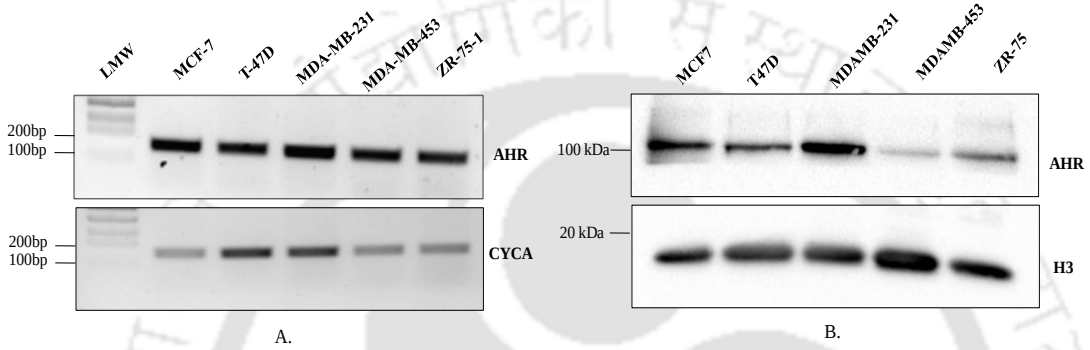


Fig. 4A3: Gene expression profiling in different breast cancer cell lines via semi–quantitative PCR. Agarose gel image of AhR gene and CYA (normalizing gene). B. Protein expression profiles in different breast cancer cell lines. Histone (H3) served as a normalizing gene.

Appendix V

Cloning of ESR1 LBD ORF with His-tag in PET28a (+) vector:

To study the significance of ERα LBD and its biophysical interaction analyses with Karanjin. We cloned ESR1 LBD into a bacterial expression vector PET28a (+) (Fig. 5A1). To amplify the ORF, we have used cDNA of T47D cell line. cDNA isolated from T47D cells was used as template. 0.4 U of Phusion High fidelity DNA polymerase was used for PCR, in a 20 µL reaction. With the following conditions: Initial denaturation of 1 min at 98 C followed by 30 cycles of denaturation at 98 C for 10 sec, annealing at 64 C for 30 sec and extension at 72 C for 30 sec for 5 min of final extension at 72 C. PCR amplified ORF was cloned into PET28a (+) by using Gibson method. Clones were checked by restriction digestion with Nde I and Xho I Moreover, PCR confirmed with Sequencing Primers T7 Promoter Fw and T7 terminator Rw. Finally, ESR1 LBD construct was sequence confirmed (Fig. 5A2 & Fig. 5A3).

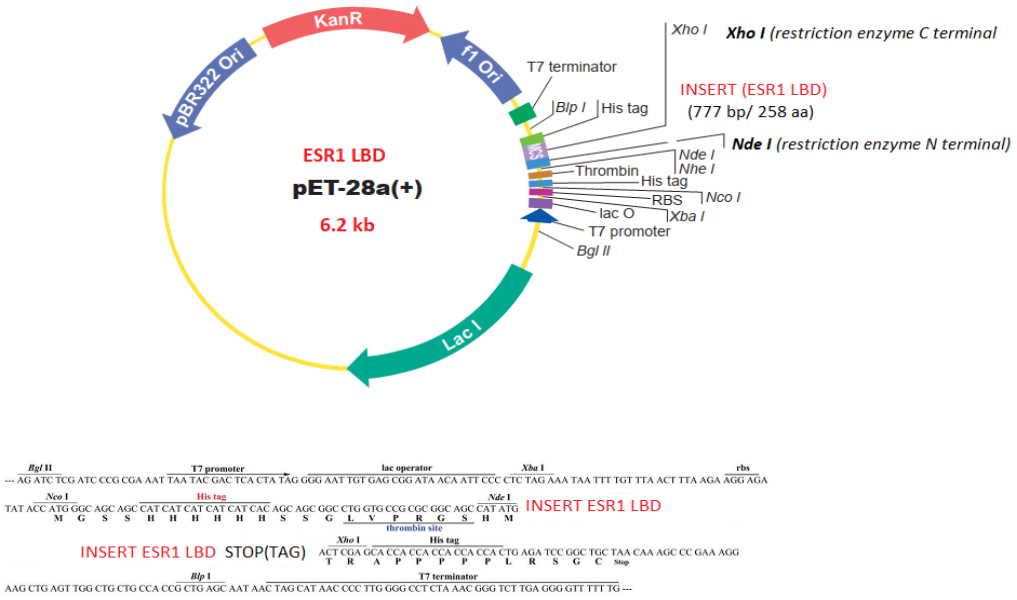


Fig. 5A1. Vector-insert map ESR1 LBD PET28a (+) for expression of ERα LBD.

Clone 2 B (translated sequence)

MIKRSKKNSLALSLTADQMVSALLDAEPPILYSEYDPTRPFSEASMMGLLTNLADRELVHMINWAKRVPG
FVDLTLHDQVHLLECAWLEILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFR
MMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHQRLAQ
LLLILSHIRHMSNKGMEHLYSMKCKNVVPLYDLLLEMLDAHRLHAPTS

3ERT (PDB SEQUENCE OF ESR1 LDB)

MIKRSKKNSLALSLTADQMVSALLDAEPPILYSEYDPTRPFSEASMMGLLTNLADRELVHMINWAKRVPG
FVDLTLHDQVHLLECAWLEILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFR
MMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQHQRLAQ
LLLILSHIRHMSNKGMEHLYSMKCKNVVPLYDLLLEMLDAHRLHAPTS

```
100.0% identity in 258 residues overlap; Score: 1322.0; Gap frequency: 0.0%

UserSeq1       22 MIKRSKKNSLALSLTADQMVSALLDAEPPILYSEYDPTRPFSEASMMGLLTNLADRELVH
UserSeq2        4 MIKRSKKNSLALSLTADQMVSALLDAEPPILYSEYDPTRPFSEASMMGLLTNLADRELVH
                  ************************************************************

UserSeq1       82 MINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLVWRSMEHPGKLLFAPNLLLDRNQGK
UserSeq2       64 MINWAKRVPGFVDLTLHDQVHLLECAWLEILMIGLVWRSMEHPGKLLFAPNLLLDRNQGK
                  ************************************************************

UserSeq1      142 CVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIH
UserSeq2      124 CVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSIILLNSGVYTFLSSTLKSLEEKDHIH
                  ************************************************************

UserSeq1      202 RVLDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPL
UserSeq2      184 RVLDKITDTLIHLMAKAGLTLQQQHQRLAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPL
                  ************************************************************

UserSeq1      262 YDLLLEMLDAHRLHAPTS
UserSeq2      244 YDLLLEMLDAHRLHAPTS
                  ******************
```

Fig. 5A2. Sequence confirmed ESR1 LBD clone.

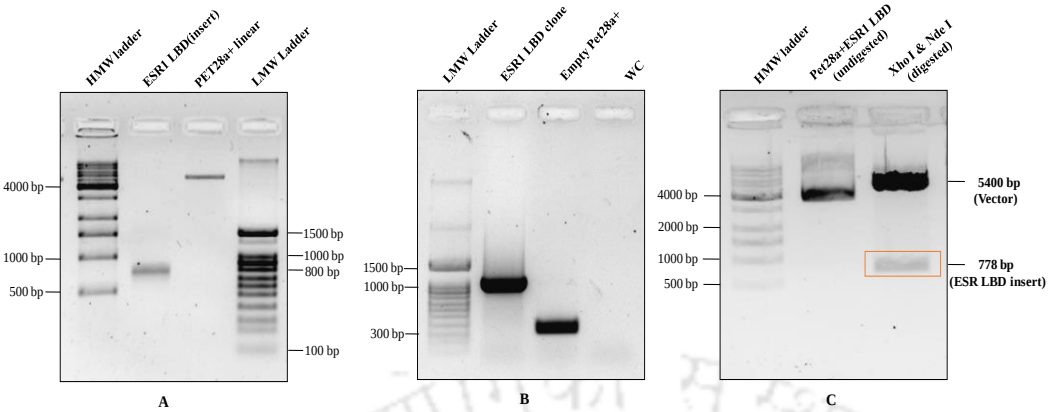


Fig. 5A3: Cloning and clone confirmation through PCR and restriction digestion A. Insert and linearized vector. B.PCR amplified product of ESR1 clone with T7 Fw and T7 terminator primers. C. Restriction digestion of ESR1 clone with Nde I and Xho I enzymes.

ESR1 LBD clone was further expressed in the BL21(DE3) strain of E. coli. Protein was induced with IPTG (0.2mM) for a time period of 4 h at 23 °C in a shaking incubator at 180 RPM. Protein was expressed in soluble fraction and confirmed by western blotting by screening with Anti-His antibody. ESR1 LBD protein was observed around 30kDa (Fig. 5A4).

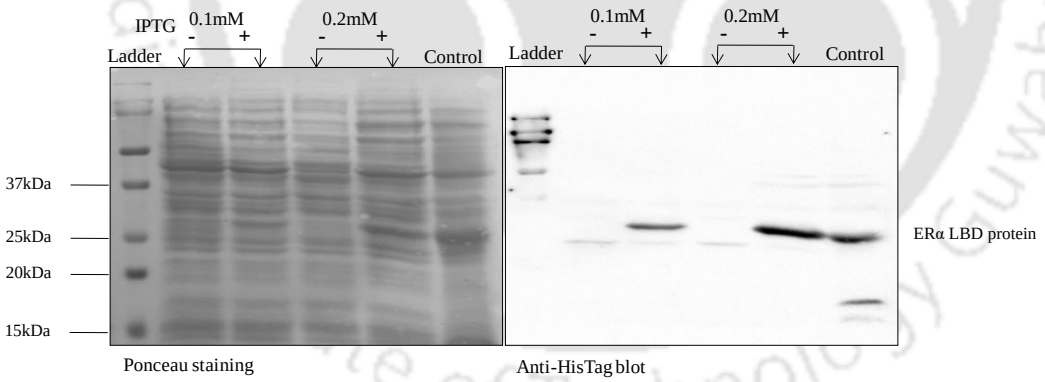


Fig. 5A4: Western blotting with Anti-His antibody to confirm the 6X-His-tag bound ESR1 LBD in the soluble fraction of clone.

Purification of ESR1 LBD protein:

Soluble expressed lysate fraction was purified using NI-NTA beads in a gravity column using the following protocol:

- **Column Preparation**-1.2 ml of bead was added to the column and washed with MQ water thrice at an interval of 15 mins

- **Equilibrate**-Equilibrated with Lysis buffer (0.5M NDSB, 100mM NaCl, 50 mM Tris-Cl, pH 7.4, 10 mM ßME) 5ml each twice for 15 mins.
- **Add Protein**-8ml of supernatant was added to the column and kept at 4°c in a rocker for three and half hours. Flow through was collected
- **Wash**-Once washed with lysis buffer, given 6ml Wash buffer 1(10mM imidazole) and 2 (20mM imidazole, 0.5M NDSB, 100mM NaCl, 50 mM Tris-Cl. pH 7.4, 3 mM ßME) wash given twice 6 ml each at an interval of 15 mins kept in shaking.
- **Elution**-Not kept for shaking, gently mixed the beads with pipette.1.5ml each of elutes were collected. Elutes were collected with Imidazole 250 mM containing elution buffer (0.5M NDSB, 100mM NaCl, 50 mM Tris-Cl. pH 7.4, 3 mM ßME) and eluted protein were further dialyzed and stored in -80 °C for validation and use (Fig. 5A5).

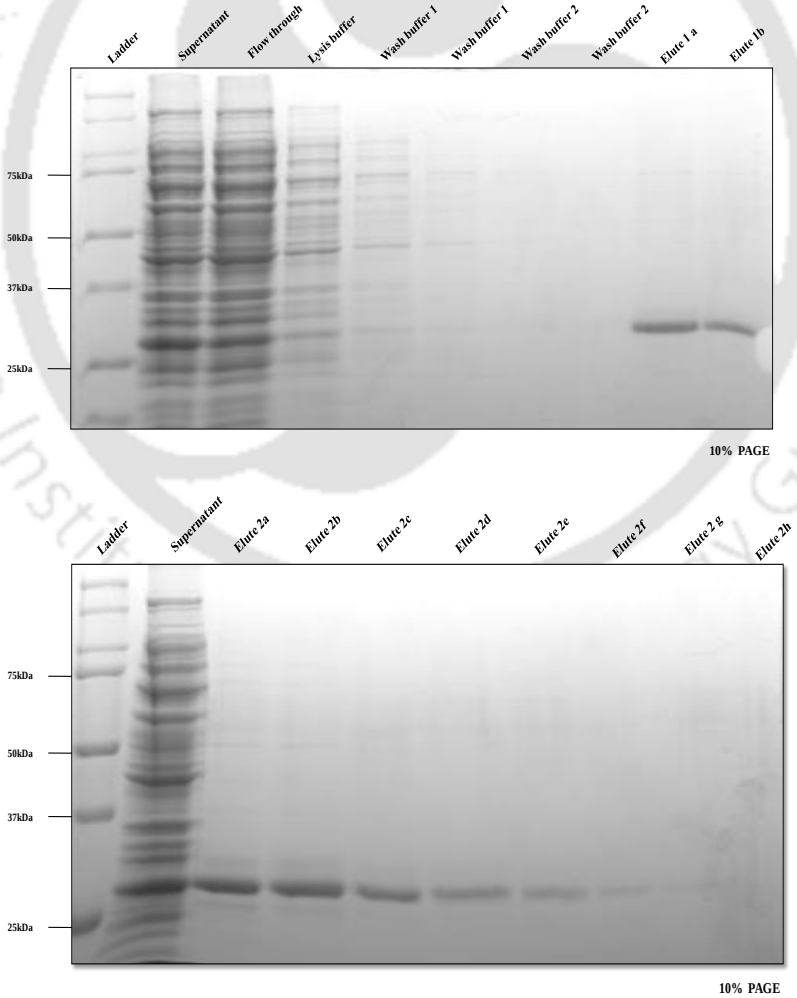
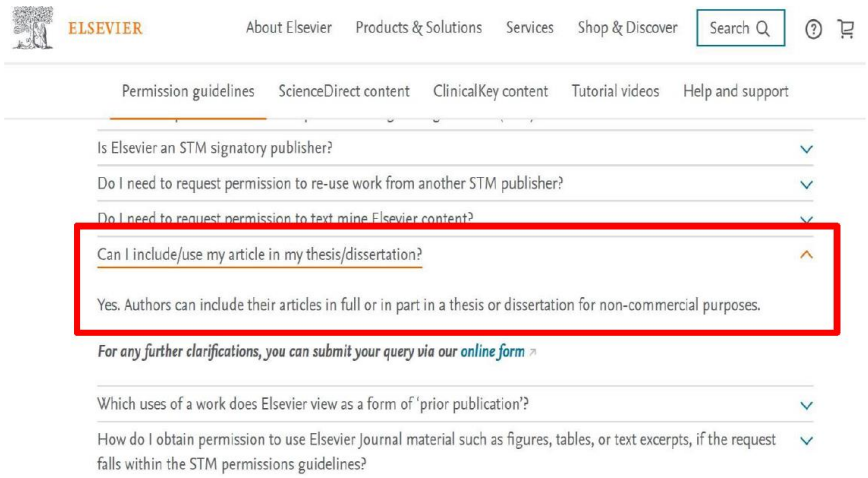


Fig. 5A5: Coomassie blue-stained 10% SDS–PAGE of 6x-His-ERα LBD protein.

Appendix VI

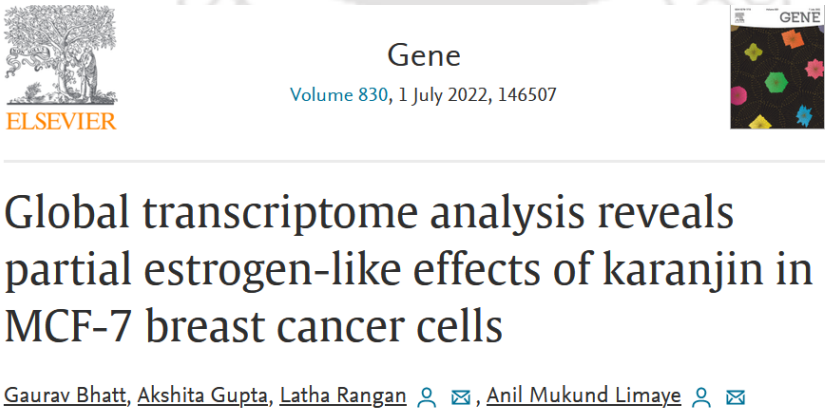
Copyrights and permissions

Chapters 2 & 4: Author of the publication and no permission required



Chapter 2: Review of literature

Chapter 4: Genome-wide transcriptomic effects of karanjin in MCF-7 breast cancer cells



List of Figures

List of Tables

List of abbreviations

A549	Human lung adenocarcinoma cell line
ABCG2	ATP binding cassette subfamily G member 2
ADAMTS19	ADAM Metallopeptidase with Thrombospondin Type 1 Motif 19
AF	Activation function
AhR	Aryl hydrocarbon receptor
AHRR	AhR repressor
AIs	Aromatase inhibitors
Akt	Protein Kinase B
AMPK	AMP-activated protein kinase
ANOVA	Analysis of variance
AP	Activator protein
AREG	Amphiregulin
ARNT	AhR nuclear translocator
ATP	Adenosine triphosphate
ATPase	Adenosine triphosphatase
B3GALT5	Beta-1,3-galactosyltransferase 5
Bax	BCL2-Associated X Protein
Bcl-2	B-cell lymphoma 2
BHLH	Basic helix-loop-helix
BMF	Bcl2 Modifying Factor
BRINP2	BMP/retinoic acid-inducible neural-specific protein 2
BSA	Bovine serum albumin
cAMP	Cyclic adenosine monophosphate
CD44	CD44 molecule (Indian blood group)
CDC2	Cyclin dependent kinase 2
CDC25B	Cell Division Cycle 25B
CDKs	Cyclin-dependent kinases
cDNA	Complementary Deoxyribonucleic acid
CENPF	Centromere protein F
CHARMM	Chemistry at Harvard Macromolecular Mechanics
ChIP	Chromatin immunoprecipitation
ChIP-Seq	Chromatin immunoprecipitation sequencing
CHST1	Carbohydrate sulfotransferase 1
CREG2	Cellular repressor of E1A-stimulated genes 1

csFBS	Charcoal stripped FBS
CSTA	Cystatin A
CycA	Cyclophilin A
CYP	Cytochrome P450
CYP1A1	Cytochrome P450 family 1 subfamily A member 1
CYP1B1	Cytochrome P450 family 1 subfamily B member 1
DBD	DNA binding domain
DCIS	Ductal carcinoma in-situ
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
DPBS	Dulbecco's phosphate-buffered saline
E2	17β-Estradiol
ECL	Enhanced chemiluminescence
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
ER	Estrogen receptor
ERE	Estrogen response elements
EROD	Ethoxyresorufin-O-deethylase
ERα	Estrogen receptor alpha
ERβ	Estrogen receptor beta
EtBr	Ethidium bromide
EtOH	Ethanol
FastQC	Fast Quality check
FBS	Fetal bovine serum
FDR	False Discovery Rate
fGSEA	Fast preranked gene set enrichment analysis
FICZ	6-formylindolo[3,2-b] carbazole
FOSL2	FOS Like 2, AP-1 Transcription Factor Subunit
G2/M	G2/entry into mitosis
GEO	Gene Expression Omnibus
GLUT4	Glucose transporter type 4
GPER1	G-protein coupled estrogen receptor 1
GRCh38	Genome Reference Consortium Human Build 38 Organism
H3	Histone
HCl	Hydrochloride
HEPES	4-(-2hydroxyethyl)-piperazineethanesulfonic acid

HER2	Human epidermal growth factor receptor 2
HIF-2α	Hypoxia-inducible factor 2α
HOXB2	Homeobox B2
HRMS	High-resolution mass spectrometry
HRP	Horseradish peroxidase
Hsp	Heat shock protein
IARC	International Agency for Research on Cancer
IC50	Half maximal inhibitory concentration
IDC	Invasive ductal carcinoma
IGFBP1	Insulin-like growth factor-binding protein 1
IgG	Immunoglobulin G
IL8	Interleukin 8
IP	Immunoprecipitation
IRS-1	Insulin Receptor Substrate-1
IUPAC	International Union of Pure and Applied Chemistry
I-κB	Inhibitor of nuclear factor kappa B
kb	Kilobase
kDa	Kilo Dalton
KRAS	Kirsten rat sarcoma virus
LBD	Ligand binding domain
LBP	Ligand binding pocket
LFC	Log fold change
MAPK	Mitogen-activated protein kinases
MCF-7	Michigan Cancer Foundation – 7
MD	Molecular dynamics
MDA-MB-231	M.D. Anderson - Metastatic Breast 231.
mL	Milliliter
mRNA	Messenger ribonucleic acid
MTT	Methylthiazolyldiphenyl-tetrazolium bromide
MYC	MYC proto-oncogene, bHLH transcription factor
NaCl	Sodium chloride
NAMD	Nanoscale Molecular Dynamics
NCBI	National Center for Biotechnology Information
NCCS	National Centre for Cell Science
NCOA3	Nuclear Receptor Coactivator 3
nER	Nuclear estrogen receptors

NES	Normalized enrichment scores
NF-κB	Nuclear factor-κB
ng	Nanogram
nM	Nanomolar
NMR	Nuclear magnetic resonance
NOS	Nitric oxide synthase
NQO1	NAD(P)H quinone dehydrogenase 1
NTD	N terminal domain
PAGE	Polyacrylamide gel electrophoresis
PAHs	Polycyclic aromatic hydrocarbons
PBS	Phosphate-buffered saline
PBST	PBS containing 0.05% Tween 20
PCA	Principal component analysis
PCBs	Polychlorinated biphenyls
PCDFs	Polychlorinated dibenzofurans
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PHLDA1	Pleckstrin homology-like domain family A member 1
PI3K	Phosphoinositide 3-kinases
PKA	Protein kinase A
PR	Progesterone receptor
PS2/TFF-1	Trefoil factor 1
PTP	Protein-tyrosine phosphatase
QC	Quality check
qRT-PCR	Quantitative reverse transcription PCR
RIN	RNA integrity number
RMSD	Root Mean Square Deviation
RNA	Ribonucleic acid
RNAiMAX	RNAi-specific cationic lipid formulation
RNase	Ribonuclease
RNA-seq	RNA Sequencing
ROS	Reactive oxygen species
RP-HPLC	Reverse-phase high-performance liquid chromatography
RPKM	Reads per kilobases million
RPMI-1640	Roswell Park Memorial Institute-1640 medium

RT-PCR	Reverse transcription-polymerase chain reaction
SAhRMs	Selective AhR modulators
Scr	Scrambled
SD	Standard deviation
SDF	Standard delay format
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
SERDs	Selective estrogen receptors down-regulators
SERMs	Selective estrogen receptor modulators
siRNA	Short interfering ribonucleic acid
SLC7A5	Solute carrier family 7-member 5
SMILES	Simplified molecular-input line-entry system
SRA	Sequence Read Archival
SRL	Sisco research laboratory
STC2	Stanniocalcin 2
TAM	4-hydroxytamoxifen
TBST	Tris-buffered saline containing 0.05% Tween 20
TCA	Trichloroacetic acid
TCDD	2,3,7,8 -Tetrachlorodibenzo-p-dioxin
TE	Tris EDTA
TIP3P	Transferable intermolecular potential 3P
TIPARP	TCDD-inducible poly [ADP-ribose] polymerase
TLC	Thin-layer chromatography
Tm	Melting temperature
TNBC	Triple-Negative Breast Cancer
TNFα	Tumor necrosis factor α
TukeyHSD	Tukey Honest Significant Differences
UV	Ultraviolet
WHO	World Health Organization
Wnt	Wingless-related integration site
XRE	Xenobiotic response elements
βNF	Beta-naphthoflavone
µl	Microliter
µm	Micrometer
µM	Micromolar
4OHT	4-Hydroxytamoxifen

Research outputs

List of publications

1. **G Bhatt**, A Gupta, L Rangan, A M Limaye, Global transcriptome analysis reveals partial estrogen-like effects of karanjin in MCF-7 breast cancer cells, **Gene**, **2022,** 830,146507
2. **G Bhatt**, A Singh, N Gujre, S Mitra, R Swaminathan, AM Limaye, Karanjin, **Phytochemistry, 2021,** 183, 112641

Publication from other collaborative research work

1. NV Puranik, P Srivastava, **G Bhatt**, DJS John Mary, AM Limaye, Determination and analysis of agonist and antagonist potential of naturally occurring flavonoids for estrogen receptor (ERα) by various parameters and molecular modelling approach, **Scientific reports, 2019 ,**9 (1), 1-11
2. ND Devi, C Mukherjee, **G Bhatt**, L Rangan, VV Goud, Co-cultivation of microalgae-cyanobacterium under various nitrogen and phosphorus regimes to concurrently improve biomass, lipid accumulation and easy harvesting, **Biochemical Engineering Journal, 2022,** 188, 108706

NEWGEN-IEDC project grant:

Bhatt G, Gupta MK, Nuzelu, Singh R, Rangan L, (2023), Development of safe and cost-effective plant derived nucleic acid staining kit.

List of Poster and Oral presentation

- **Gaurav Bhatt**, Latha Rangan, Anil Mukund Limaye, "Karanjin: An aryl hydrocarbon receptor modulator in breast cancer cells". International Conference on Biotechnology, Sustainable Bioresources and Bioeconomy, BSBB 2022, pp 321. Indian Institute of technology Guwahati, Assam, India December 7-11, 2022.

- **Gaurav Bhatt**, Latha Rangan, Anil Mukund Limaye, "Karanjin- a novel selective estrogen receptor modulator ". International Seminar on Medicinal Plant Production and advancement in herbal medicine, nutraceuticals, cosmetics, and other herbal product formulations technologies, IIT Guwahati September 28-29th October, 2022. (**ORAL**)

- **Gaurav Bhatt**, Akshita Gupta, Latha Rangan, Anil Mukund Limaye. "Karanjin exerts partial estrogen-like effects in breast cancer cells." International Conference on Recent Trends and Future Opportunities in Pharmaceuticals. PHARMACON 2022, pp 110.

National Institute of Pharmaceutical Education and Research (NIPER) S.A.S. Nagar (Mohali), Punjab, India November 10-12, 2022.

- **Gaurav Bhatt**, Anil Mukund Limaye, Latha Rangan, "Karanjin mediated regulation of estrogen target genes in MCF-7 breast cancer cells". International conference on Biotechnology for Sustainable Agriculture, Environment and Health, Jaipur, Rajasthan, India, April 04-08, 2021.

- **Gaurav Bhatt**, Abhishek Kumar, Priyadarshi Satpati, Anil Mukund Limaye, Latha Rangan "Insight into Estrogen receptor α (ERα) and ligands (estrogen, karanjin) interactions using computer simulations", International conference on Nutraceuticals and chronic diseases (INCD), IIT Guwahati September 21-23, 2019. **(BEST POSTER).**

List of Workshops attended

- **Computer aided drug design for human pathogens**
  The International seminar cum workshop on "Computer aided drug design for human pathogens," Tezpur university, 12-17th Feb,2018.
- **Confocal Laser Scanning Microscopy**
  Workshop on Confocal Laser Scanning Microscopy and its applications, Biotech research park 2-3th Nov,2017.
- **GCMS & HPLC**
  Workshop on Analytical instruments and their applications (GCMS & HPLC), 9-10th Nov,2017
- **Flow Cytometry**
  20th Indo US Flow Cytometry Symposium cum Workshop on "Applications of flowcytometry in biotechnology" from 13th to 16th March held in IIT, Guwahati, India, 13-16th March, 2019.
- **Flow Cytometry**
  9th Annual TCS Event & Flow Cytometry Workshop on Flow Application in Basic Applied & Clinical Biology, November 03 – 05, IIT Guwahati & Dr. B Barooah Cancer Institute, Guwahati, 3-5th Nov, 2016
- Indo-Japan Workshop on Translational Agriculture Avenues for International Cooperation, 2017.