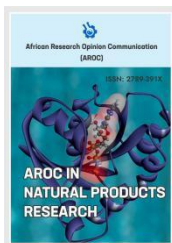


REVIEW ARTICLE

# Artificial Intelligence-Guided Discovery of Natural Products Targeting the Cancer–Diabetes Metabolic Axis: Multi-Omics, Network Pharmacology, and Precision Phytotherapy

Mufutau Adewale Lawal<sup>1</sup>, Dennis Yisheng Daniel<sup>2</sup>, Aliyah Omotayo Abdulkabir<sup>2</sup>, Deborah Arewa Popoola<sup>3</sup>, and Bolaji Mubarak Ayeyemi<sup>4</sup>



<sup>1</sup>Usman Danfodio University Teaching Hospital Sokoto, Nigeria.

<sup>2</sup>Ahmadu Bello University Teaching Hospital, Zaria, Kaduna State, Nigeria

<sup>3</sup>Faculty of Clinical sciences, College of Medicine, Kaduna State University, Kaduna State.

<sup>4</sup>Department of Computational Data Science and Engineering, North Carolina Agricultural and Technical State University, Greensboro, North Carolina, USA.

Corresponding Author: Mufutau A. Lawal  
Email: [25210704003@udusok.edu.ng](mailto:25210704003@udusok.edu.ng)

## ABSTRACT

**Background:** Cancer and type 2 diabetes mellitus are increasingly recognized as metabolically intertwined diseases sharing fundamental biological drivers, including insulin resistance, chronic low-grade inflammation, oxidative stress, dysregulated lipid and glucose metabolism, immune dysfunction, and aberrant growth-factor signaling. Epidemiological data consistently demonstrate that individuals with type 2 diabetes face a substantially elevated risk for several malignancies, including hepatocellular, pancreatic, colorectal, breast, and endometrial cancers, with excess risk attributed largely to hyperinsulinemia, hyperglycemia-driven metabolic reprogramming, and pro-inflammatory cytokine signaling. Natural products, encompassing polyphenols, alkaloids, terpenoids, flavonoids, saponins, and phenolic acids, offer intrinsic multi-target pharmacological activity against pathways operative in both diseases; however, their clinical translation has been severely limited by poor mechanistic characterization, complex chemical composition, variable bioavailability, insufficient patient stratification, and a lack of rigorous multi-disease validation. Advances in artificial intelligence (AI), multi-omics integration, molecular docking, network pharmacology, and machine learning now provide a transformative computational infrastructure to identify bioactive compounds, predict molecular targets, map disease-specific biological networks, and prioritize natural products for cancer–diabetes comorbidity. This review systematically synthesizes the biological foundations of the cancer–diabetes metabolic axis, classifies natural products by mechanism of dual activity, and critically evaluates current AI-enabled discovery frameworks. We further propose a comprehensive precision phytotherapy paradigm integrating patient metabolic phenotyping, tumor omics profiling, AI-driven compound–target matching, and pharmacokinetic optimization to accelerate mechanism-based, patient-stratified natural-product intervention for cancer–diabetes comorbidity.

**Keywords:** Artificial intelligence; Natural products; Cancer; Diabetes; Metabolic axis; Multi-omics; Network Pharmacology; Precision phytotherapy; Molecular docking; Insulin resistance

Received: 25 April 2026, Revised: 10 May 2026, Published: May 2026

**Citation:** M.A., Lawal, D.Y., Daniel, A.O., Abdulkabir, D.A., Popoola, & B.M., Ayeyemi (2026). *Artificial Intelligence-Guided Discovery of Natural Products Targeting the Cancer-Diabetes Metabolic Axis: Multi-Omics, Network Pharmacology, and Precision Phytotherapy*. *AROC In Natural Products Research* 6(1):01-23, <https://doi.org/10.53858/arocnpr06010123>

## 1. Introduction

Cancer and type 2 diabetes mellitus (T2DM) represent two of the most pressing public health challenges of the twenty-first century, together imposing enormous morbidity, mortality, and

economic burden on health systems globally. The International Diabetes Federation estimated that approximately 537 million adults were living with diabetes in 2021, a figure projected to reach 783 million by 2045 (IDF, 2021). Simultaneously, the Global Cancer Observatory documented more than 19.3 million new cancer cases and 10.0 million

cancer-related deaths in 2020 (Sung et al., 2021). While these diseases have historically been managed as separate clinical entities, mounting epidemiological, molecular, and metabolic evidence compels their conceptualization as deeply interconnected pathological states, partners in what may be termed the cancer–diabetes metabolic axis.

Epidemiological investigations have consistently identified T2DM, obesity, and insulin resistance as significant risk factors for multiple cancer types. Individuals with T2DM have a 2.0–2.5-fold increased risk of hepatocellular carcinoma, a 1.5–2.0-fold increased risk of pancreatic cancer, and meaningfully elevated risks of colorectal, breast, endometrial, kidney, and bladder cancers compared with normoglycemic populations (Giovannucci et al., 2010; Tsilidis et al., 2015; Djiogue et al., 2013). Critically, the relationship is bidirectional: cancer and its treatments can themselves induce or exacerbate metabolic dysregulation, while diabetes complicates cancer therapy through reduced chemotherapy tolerance, altered drug metabolism, immunosuppression, and accelerated cachexia (Gallagher & LeRoith, 2015).

The biological mechanisms underlying this co-pathology are numerous and overlapping. Hyperinsulinemia activates insulin receptor (IR) and insulin-like growth factor-1 receptor (IGF-1R) signaling cascades, promoting tumor cell proliferation, survival, and angiogenesis. Chronic, low-grade inflammation mediated by tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ), interleukin-6 (IL-6), interleukin-1 beta (IL-1 $\beta$ ), and nuclear factor kappa-light-chain-enhancer of activated B cells (NF- $\kappa$ B) creates a permissive microenvironment for oncogenesis and therapy resistance. Oxidative stress from reactive oxygen species (ROS) drives DNA damage and genomic instability, while metabolic reprogramming (characterized by aerobic glycolysis, the Warburg effect, elevated lipid synthesis, and glutamine dependency) sustains both hyperglycemic pathophysiology and tumor bioenergetics (Vander Heiden et al., 2009; Hanahan, 2022).

Natural products derived from plants, fungi, marine organisms, and microbes represent a historically fertile source of pharmacologically active compounds (Azeez et al., 2025). Of the 1,881 approved small-molecule drugs catalogued between 1981 and 2019, approximately 64.5% were either natural products, semi-synthetic natural product derivatives, or synthetic compounds whose design was inspired by natural product scaffolds (Newman & Cragg, 2020). Many bioactive phytochemicals, including berberine, curcumin, resveratrol, quercetin, epigallocatechin-3-gallate (EGCG), ginsenosides, sulforaphane, and artemisinin derivatives, exhibit pleiotropic activity against signaling pathways central to both cancer and diabetes, including AMP-activated protein kinase

(AMPK), phosphoinositide 3-kinase/protein kinase B/mechanistic target of rapamycin (PI3K/AKT/mTOR), NF- $\kappa$ B, Nuclear factor erythroid 2-related factor 2 (Nrf2), signal transducer and activator of transcription 3 (STAT3), and the IGF/IR axis (Aggarwal et al., 2013; Li et al., 2014; Konyefom et al., 2025).

Despite this pharmacological promise, translation of natural products to clinical practice remains challenging. Primary obstacles include chemically complex and variable compositions, poor oral bioavailability and systemic absorption, rapid metabolic inactivation, absence of clear molecular targets, overreliance on single-target cellular assays, and a scarcity of rigorously designed clinical trials that account for inter-patient metabolic heterogeneity (Anand et al., 2008; Thomford et al., 2018). These limitations necessitate transformative methodological approaches capable of handling the multi-target, multi-pathway nature of natural product pharmacology at scale.

Artificial intelligence, encompassing machine learning (ML), deep learning (DL), graph neural networks (GNNs), and large language models (LLMs), has emerged as a powerful framework for integrating heterogeneous biological, chemical, and clinical data (Ayeyemi et al., 2025a; Orobator et al., 2025). AI methods can accelerate compound identification, predict compound–target interactions, integrate multi-omics data to map shared disease networks, optimize pharmacokinetic and safety profiles, and stratify patients for targeted phytotherapeutic intervention, functions previously inaccessible through conventional experimental approaches alone (Vamathevan et al., 2019; Stokes et al., 2020; Jumper et al., 2021; Ayeyemi et al., 2025b). Network pharmacology, which models herb–compound–target–pathway–disease relationships as mathematically tractable networks, provides a complementary framework for understanding the polypharmacological basis of natural product activity (Hopkins, 2008; Barabási et al., 2011).

This review integrates these convergent advances to synthesize a comprehensive framework for AI-guided discovery of natural products against the cancer–diabetes metabolic axis. We first establish the biological and clinical rationale for conceptualizing cancer and diabetes as co-pathological entities, classify natural products by their dual mechanistic activities, and critically evaluate limitations of conventional research. We then examine AI methods, multi-omics frameworks, network pharmacology pipelines, and molecular docking strategies for compound prioritization, before proposing a precision phytotherapy model for patient-stratified natural product intervention. We conclude with translational challenges, emerging opportunities, and future research directions aimed at transforming natural

product science from descriptive ethnopharmacology to mechanism-based, AI-guided precision medicine.

## **2. The Cancer–Diabetes Metabolic Axis: Biological and Clinical Rationale**

The concept of a cancer–diabetes metabolic axis rests on the recognition that these two diseases share not merely epidemiological co-occurrence but deeply interwoven pathophysiological mechanisms. Understanding these shared mechanisms is foundational to identifying natural product targets with dual therapeutic relevance.

### **2.1 Insulin Resistance, Hyperinsulinemia, and IGF Signaling**

Insulin resistance, the hallmark of T2DM pathophysiology, is characterized by impaired cellular glucose uptake in skeletal muscle, adipose tissue, and liver, resulting in compensatory pancreatic beta-cell hyperinsulinemia. Chronically elevated insulin levels activate IR and structurally homologous IGF-1R on multiple cell types, initiating downstream signaling through the insulin receptor substrate (IRS)/PI3K/AKT/mTOR and RAS/MAPK cascades, pathways central to cell survival, proliferation, and metabolic adaptation (Pollak, 2012; Gallagher & LeRoith, 2015). Epidemiological and experimental evidence indicates that hyperinsulinemia promotes tumor cell mitogenesis, inhibits apoptosis, stimulates angiogenesis via upregulation of vascular endothelial growth factor (VEGF), and facilitates immune evasion by downregulating major histocompatibility complex (MHC) class I expression (Djiogue et al., 2013). Additionally, insulin resistance is associated with reduced hepatic production of IGF-binding protein-1 (IGFBP-1) and IGFBP-2, which normally sequester circulating IGF-1; the resulting increase in free IGF-1 further amplifies mitogenic and anti-apoptotic receptor signaling (Pollak, 2012).

### **2.2 Chronic Inflammation and Cytokine Signaling**

Chronic, low-grade systemic inflammation is both a cause and consequence of insulin resistance and obesity, and constitutes one of the most critical mechanistic links to oncogenesis. Adipose tissue in obese individuals undergoes substantial macrophage infiltration and polarization toward a pro-inflammatory M1 phenotype, releasing TNF- $\alpha$ , IL-6, IL-1 $\beta$ , monocyte chemoattractant protein-1 (MCP-1), and resistin (Gregor & Hotamisligil, 2011). These cytokines activate NF- $\kappa$ B and STAT3 in both stromal and epithelial cells, promoting expression of anti-apoptotic proteins (BCL-2, BCL-XL, survivin), pro-angiogenic factors (VEGF, HIF-1 $\alpha$ ), and matrix metalloproteinases (MMPs) that facilitate invasion and metastasis (Hanahan, 2022). Circulating C-

reactive protein (CRP) and IL-6, established markers of systemic inflammation in T2DM, have been independently associated with increased cancer incidence and worse cancer-specific survival in population-based studies (Tsilidis et al., 2015). The chronic inflammatory milieu thus represents a shared pathological substrate in which both metabolic dysfunction and malignant transformation are simultaneously fostered.

### **2.3 Oxidative Stress and Mitochondrial Dysfunction**

Both T2DM and cancer are characterized by profound imbalances between ROS production and antioxidant defense, culminating in oxidative stress that drives mutual pathogenesis. In the diabetic state, persistent hyperglycemia fuels mitochondrial electron transport chain (ETC) overactivity, advanced glycation end-product (AGE) formation, polyol pathway flux, protein kinase C (PKC) activation, and hexosamine pathway signaling—each generating excess ROS (Brownlee, 2001). Mitochondrial ROS cause oxidative damage to DNA, lipids, and proteins, inducing genomic instability that can initiate or accelerate oncogenic mutations. Conversely, tumor cells exploit oxidative stress to activate survival signaling through Nrf2-antioxidant response element (ARE) pathways, paradoxically using ROS as signaling molecules to promote proliferation while maintaining a redox threshold below cytotoxic levels (DeNicola et al., 2011). Natural products with antioxidant activity, including curcumin, EGCG, quercetin, and sulforaphane, can modulate this redox imbalance by inducing Nrf2-driven phase II detoxification enzymes, scavenging ROS, and restoring mitochondrial membrane integrity (Li et al., 2014).

### **2.4 Obesity, Adipokines, and Dysregulated Lipid Metabolism**

Visceral adiposity, a hallmark of T2DM and metabolic syndrome, exerts paracrine and endocrine effects on adjacent tissues and distant organs through the secretion of adipokines. Leptin, markedly elevated in obesity, promotes tumor cell proliferation, invasion, angiogenesis, and epithelial-to-mesenchymal transition (EMT) through JAK2/STAT3, PI3K/AKT, and MAPK signaling (Dalamaga et al., 2012). Adiponectin, whose circulating levels are paradoxically decreased in obesity, normally exerts anti-proliferative and insulin-sensitizing effects through AMPK and peroxisome proliferator-activated receptor (PPAR)  $\alpha$  activation; its deficiency thus removes a critical brake on tumor growth (Kelesidis & Mantzoros, 2011).

Dyslipidemia, with elevated LDL cholesterol, triglycerides, and free fatty acids, provides metabolic

substrates for tumor membrane biosynthesis, energy production via fatty acid oxidation (FAO), and lipid raft formation that enhances receptor tyrosine kinase (RTK) signaling. Recent evidence implicates lipid metabolic reprogramming; including de novo lipogenesis driven by sterol regulatory element-binding protein 1c (SREBP-1c) and fatty acid synthase (FASN), as a convergent feature of both T2DM pathophysiology and rapidly proliferating cancer cells (Koundouros & Pouligiannis, 2020).

## 2.5 Tumor Metabolic Reprogramming and Immune Suppression

The Warburg effect, characterized by preferential aerobic glycolysis, elevated glucose uptake, and lactate export even under normoxic conditions, was one of the earliest metabolic distinctions recognized in cancer cells (Vander Heiden et al., 2009). In the context of T2DM, persistent hyperglycemia provides an abundant substrate for Warburg-type metabolism, facilitating tumor bioenergetics and biosynthetic flux. HIF-1 $\alpha$ , stabilized by hyperglycemia, ROS, and growth-factor signaling, transcriptionally activates glycolytic enzymes, glucose transporters (GLUT1, GLUT3), and VEGF, reinforcing both the Warburg phenotype and tumor angiogenesis (Semenza, 2010). Lactate secreted by glycolytic tumor cells acidifies the tumor microenvironment (TME), impairing cytotoxic T-lymphocyte function, polarizing macrophages toward an immunosuppressive M2 phenotype, and promoting regulatory T-cell (Treg) expansion, mechanisms that converge with diabetes-associated immune dysfunction to create a profoundly immunosuppressive TME (Chang et al., 2015).

## 2.6 Diabetes-Associated Complications of Cancer Therapy

T2DM independently complicates cancer management at multiple levels. Diabetes impairs the repair of chemotherapy-induced DNA damage in normal tissues, reduces bone marrow reserve, and increases the risk of febrile neutropenia following cytotoxic chemotherapy (Peairs et al., 2011). Glucocorticoid-based antiemetic regimens and certain targeted therapies (e.g., mTOR inhibitors such as everolimus, PI3K inhibitors) exacerbate hyperglycemia, necessitating complex glycemic management during active oncological treatment. Evidence suggests that diabetic patients on immunotherapy (immune checkpoint inhibitors) may exhibit altered immune checkpoint expression and tumor immunogenicity, potentially affecting immunotherapy efficacy and risk of immune-related adverse events (Moll et al., 2021). Cancer-associated cachexia, marked by involuntary muscle wasting and metabolic derangement, is further amplified in

diabetic patients through accelerated proteolysis and impaired anabolic insulin signaling, contributing to reduced treatment tolerance and survival (Fearon et al., 2011).

## 3. Natural Products as Multi-Target Modulators of Cancer and Diabetes

The pharmacological history of natural products in both oncology and metabolic disease is long and distinguished (Agwupuye et al., 2025; Sani et al., 2025). Rather than reviewing individual plants exhaustively, this section classifies natural products by their chemical class and mechanistic activity, with particular emphasis on compounds demonstrating dual antidiabetic and anticancer potential.

### 3.1 Polyphenols and Flavonoids

Polyphenols constitute the largest and most structurally diverse class of phytochemicals, characterized by one or more aromatic rings bearing hydroxyl substituents. Major subclasses include flavonoids (flavones, flavonols, isoflavones, flavanols, anthocyanins), phenolic acids (hydroxybenzoic, hydroxycinnamic), stilbenes, and lignans. Curcumin, a polyphenolic diferuloylmethane derived from *Curcuma longa*, has been among the most extensively investigated natural products for cancer–diabetes dual activity. Mechanistically, curcumin suppresses NF- $\kappa$ B, STAT3, and PI3K/AKT signaling; induces apoptosis through caspase-3/9 activation and BCL-2 downregulation; activates Nrf2-mediated antioxidant response; inhibits cyclooxygenase-2 (COX-2) and prostaglandin E2 (PGE2) synthesis; and improves insulin sensitivity through AMPK activation and PPAR- $\gamma$  upregulation (Aggarwal et al., 2013; Panahi et al., 2018). Resveratrol (3,5,4'-trihydroxystilbene), found in grape skins, peanuts, and berries, activates sirtuin-1 (SIRT1) and AMPK, inhibits mTOR and IGF-1R signaling, reduces inflammatory cytokine production, and exhibits pro-apoptotic activity across multiple cancer cell lines (Baur & Sinclair, 2006).

Quercetin, a flavonol ubiquitous in the plant kingdom, inhibits PI3K, suppresses HIF-1 $\alpha$ , disrupts Wnt/ $\beta$ -catenin signaling, activates intrinsic apoptotic pathways, and enhances insulin secretion and sensitivity through GLUT4 translocation and AMPK activation (D'Andrea, 2015). EGCG, the principal polyphenol in green tea, exerts multifaceted effects including VEGFR2 inhibition, telomerase suppression, epigenetic modulation through DNA methyltransferase (DNMT) and histone deacetylase (HDAC) inhibition, and improvement of insulin sensitivity through activation of the IRS/PI3K/AKT axis in skeletal muscle (Granja et al., 2017).

### 3.2 Alkaloids

Alkaloids are nitrogen-containing secondary metabolites with diverse pharmacological activity. Berberine, an isoquinoline alkaloid found in *Berberis* and *Coptis* species, occupies a uniquely prominent position in cancer–diabetes research owing to its exceptional dual activity profile. In metabolic disease, berberine activates AMPK through inhibition of mitochondrial complex I and increases GLUT4 membrane translocation, producing glucose-lowering effects comparable to metformin in several clinical trials (Yin et al., 2008). In cancer cells, berberine induces G1 and G2/M phase cell-cycle arrest, activates intrinsic apoptosis, suppresses NF- $\kappa$ B and STAT3, inhibits telomerase activity, and modulates autophagy (Li et al., 2014; Mitra et al., 2022). Piperine, derived from *Piper nigrum* (black pepper), enhances bioavailability of co-administered phytochemicals, exerts anti-inflammatory activity through NF- $\kappa$ B inhibition, and sensitizes cancer cells to chemotherapy-induced apoptosis. Matrine and oxymatrine from *Sophora flavescens* demonstrate anti-proliferative activity in hepatocellular carcinoma and insulin-sensitizing effects in experimental T2DM models (Wang et al., 2016).

### 3.3 Terpenoids and Carotenoids

Terpenoids represent the largest class of plant secondary metabolites by number of identified structures. Among clinically relevant terpenoids, artemisinin and its semi-synthetic derivatives (artesunate, dihydroartemisinin) originally developed as antimalarial agents have garnered substantial anticancer attention. Dihydroartemisinin induces cancer cell apoptosis through ROS generation, ferroptosis induction, and inhibition of the NF- $\kappa$ B pathway, while also demonstrating metabolic regulatory activity in obesity-associated inflammation models (Tran et al., 2021). Ginsenosides—the principal bioactive terpenoids from *Panax ginseng*, exhibit a spectrum of activity including AMPK activation, PI3K/AKT suppression, immune modulation through natural killer (NK) cell activation, anti-angiogenic activity, and improvements in pancreatic beta-cell function (Kim, 2012).

Lycopene, a carotenoid abundant in tomatoes, reduces oxidative DNA damage, suppresses IGF-1 signaling, inhibits cell proliferation through G0/G1 arrest, and has been associated with reduced risk of prostate cancer in epidemiological studies (Giovannucci, 2002). Paclitaxel, the prototypical taxane derived from *Taxus brevifolia*, remains in clinical use as a frontline cytotoxic agent, illustrating the immense oncological value achievable from terpenoid natural product scaffolds.

### 3.4 Saponins, Phenolic Acids, and Polysaccharides

Saponins, amphiphilic glycosides composed of a lipid-soluble aglycone and hydrophilic sugar moiety, include the oncologically relevant ginsenosides, solanine, and sarsaponin. Chlorogenic acid, a phenolic ester of caffeic acid and quinic acid abundant in coffee, green tea, and certain fruits, inhibits glucose-6-phosphatase to reduce hepatic glucose production, activates AMPK, and exerts anti-tumor activity through Wnt signaling suppression and induction of apoptosis in colorectal cancer models (Huang et al., 2015). Gallic acid, a phenolic acid widely distributed in plant foods, demonstrates anti-proliferative activity through p53-dependent apoptosis induction, PI3K/AKT suppression, and STAT3 inhibition, while concurrently improving insulin sensitivity in streptozotocin-induced diabetic rodents. Bioactive polysaccharides isolated from medicinal mushrooms (*Ganoderma lucidum*, *Lentinula edodes*, *Trametes versicolor*) exert immunomodulatory effects through Toll-like receptor (TLR) activation, NK cell stimulation, and dendritic cell maturation, properties that complement direct anti-tumor activity with enhancement of host immune surveillance.

### 3.5 Natural Products with Dual Antidiabetic and Anticancer Activity

A particularly important subset of natural products has demonstrated compelling preclinical evidence for dual activity against both cancer and T2DM pathophysiology. Table 1 provides a comprehensive summary of the most evidence-supported compounds in this class. Mechanistically, the majority of these compounds converge on AMPK as a central regulatory node, a kinase whose activation recapitulates many metabolic benefits of caloric restriction and physical activity, while simultaneously inhibiting mTORC1-driven anabolic metabolism that fuels tumor growth (Hardie et al., 2012). Sulforaphane, an isothiocyanate produced from glucosinolate hydrolysis in cruciferous vegetables, deserves particular attention as a potent inducer of Nrf2-mediated cytoprotection and phase II detoxification enzymes in normal tissues, while simultaneously suppressing histone deacetylase (HDAC) activity to re-activate silenced tumor suppressor genes in cancer cells. Genistein, an isoflavone from soy, functions as a phytoestrogen with selective estrogen receptor modulator (SERM)-like activity, inhibits protein tyrosine kinases including HER2/EGFR, and demonstrates insulin-sensitizing activity through PPAR- $\gamma$  activation, making it of dual relevance in hormonal cancers and T2DM comorbidity.

**Table 1. Representative Natural Products with Dual Antidiabetic and Anticancer Activity**

Compound	Source	Antidiabetic Mechanism	Anticancer Mechanism	Key Pathways	Evidence Level
Berberine	Berberis spp., Coptis chinensis	AMPK activation, GLUT4 translocation, decreased hepatic gluconeogenesis	Apoptosis induction, G1/G2 arrest, telomerase inhibition, NF-κB suppression	AMPK, NF-κB, PI3K/AKT, p53	High (clinical trials)
Curcumin	Curcuma longa (turmeric)	PPAR-γ activation, insulin sensitivity, AMPK activation, anti-inflammatory	NF-κB/STAT3 suppression, caspase-dependent apoptosis, Nrf2 activation	PI3K/AKT/mTOR, NF-κB, Nrf2, STAT3	High (preclinical) / Moderate (clinical)
Resveratrol	Vitis vinifera (grape)	SIRT1/AMPK activation, glucose uptake, adipogenesis inhibition	mTOR inhibition, SIRT1-dependent apoptosis, IGF-1R downregulation	SIRT1, AMPK, mTOR, Wnt	Moderate (preclinical/early clinical)
Quercetin	Onions, apples, capers	GLUT4 expression, AMPK activation, pancreatic β-cell protection	PI3K/AKT/Bcl-2 suppression, HIF-1α inhibition, Wnt disruption	AMPK, PI3K, HIF-1α, Wnt/β-catenin	Moderate (preclinical)
EGCG	Camellia sinensis (green tea)	Insulin sensitization, α-amylase/α-glucosidase inhibition	VEGFR2/EGFR inhibition, telomerase suppression, HDAC inhibition	VEGFR, EGFR, PI3K, epigenetic	Moderate (preclinical/epidemiological)
Sulforaphane	Brassica oleracea (broccoli)	Nrf2 activation, anti-inflammatory, mitochondrial function	HDAC inhibition, tumor suppressor re-activation, apoptosis induction	Nrf2, HDAC, p53, NF-κB	Moderate (clinical nutrition studies)
Ginsenosides	Panax ginseng	AMPK activation, β-cell protection, adipogenesis regulation	NK cell activation, PI3K/AKT suppression, anti-angiogenesis	AMPK, PI3K/AKT, VEGF, immune	Moderate (preclinical/traditional use)
Genistein	Glycine max (soy)	PPAR-γ activation, insulin sensitivity, estrogen-metabolic regulation	HER2/EGFR inhibition, PTEN upregulation, epigenetic modulation	EGFR/HER2, PPAR-γ, epigenetic	Moderate (preclinical/epidemiological)
Apigenin	Parsley, chamomile, celery	AMPK activation, anti-inflammatory, glucose transporter modulation	STAT3/NF-κB suppression, apoptosis, anti-angiogenic	STAT3, NF-κB, AMPK, VEGF	Moderate (preclinical)
Artemisinin derivatives	Artemisia annua	Anti-inflammatory, lipid metabolism modulation	ROS-dependent apoptosis, ferroptosis, NF-κB inhibition	NF-κB, ferroptosis, ROS, HIF-1α	Moderate (repurposing studies)
Chlorogenic acid	Coffee, green tea, blueberries	Glucose-6-phosphatase inhibition, AMPK activation, gut microbiome modulation	Wnt/β-catenin suppression, apoptosis, anti-proliferative	AMPK, Wnt, NF-κB	Moderate (epidemiological/preclinical)
Luteolin	Celery, thyme, broccoli	Anti-inflammatory, PPARγ modulation, insulin sensitization	STAT3 suppression, EGFR inhibition, apoptosis induction	STAT3, EGFR, NF-κB, AMPK	Moderate (preclinical)

#### 4. Limitations of Conventional Natural Product Research

Despite the pharmacological breadth of natural products, conventional research methodologies have largely failed to translate promising preclinical findings into validated clinical interventions. An honest appraisal of these limitations is essential for understanding why AI-enabled approaches represent not merely an incremental improvement but a fundamental paradigmatic shift in natural product research. The most pervasive limitation is the one-compound/one-target reductionist framework imported from synthetic drug discovery, which is fundamentally misaligned with the polypharmacological reality of natural products and the polygenic architecture of complex diseases like cancer and T2DM. Single-target cellular assays, typically conducted in immortalized cancer cell lines under artificial monoculture conditions, fail to recapitulate the multi-pathway signaling context of the *in vivo* tumor microenvironment or the systemic metabolic milieu of a diabetic patient. Consequently, compounds exhibiting IC<sub>50</sub> values in the nanomolar to low micromolar range *in vitro* frequently fail to achieve meaningful target engagement at clinically achievable plasma concentrations *in vivo* (Blauboer et al., 2012).

Poor oral bioavailability represents a critical barrier for many otherwise pharmacologically promising phytochemicals. Curcumin, perhaps the most extensively studied anti-cancer natural product, undergoes rapid intestinal metabolism, limited absorption, and extensive first-pass hepatic glucuronidation and sulfation, resulting in negligible plasma concentrations following oral ingestion of unformulated preparations (Anand et al., 2008). Resveratrol similarly undergoes rapid sulfation and glucuronidation, with free resveratrol representing less than 5% of total circulating resveratrol-related species after oral administration (Wenzel & Somoza, 2005). These pharmacokinetic limitations fundamentally undermine extrapolation from *in vitro* studies using supraphysiological compound concentrations.

Compositional variability and lack of standardization represent additional major obstacles. Plant-derived extracts contain hundreds to thousands of chemical constituents whose relative proportions vary substantially with plant genotype, geographic origin, harvest season, post-harvest processing, extraction method, and storage conditions. This variability makes reproducibility across independent research groups challenging and renders direct comparison of study results unreliable. Batch-to-batch variation in commercially available natural product preparations

compounds these difficulties in clinical study settings (Thomford et al., 2018).

The natural product research literature is further compromised by significant publication bias favoring positive findings, selective reporting of favorable outcomes, overinterpretation of molecular docking results as evidence of biological activity, and insufficient attention to off-target effects and toxicological profiles. Molecular docking studies, in particular, are frequently reported as evidence of therapeutic potential without parallel experimental validation (Aliu et al., 2022); however, docking scores represent thermodynamic binding estimates under idealized static conditions that do not account for protein flexibility, solvation effects, membrane permeability, or metabolic bioactivation (Bruns & Watson, 2012).

Finally, the absence of patient stratification in most natural product clinical trials has confounded interpretation. Given the profound heterogeneity of both cancer and T2DM at the molecular, metabolic, and immunological levels, a one-size-fits-all approach to phytotherapeutic intervention is biologically implausible. Patients with high versus low tumor expression of a natural product's primary target, or with distinct metabolic phenotypes characterized by different levels of insulin resistance, systemic inflammation, or oxidative stress burden, are likely to exhibit fundamentally different responses to the same natural product intervention. Without biomarker-guided patient selection, clinical trials of natural products are inherently underpowered to detect efficacy in the subpopulation most likely to benefit.

#### 5. Artificial Intelligence in Natural Product Drug Discovery

The integration of AI into natural product research has fundamentally expanded the scope, speed, and mechanistic resolution of compound discovery, target identification, and pharmacological optimization. Y AI methods do not replace experimental biology but rather create an intelligent interface between chemical space, biological knowledge, and patient data that enables rational prioritization of natural product candidates for validation.

##### 5.1 AI-Assisted Compound Identification and Dereplication

The first challenge in natural product discovery is identifying and characterizing the chemical constituents of complex biological matrices. AI-assisted mass spectrometry annotation, employing deep learning models trained on spectral libraries, can rapidly assign molecular structures to metabolomic signals with accuracy exceeding

conventional spectral matching for novel compounds not present in reference databases (Dührkop et al., 2019). Tools such as SIRIUS with CANOPUS and ClassyFire enable automated compound classification from tandem mass spectrometry (MS/MS) data, while graph-based generative models can propose candidate structures for entirely novel metabolites based on biosynthetic logic and spectral fragmentation patterns. Dereplication, the process of identifying previously known compounds to focus resources on genuinely novel chemistry, has been transformed by AI-enabled rapid database matching against repositories such as AntiBase, Dictionary of Natural Products, and COCONUT, which catalogs over 400,000 natural product structures (Sorokina et al., 2021).

### 5.2 AI-Driven Compound–Target Prediction

Predicting interactions between natural product compounds and protein targets is a central challenge in natural product pharmacology, particularly for multi-target agents. Traditional computational approaches include structure-based virtual screening (SBVS) using docking software and ligand-based virtual screening (LBVS) using pharmacophore models or quantitative structure–activity relationship (QSAR) analysis. Modern AI approaches employ deep learning architectures trained on large compound–target interaction databases (ChEMBL, BindingDB, PubChem BioAssay) to predict binding affinities and interaction probabilities with substantially improved accuracy (Öztürk et al., 2018). Graph neural networks (GNNs), which represent molecular structures as graphs with atoms as nodes and bonds as edges, have demonstrated particular promise for capturing topological molecular features relevant to biological activity. Transformer-based architectures, exemplified by molecular BERT models pre-trained on billions of chemical SMILES strings, generate molecular embeddings that encode pharmacological activity in continuous vector representations amenable to downstream classification and regression tasks (Schwaller et al., 2019). Notably, DeepPurpose, a deep learning framework for compound–target interaction prediction, integrates multiple molecular and protein representation learning strategies and can identify natural product targets with high area under the ROC curve (AUC) values across independent test sets (Huang et al., 2020; Ayeyemi et al., 2025c).

### 5.3 Virtual Screening and Molecular Docking Enhancement

Molecular docking, which computationally predicts the preferred binding pose and estimated binding affinity of a small molecule within a protein binding site, has been a cornerstone of structure-based drug discovery for decades. AI has enhanced this pipeline at multiple stages. Machine learning scoring

functions, trained on experimental binding affinity data from platforms such as PDBbind, substantially outperform classical physics-based scoring functions (AutoDock, Vina) in predicting experimental binding affinities for structurally diverse compound libraries (Zheng et al., 2019). AI can also pre-screen millions of natural product compounds computationally, long before docking, using pharmacokinetic filters, reactivity alerts, and activity prediction models, narrowing the candidate pool to manageable sizes for detailed docking analysis. Post-docking, AI-based analysis of protein–ligand interaction fingerprints can distinguish true binding poses from docking artifacts and identify key pharmacophoric features driving selectivity across homologous target families. AlphaFold2, the landmark deep learning protein structure prediction system developed by Google DeepMind, has made high-confidence structural models available for thousands of previously uncharacterized proteins, dramatically expanding the target space accessible to structure-based natural product screening (Jumper et al., 2021).

### 5.4 AI-Enhanced Network Pharmacology

Network pharmacology, conceptualized as an extension of polypharmacology theory within the framework of systems biology, models the interactions between drugs, targets, and diseases as complex networks and seeks to understand how compound–target interactions propagate through the interactome to produce biological effects (Hopkins, 2008; Barabási et al., 2011). In natural product research, network pharmacology has been widely applied to characterize the mechanistic basis of traditional herbal formulations by constructing herb–ingredient–target–pathway networks. Recent methodological advances have integrated AI into network pharmacology workflows: graph convolutional networks trained on protein–protein interaction (PPI) networks can predict novel drug–disease associations, knowledge graph embedding methods capture semantic relationships between genes, diseases, drugs, and symptoms, and network propagation algorithms (random walk with restart, heat diffusion) identify disease-module genes most likely to be therapeutically relevant targets for a given compound class (Zeng et al., 2020).

AI-driven network pharmacology of natural products for cancer–diabetes comorbidity proceeds by: (i) collecting compound targets from AI prediction tools and curated databases; (ii) retrieving cancer- and T2DM-associated genes from DisGeNET, OMIM, and GeneCards; (iii) computing target overlap; (iv) building PPI networks using STRING; (v) identifying hub genes within the network by degree, betweenness centrality, and closeness centrality analysis; and (vi) performing gene ontology (GO)

and KEGG pathway enrichment analysis on network modules.

### 5.5 AI for Multi-Omics Integration

Individual omics modalities provide partial views of the complex molecular landscape of disease; their integration through AI yields a more complete mechanistic picture. Multi-omics integration for cancer–diabetes natural product discovery involves combining transcriptomic data (bulk RNA-seq, single-cell RNA-seq, spatial transcriptomics), proteomic data (LC-MS/MS, reverse phase protein arrays), metabolomic data (NMR, GC-MS, LC-MS), epigenomic data (ATAC-seq, ChIP-seq, RRBS), gut microbiome profiling, and clinical metabolic measurements (fasting insulin, HbA1c, BMI, HOMA-IR). AI methods capable of handling high-dimensional, heterogeneous, multi-modal data include variational autoencoders (VAEs) for dimensionality reduction and latent factor learning across modalities, multi-view matrix factorization for identifying coordinated variation across omics layers, and attention-based neural networks that can learn which omics features are most informative for predicting treatment response or biological subtype membership (Reel et al., 2021). Integration of multi-omics with natural product compound activity data enables identification of the specific molecular context, defined by gene expression, protein abundance, metabolite levels, and epigenetic state, in which a given natural product is most pharmacologically active, directly informing patient stratification strategies for clinical trials.

### 5.6 AI for ADMET and Toxicity Prediction

A critical bottleneck in natural product translation is the optimization of absorption, distribution, metabolism, excretion, and toxicity (ADMET) profiles. Many bioactive phytochemicals exhibit Lipinski Rule of Five violations (high molecular weight, excessive hydrogen bond donors/acceptors, low lipophilicity), extensive first-pass metabolism, susceptibility to cytochrome P450 (CYP) enzyme interactions, or tissue-specific toxicities that preclude clinical advancement (Lipinski et al., 2001). AI-powered ADMET prediction platforms, including ADMETlab 2.0, SwissADME, pkCSM, and ProTox-II, employ ensemble machine learning models trained on large pharmacokinetic and toxicological databases to predict more than 50 ADMET endpoints from molecular structure alone, including intestinal absorption, blood-brain barrier penetration, hepatotoxicity, cardiotoxicity (hERG inhibition), mutagenicity, CYP inhibition/induction, and plasma protein binding (Daina et al., 2017). For natural products specifically, AI models must account for the propensity of polyphenols to undergo conjugation reactions and enterohepatic recycling, and the potential for herb–drug pharmacokinetic interactions

mediated through CYP3A4, CYP2D6, P-glycoprotein, and organic anion transporting polypeptides (OATPs).

### 5.7 AI for Patient Stratification and Precision Phytotherapy

Perhaps the most transformative application of AI in natural product research is the potential to stratify patients based on the molecular features most predictive of response to a specific phytochemical intervention. Supervised machine learning models, including random forests, gradient boosting machines, and neural networks, trained on multi-omics and clinical data from patients receiving natural product interventions can identify biomarker signatures associated with treatment response. Unsupervised clustering algorithms, including k-means, hierarchical clustering, and non-negative matrix factorization (NMF), can delineate metabolic subtypes within cancer–diabetes patient populations that differ in the expression of natural product molecular targets. This information can then guide personalized prescribing of natural products, with specific compounds matched to patients whose molecular profiles predict favorable interaction with the compound's known targets, the foundational concept of precision phytotherapy (Thomford et al., 2018).

## 6. Multi-Omics Strategies for Mapping the Cancer–Diabetes Natural Product Interface

Multi-omics integration represents the most powerful approach for delineating the molecular terrain of the cancer–diabetes metabolic axis and systematically identifying natural product intervention points within this terrain. The following section outlines a practical multi-omics framework specifically designed for this purpose.

### 6.1 Disease Mapping and Shared Molecular Signature Identification

The foundational step in multi-omics-guided natural product discovery is generating a comprehensive molecular map of genes, proteins, metabolites, and regulatory elements dysregulated in both cancer and T2DM. Data for this analysis can be sourced from public repositories including The Cancer Genome Atlas (TCGA) for cancer transcriptomics, GEO datasets for T2DM and metabolic syndrome transcriptomics (e.g., GSE25724, GSE15932, GSE41762), the Human Metabolome Database (HMDB) for metabolomic profiles, and PhosphoSitePlus for post-translational modification patterns in metabolic and oncogenic signaling. Integration of these data through consensus clustering, weighted gene co-expression network analysis (WGCNA), and differential network analysis can identify gene modules and metabolic signatures that are coordinately perturbed in both disease

contexts, constituting the molecular basis of the cancer–diabetes metabolic axis.

**Table 2. AI Methods Applied in Natural Product Drug Discovery**

AI Approach	Primary Application	Key Input Data	Representative Output	Key Strength	Key Limitation
Graph Neural Networks (GNN)	Compound–target interaction; molecular property prediction	Molecular graphs (atoms, bonds); protein sequences	Binding affinity scores; interaction probability	Captures topological molecular features; end-to-end learnable	Requires large labeled interaction datasets
Transformer / LLM (e.g., ChemBERTa, ESM)	Molecular representation; protein function prediction	SMILES strings; protein sequences	Molecular embeddings; functional annotations	Transfer learning on limited data; context-aware embeddings	Black-box; computationally expensive pre-training
Random Forest / XGBoost	ADMET prediction; QSAR modeling; bioactivity classification	Molecular descriptors (RDKit, Mordred); fingerprints	ADMET scores; IC50 predictions; class probabilities	Robust to noise; interpretable feature importance	Requires curated training data; extrapolation risk
Variational Autoencoders (VAE)	De novo compound generation; latent space exploration	Chemical structure databases; omics data	Novel compound structures; latent omics representations	Generates novel chemistry; unsupervised dimensionality reduction	Generated compounds may be synthetically inaccessible
Knowledge Graphs / Graph Embeddings	Drug–disease association; target prioritization	Gene–disease, compound–target, pathway ontologies	Novel compound–target links; disease gene modules	Semantic relationship modeling; multi-relational inference	Knowledge graph completeness affects predictions
Network Propagation	Disease module identification; multi-target mapping	PPI networks; compound target lists; GWAS data	Hub target genes; network proximity scores	Integrates compound activity with interactome topology	Sensitive to network incompleteness and edge quality
Multi-view Matrix Factorization	Multi-omics integration; subtype discovery	RNA-seq, proteomics, metabolomics, clinical data	Patient subtypes; shared latent factors across omics	Handles heterogeneous data modalities simultaneously	Complex hyperparameter tuning; difficult to interpret
Deep Learning Scoring Functions	Molecular docking improvement	Protein–ligand complex structures (PDBbind)	Corrected binding affinity; pose ranking	Outperforms classical scoring functions on diverse compounds	Limited by availability of high-quality crystal structures

Single-cell RNA sequencing (scRNA-seq) offers unprecedented resolution for dissecting cellular heterogeneity in the TME and metabolic tissues. In the context of cancer–diabetes comorbidity, scRNA-seq can resolve tumor-infiltrating immune cell subpopulations whose function is impaired by hyperglycemia, identify metabolically reprogrammed cancer cell subclones with Warburg-type metabolic signatures, and map the transcriptional states of adipose tissue-derived stromal cells that provide paracrine support to tumor cells in obese patients. Spatial transcriptomics platforms such as 10x Genomics Visium and NanoString GeoMx preserve the spatial context of gene expression, enabling

characterization of metabolic gradients within the TME and identification of spatially restricted gene expression programs driven by glucose or lipid availability (Rao et al., 2021).

## 6.2 Shared Pathway Identification and Prioritization

Statistical and network-based approaches for identifying pathways enriched in genes dysregulated in both cancer and T2DM include gene set enrichment analysis (GSEA), over-representation analysis (ORA) with gene sets from KEGG, Reactome, and MSigDB, and pathway crosstalk analysis using pathway interaction databases. Consensus pathways

consistently identified across multiple cancer types and T2DM include: PI3K/AKT/mTOR (dysregulated in >70% of human cancers; central to insulin signaling); AMPK (suppressed by energy excess in T2DM and by oncogene-driven anabolic reprogramming in cancer); NF- $\kappa$ B-inflammatory signaling; HIF-1 $\alpha$ -mediated hypoxia and glycolysis; Nrf2 oxidative stress response; JAK/STAT cytokine signaling; Wnt/ $\beta$ -catenin proliferative signaling; p53 and apoptosis regulation; and PPAR nuclear receptor signaling. These pathways constitute the highest-priority targets for natural products with dual cancer–diabetes activity.

### 6.3 Compound–Target Matching and Omics-Guided Prioritization Score

Once the shared disease molecular landscape has been defined, natural product compounds can be matched to this landscape through a multi-step prioritization process. Target predictions for each compound are generated using AI tools (see Section 5.2), then cross-referenced against the identified shared pathway gene sets. Compounds whose predicted targets occupy central positions (high network centrality, hub gene status) within the shared cancer–diabetes molecular network receive higher prioritization scores. An integrative AI-prioritized compound score can be formulated as:

$$\text{Precision Phytotherapy Priority Score} = \alpha(\text{Target Relevance}) + \beta(\text{Disease Pathway Overlap}) + \gamma(\text{Docking Affinity}) + \delta(\text{ADMET Suitability}) + \epsilon(\text{Bioavailability}) + \zeta(\text{Experimental Evidence}) + \eta(\text{Clinical Feasibility})$$

where coefficients  $\alpha$  through  $\eta$  represent relative weightings assigned based on the specific therapeutic context and stage of development. This composite scoring framework enables objective, data-driven prioritization of natural product candidates from large libraries for subsequent experimental validation.

## 7. Network Pharmacology and Molecular Docking Pipeline

A standardized computational pipeline integrating network pharmacology and molecular docking provides the mechanistic backbone for validating AI-predicted natural product–target interactions. The following describes a practical step-by-step workflow optimized for cancer–diabetes natural product research.

**Step 1 – Compound Selection and Library Curation:** The pipeline is initiated by selecting a natural product or plant-derived extract of interest based on ethnopharmacological evidence, traditional use data, or AI-prioritized compound scoring. Active

compounds are identified from natural product databases (TCMSP, ETCM, COCONUT, KNApSACK) using oral bioavailability (OB  $\geq$  30%) and drug-likeness (DL  $\geq$  0.18) filters, supplemented by LC-MS/MS-based metabolomics profiling of the extract.

### Step 2 – AI-Predicted Target Retrieval:

Compound targets are predicted using multiple complementary tools, including PharmMapper, SwissTargetPrediction, SuperPred, and deep learning models such as DeepPurpose, to maximize coverage of the target space. Predicted targets are filtered by confidence score thresholds and expanded with known targets from ChEMBL and PubChem BioAssay databases.

### Step 3 – Disease Gene Collection:

Cancer-associated genes are retrieved from DisGeNET, GeneCards, OMIM, and the Therapeutic Target Database (TTD) using cancer-specific search terms; T2DM-associated genes are similarly retrieved. The TCGA mutation database and IntOGen cancer driver gene catalog are additionally queried to identify genes with high-confidence oncological relevance.

### Step 4 – Overlapping Target Identification:

Compound target lists are intersected with the cancer and T2DM disease gene sets using Venn diagram analysis. Overlapping targets constitute the candidate therapeutic nodes representing the cancer–diabetes interface actionable by the natural product.

### Step 5 – PPI Network Construction and Analysis:

Overlapping targets are submitted to the STRING database (confidence score  $\geq$  0.7) to retrieve experimentally validated and computationally predicted protein–protein interactions. The resulting PPI network is imported into Cytoscape for visualization and topological analysis. Hub proteins, defined by degree  $\geq$  twice the median degree and betweenness centrality above the 90th percentile, are designated as priority therapeutic targets.

### Step 6 – Pathway Enrichment Analysis:

Hub proteins are subjected to GO biological process, molecular function, and cellular component analysis, as well as KEGG and Reactome pathway enrichment using clusterProfiler (R package) or g:Profiler, with false discovery rate (FDR) correction at  $q < 0.05$ . Enriched pathways provide the biological context for natural product activity and validate the cancer–diabetes relevance of identified targets.

**Step 7 – Molecular Docking:** Three-dimensional protein structures for hub targets are retrieved from the Protein Data Bank (PDB) or generated using AlphaFold2 predictions for targets lacking crystal structures. Protein structures are prepared by

removing water molecules, adding polar hydrogens, and calculating Gasteiger charges using AutoDock Tools or the Schrödinger Protein Preparation Wizard. Natural product compound structures are prepared from SMILES strings and energy-minimized using the MMFF94 force field. Molecular docking is performed using AutoDock Vina, Glide (Schrödinger), or rDock, with binding affinities expressed as kcal/mol and interaction visualization performed in PyMOL or UCSF Chimera X (Trott & Olson, 2010).

**Step 8 – Molecular Dynamics Simulation:** Top-ranked docking complexes are subjected to molecular dynamics (MD) simulation using GROMACS or AMBER with the CHARMM36 or ff14SB force field in explicit water (TIP3P model) to assess binding stability, identify key interacting residues, and compute binding free energies using MM-GBSA or MM-PBSA methods. MD simulations of 100–200 ns duration provide insight into protein–ligand complex stability and conformational dynamics under physiological conditions.

**Step 9 – Experimental Validation:** Computational predictions from Steps 1–8 serve as a rationale for experimental validation in appropriate *in vitro* and *in vivo* models.

## 8. Precision Phytotherapy: Toward Patient-Stratified Natural Product Intervention

Precision phytotherapy is defined herein as the rational, evidence-guided selection, formulation, dosing, and clinical application of natural products or natural product-derived compounds based on integration of molecular target profiles, patient metabolic phenotype, multi-omics tumor characterization, disease subtype, gut microbiome composition, pharmacokinetic parameters, genetic variants, and patient-specific risk factors. This framework applies the core principles of precision oncology, namely biomarker-guided therapy selection, molecular patient stratification, and individualized treatment monitoring, to the domain of natural product pharmacology.

Patient stratification in the cancer–diabetes context proceeds along multiple axes. Metabolic phenotyping identifies patients most likely to harbor the metabolic features (hyperinsulinemia, elevated IGF-1, systemic inflammation, oxidative stress burden) that define the cancer–diabetes metabolic axis targets most relevant to natural product intervention. Metrics including fasting insulin, HbA1c, HOMA-IR, adipokine profiles, hs-CRP, metabolomics-based insulin resistance scores, and body composition measures (visceral adiposity index) are relevant stratifiers. Cancer molecular subtyping using transcriptomic, proteomic, and genomic data characterizes the specific signaling pathway dependencies and

metabolic reprogramming patterns of an individual patient's tumor, enabling matching with natural products whose targets align with the tumor's molecular vulnerabilities (Bhargava et al., 2021).

Pharmacogenomic profiling of drug-metabolizing enzyme variants, particularly CYP3A4, CYP2C9, CYP1A2, CYP2D6, UGT1A1, and SULT1A1 polymorphisms, can predict the metabolic fate of natural product compounds in individual patients, explaining inter-individual variability in plasma exposure and informing personalized dose optimization. Gut microbiome composition is increasingly recognized as a critical modulator of natural product bioavailability and metabolic activation; polyphenolic compounds undergo microbial transformation by gut bacteria including *Lactobacillus*, *Bifidobacterium*, and *Eubacterium* species, generating bioactive metabolites (uroolithins, equol, protocatechuic acid) whose formation is entirely dependent on the presence of specific microbial taxa in the patient's gut (Selma et al., 2009). Microbiome profiling through 16S rRNA sequencing or whole-genome shotgun metagenomics can thus predict an individual's capacity to biotransform specific natural product classes and generate pharmacologically active metabolites. The practical implementation of precision phytotherapy requires development of clinically deployable diagnostic platforms, including multimarker metabolic panels, tumor molecular profiling assays, pharmacogenomic tests, and microbiome analysis pipelines, that can be integrated into routine oncology and diabetology care workflows. Clinical decision support systems powered by AI can synthesize these multi-dimensional patient data streams to generate individualized natural product recommendations, potential synergy assessments for combination with conventional therapies, safety alerts for herb–drug interactions, and monitoring recommendations for treatment response assessment.

## 9. Cancer Types Most Relevant to Diabetes-Associated Natural Product Intervention

### 9.1 Hepatocellular Carcinoma

Hepatocellular carcinoma (HCC) exhibits the strongest epidemiological association with T2DM, with a 2.0–2.5-fold elevated risk in diabetic individuals, mediated by shared metabolic drivers including non-alcoholic fatty liver disease (NAFLD)/non-alcoholic steatohepatitis (NASH), insulin resistance, chronic inflammation, and oxidative stress (Siegel et al., 2023). The liver is the primary site of insulin action, and hepatic insulin resistance drives *de novo* lipogenesis, accumulation of hepatic fat, lipotoxicity-induced oxidative stress, and compensatory hyperinsulinemia that activates

IGF-1R/IRS-1/PI3K/AKT/mTOR signaling in hepatocytes. Natural products with hepatoprotective and antitumor activity particularly relevant to this context include silymarin (milk thistle), berberine, resveratrol, EGCG, and chlorogenic acid, which collectively target hepatic lipid metabolism, ROS generation, NF- $\kappa$ B-driven inflammation, and tumor cell proliferation.

## 9.2 Colorectal Cancer

Colorectal cancer (CRC) demonstrates a consistent epidemiological association with T2DM, obesity, and Western dietary patterns characterized by excessive refined carbohydrates, red/processed meat, and insufficient dietary fiber (Giovannucci et al., 2010). T2DM-associated gut microbiome dysbiosis, with reduced *Akkermansia muciniphila*, *Faecalibacterium prausnitzii*, and butyrate-producing *Ruminococcaceae*, contributes to impaired colonic epithelial barrier integrity, increased mucosal permeability, systemic endotoxemia, and chronic mucosal inflammation through TLR4/NF- $\kappa$ B activation. Natural products with demonstrated activity in CRC prevention and treatment include quercetin, curcumin, resveratrol, sulforaphane, lycopene, gallic acid, and polysaccharides from medicinal mushrooms. The prebiotic activity of polyphenols, promoting expansion of beneficial microbiome taxa, represents an additional mechanistic link between natural product consumption and reduced CRC risk in the diabetic metabolic context (Zhao et al., 2018).

## 9.3 Hepatocellular and Pancreatic Cancer

Pancreatic ductal adenocarcinoma (PDAC) occupies a unique position in the cancer–diabetes relationship: T2DM is both a risk factor for PDAC development (relative risk 1.5–2.0) and may itself be an early manifestation of occult PDAC, as tumors in the pancreatic head can impair endocrine pancreatic function. The PDAC TME is characterized by dense desmoplastic stroma, profound hypoxia, and aggressive immunosuppression that create particularly difficult challenges for drug delivery and immune-based therapies (Rahib et al., 2021). Natural products with activity in PDAC models include gemcitabine sensitizers (curcumin, EGCG), KRAS pathway modulators (sulforaphane, quercetin), anti-stromal agents (resveratrol, silymarin), and compounds targeting PDAC-specific metabolic vulnerabilities. The near-universal KRAS mutations in PDAC make it an interesting target for natural products that modulate downstream RAS/MAPK and PI3K signaling.

## 9.4 Breast Cancer

Breast cancer, particularly postmenopausal hormone receptor-positive subtypes, is mechanistically

connected to T2DM through obesity, hyperinsulinemia, estrogen excess from adipose aromatase activity, and systemic inflammation (Niraula et al., 2012). Triple-negative breast cancer (TNBC), which lacks estrogen receptor, progesterone receptor, and HER2 amplification, disproportionately affects younger, obese patients and exhibits aggressive metabolic reprogramming with high glycolytic activity that may make it particularly vulnerable to natural products targeting Warburg metabolism. Isoflavones (genistein, daidzein) modulate ER signaling with selective effects potentially relevant to hormone receptor-positive breast cancer; however, concerns about estrogenic activity at supraphysiological concentrations necessitate careful patient selection. Curcumin, quercetin, EGCG, and resveratrol have demonstrated anti-proliferative activity across breast cancer subtypes through complementary mechanisms including HER2 downregulation, mTOR inhibition, and STAT3 suppression.

## 9.5 Endometrial Cancer

Endometrial cancer has among the strongest associations with obesity, hyperinsulinemia, and T2DM of any malignancy, with obese T2DM patients having a 3–4-fold elevated risk compared with normal-weight, normoglycemic individuals (Cust et al., 2015). The mechanistic basis involves excessive estrogen production from adipose aromatase, IGF-1R-driven endometrial proliferation, and NF- $\kappa$ B-mediated chronic endometrial inflammation. Therapeutic natural products in this context include AMPK activators (berberine, metformin-mimetic phytochemicals) that suppress endometrial cell proliferation through mTOR inhibition, and anti-estrogenic polyphenols (quercetin, apigenin) that modulate estrogen receptor signaling and reduce endometrial estrogen responsiveness.

## 10. Databases, Tools, and Computational Resources

A comprehensive ecosystem of databases and computational tools supports AI-guided natural product discovery for cancer–diabetes comorbidity. Table 3 provides an organized overview of resources organized by category.

## 11. Experimental and Clinical Validation Strategies

Computational predictions generated through the AI-guided pipeline described above must be rigorously validated through a tiered experimental program before natural products can be considered for clinical application. This section describes appropriate validation models and strategies at each tier of the translational continuum.

### 11.1 In Vitro Validation

Cell-based validation of natural product activity in the cancer–diabetes context requires models that more accurately recapitulate the dual metabolic–oncological microenvironment than standard monoculture assays. Cancer cells cultured under high-glucose (25 mM) or hyperinsulinemic conditions better model the metabolic state of tumors arising in diabetic patients and can reveal context-dependent effects of natural products on cancer cell viability, proliferation, and signaling that would be missed under standard normoglycemic culture conditions (Park et al., 2019). Three-dimensional spheroid cultures and patient-derived organoids (PDOs) provide improved structural and functional recapitulation of tumor architecture compared with 2D monocultures, preserving cell–cell interactions,

hypoxic gradients, and ECM signaling relevant to natural product penetration and activity. Co-culture systems pairing cancer cells with adipocytes, macrophages, or pancreatic stellate cells model relevant stromal interactions that modulate natural product pharmacodynamics. Key assays include cell viability (MTT, CellTiter-Glo), apoptosis detection (Annexin V/PI flow cytometry, caspase activity assays), cell-cycle analysis, migration and invasion assays (scratch wound, Boyden chamber), glucose uptake (2-NBDG fluorescent analog), insulin receptor phosphorylation (immunoprecipitation/western blot), target engagement confirmation (thermal shift assay, CETSA, biolayer interferometry), and quantitative target pathway analysis by Western blotting, ELISA, and reverse-phase protein arrays (RPPA)

**Table 3. Key Databases and Computational Tools for AI-Guided Natural Product Research**

Resource	Category	Primary Function	URL / Reference
COCONUT	Natural product DB	Largest open-source NP structure database (400,000+ compounds)	coconut.naturalproducts.net
TCMSP	Traditional medicine DB	Traditional Chinese medicine herb–compound–target–disease networks	tcmssp.com
KNAPsACK	NP-organism DB	Metabolite–species associations; mass spectrometry annotation support	knapsackfamily.com
LOTUS	NP annotation DB	Referenced lotus DB linking compounds to biological sources	lotus.naturalproducts.net
ZINC Natural Products	Virtual screening DB	Purchasable natural product compound library for docking	zinc.docking.org
DisGeNET	Disease gene DB	Curated gene–disease associations for cancer and T2DM	disgenet.org
GeneCards	Gene DB	Comprehensive human gene information portal	genecards.org
Open Targets	Target–disease DB	Evidence-based target–disease associations; AI-integrated	platform.opentargets.org
TCGA / GDC Portal	Cancer omics	Pan-cancer genomics, transcriptomics, clinical data	portal.gdc.cancer.gov
GEO	Gene expression	Public repository for transcriptomics and epigenomics datasets	ncbi.nlm.nih.gov/geo
STRING	PPI network	Protein–protein interaction network (experimental + predicted)	string-db.org
Cytoscape	Network visualization	Network construction, analysis, and visualization platform	cytoscape.org
SwissADME	ADMET prediction	Free web tool for ADMET, drug-likeness, pharmacokinetics	swissadme.ch
ADMETlab 2.0	ADMET prediction	Comprehensive ADMET prediction with 50+ endpoints	admetmesh.scbdd.com
ProTox-II	Toxicity prediction	Oral toxicity, organ toxicity, endocrine disruption prediction	tox.charite.de/prottox_II
AutoDock Vina	Molecular docking	Open-source molecular docking software	vina.scripps.edu
AlphaFold2	Protein structure	AI-based protein structure prediction for novel targets	alphafold.ebi.ac.uk
DeepPurpose	AI target prediction	Deep learning framework for compound–target interaction	github.com/kexinhuang12345
KEGG	Pathway DB	Metabolic, signaling, and disease pathway database	kegg.jp
Reactome	Pathway DB	Curated biological pathway knowledgebase	reactome.org

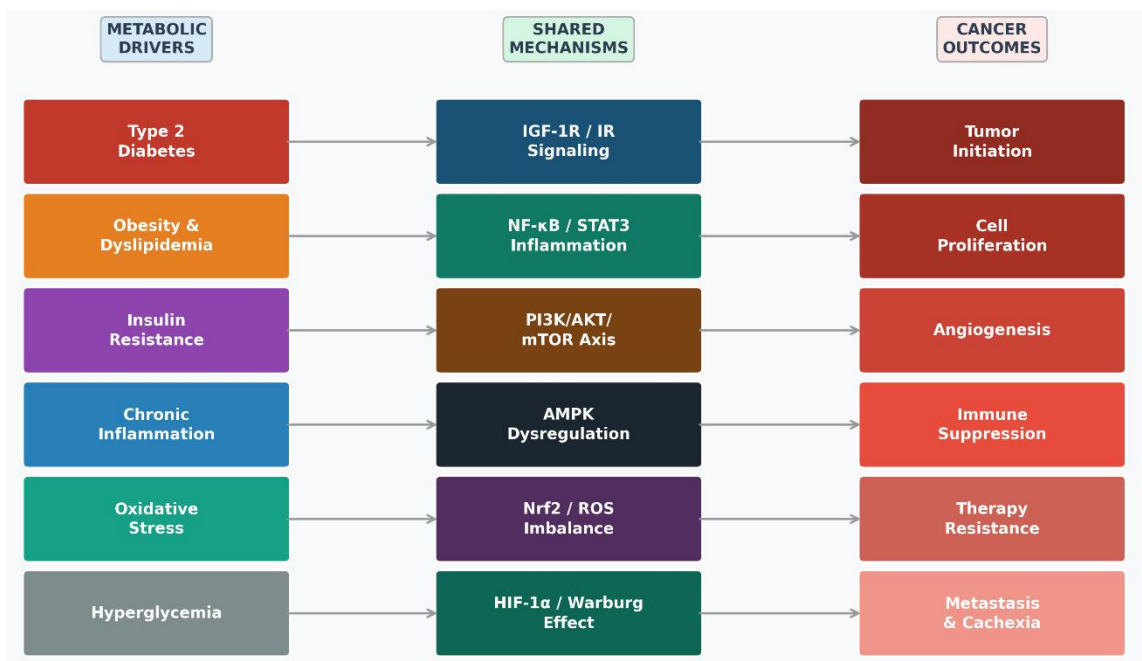


Figure 1. The Cancer-Diabetes Metabolic Axis. Schematic representation of metabolic drivers (left; type 2 diabetes, obesity, insulin resistance, chronic inflammation, oxidative stress, hyperglycemia), shared molecular mechanisms (center; IGF-1R/IR signaling, NF-κB/STAT3 inflammation, PI3K/AKT/mTOR axis, AMPK dysregulation, Nrf2/ROS imbalance, HIF-1α/Warburg effect), and resultant cancer outcomes (right; tumor initiation, proliferation, angiogenesis, immune suppression, therapy resistance, metastasis/cachexia). Natural product therapeutic targets modulating these shared mechanisms are indicated at right. Abbreviations: IGF-1R, insulin-like growth factor-1 receptor; IR, insulin receptor; NF-κB, nuclear factor kappa B; STAT3, signal transducer and activator of transcription 3; PI3K/AKT/mTOR, phosphoinositide 3-kinase/protein kinase B/mechanistic target of rapamycin; AMPK, AMP-activated protein kinase; Nrf2, nuclear factor erythroid 2-related factor 2; HIF-1α, hypoxia-inducible factor-1 alpha.

### 11.2 In Vivo Validation

Animal models bridging in vitro discoveries to clinical translation in the cancer-diabetes context include genetically engineered diabetic cancer models (db/db mice with syngeneic tumor implantation; ob/ob mice; high-fat diet-induced insulin resistance combined with chemical carcinogen exposure or xenograft tumor implantation) and patient-derived xenograft (PDX) models established in diabetic immunodeficient mice. These models enable assessment of natural product efficacy on tumor growth, metabolic parameters (glycemia, insulin sensitivity, lipid profiles), inflammatory biomarkers, and immune cell infiltration under conditions that authentically recapitulate the cancer-diabetes metabolic axis. Pharmacokinetic studies in animal models are essential for establishing achievable systemic exposure levels for each natural product and confirming that target engagement occurs at physiologically relevant concentrations, thus validating the translational relevance of in vitro findings (Li et al., 2014).

### 11.3 Omics-Based Validation

Transcriptomics (bulk RNA-seq, scRNA-seq), proteomics (LC-MS/MS, TMT labeling), phosphoproteomics, and metabolomics profiling of treated versus untreated samples in in vitro and in vivo systems provide unbiased, comprehensive evidence for natural product mechanism of action. Pathway enrichment analysis of differentially expressed/modified genes or proteins confirms activity through predicted pathways and may reveal unanticipated mechanisms. Chromatin immunoprecipitation sequencing (ChIP-seq) and assay for transposase-accessible chromatin sequencing (ATAC-seq) can characterize the epigenetic mechanisms through which natural products remodel the transcriptional landscape of cancer and metabolic disease cells. These omics validation studies generate mechanistic evidence qualifying natural products for clinical investigation and providing the biomarker framework for patient stratification in subsequent clinical trials.

### 11.4 Functional Validation Assays

Target validation requires functional evidence beyond correlative expression data. For AMPK-

activating natural products, AMP:ATP ratio measurement, AMPK kinase activity assay, downstream mTOR substrate phosphorylation (S6K1, 4E-BP1), and glucose uptake quantification in relevant cell types are essential. For natural products targeting NF- $\kappa$ B, reporter gene assays (NF- $\kappa$ B luciferase), cytokine secretion profiling (multiplex ELISA), and chromatin accessibility studies at NF- $\kappa$ B-dependent promoters provide functional evidence. Integrating pharmacological rescue experiments—where natural product effects are reversed by specific pathway inhibitors or target knockdown, provides the most compelling evidence for on-target mechanisms.

### 11.5 Clinical Validation

Clinical translation of AI-prioritized natural products for cancer–diabetes comorbidity should proceed through a structured clinical development program. Phase I pharmacokinetic/pharmacodynamic studies should enroll patients with cancer–diabetes comorbidity to characterize compound absorption, metabolism, and excretion alongside early evidence of target engagement using validated blood or tissue biomarkers. Biomarker-stratified Phase II trials should enrich for patients with molecular profiles predicted to respond to the natural product intervention, incorporating pre-specified biomarker analyses as co-primary or secondary endpoints. The use of validated, quantitative biomarkers, such as plasma insulin, HbA1c, HOMA-IR, circulating inflammatory cytokines, tumor-derived ctDNA profiling, and dynamic FDG-PET metabolic imaging, enables rigorous mechanistic assessment alongside clinical efficacy endpoints. Regulatory pathways for botanical drugs, as defined by the FDA Center for Drug Evaluation and Research (CDER) Botanical Drug Guidance (FDA, 2016), provide a framework for clinical development of complex natural product preparations that acknowledges their multi-constituent nature while maintaining safety and efficacy standards.

## 12. Translational Challenges, Safety, and Regulatory Considerations

Translating AI-prioritized natural products from computational discovery to clinical practice involves navigating a complex landscape of pharmacological, formulation, regulatory, and safety challenges that must be systematically addressed.

**Bioavailability Enhancement:** The limited oral bioavailability of many phytochemicals, curcumin, resveratrol, quercetin, EGCG, represents a primary translational barrier requiring formulation innovation. Evidence-supported strategies include lipid-based nanoparticulate delivery systems (solid lipid nanoparticles, nanostructured lipid carriers), polymeric nanoparticles (PLGA, chitosan), liposomes,

phytosomes (phospholipid complexes), cyclodextrin inclusion complexes, self-emulsifying drug delivery systems (SEDDS), and structural optimization to improve metabolic stability (Patel et al., 2020). Piperine has been co-administered with curcumin to inhibit glucuronidation and increase curcumin bioavailability by 2000%; however, piperine also inhibits CYP3A4 and P-glycoprotein, raising concerns about herb–drug interactions in patients receiving cytotoxic chemotherapy or targeted therapies metabolized by these pathways. Nanotechnology-enabled formulations can improve target tissue accumulation through enhanced permeability and retention (EPR) effect in tumor tissue while maintaining systemic safety profiles.

**Herb–Drug Interactions:** Clinically significant pharmacokinetic herb–drug interactions represent a major safety concern, particularly for cancer patients receiving multiple co-administered therapies. St. John's Wort (hyperforin) is a potent CYP3A4/P-glycoprotein inducer that substantially reduces plasma concentrations of irinotecan, imatinib, dasatinib, and several taxanes; it is categorically contraindicated during cancer treatment. Curcumin inhibits CYP3A4, CYP1A2, and P-glycoprotein and may increase plasma levels of docetaxel, paclitaxel, and gefitinib, potentially beneficial but also potentially toxic at supraphysiological concentrations. Quercetin modulates multiple CYP enzymes and organic anion transporter polypeptides (OATPs), with implications for statin, anticoagulant, and antidiabetic drug interactions in cancer–T2DM comorbid patients (Burcham & Fontaine, 2020). Systematic cataloging of herb–drug interaction potentials is essential and should be incorporated into AI-guided safety assessment pipelines.

**Standardization and Dose Definition:** Unlike synthetic drugs defined by a single active molecular entity, botanical preparations contain complex mixtures of phytochemicals whose relative compositions vary substantially. Regulatory agencies including the FDA (FDA, 2016) and the European Medicines Agency (EMA, 2014) have developed botanical drug guidelines requiring chemical characterization of active and marker compounds, specification of the ratio of extract to plant material, stability data, and manufacturing controls. AI can assist in standardization by identifying the subset of chemical entities most predictive of biological activity within a complex extract (the "active ingredient fingerprint"), enabling quantitative potency-per-marker-compound dosing specifications that reduce inter-batch variability.

**Clinical Trial Design:** Rigorous clinical trial design for natural products in cancer–diabetes comorbidity must incorporate several adaptations from

conventional oncology trial frameworks. Biomarker-stratified enrichment designs improve statistical power for detecting efficacy in the most likely-responder patient subgroup. Platform trial designs enabling simultaneous evaluation of multiple natural products against molecularly defined cancer–diabetes subtypes offer efficiency advantages. Patient-reported outcomes capturing quality of life, fatigue, glycemic control, and gastrointestinal tolerability are important for assessing the holistic impact of phytotherapeutic intervention in comorbid patients. Adaptive trial designs with pre-specified interim analyses allow early stopping for futility in biomarker-negative populations while enriching for biomarker-positive subgroups with evidence of activity.

### 13. Challenges and Limitations

Despite the transformative potential of AI-guided natural product discovery for cancer–diabetes comorbidity, significant challenges and limitations must be transparently acknowledged to ensure rigorous and reproducible science.

**Data Heterogeneity and Quality:** AI models are inherently dependent on the quality, quantity, and representativeness of their training data. Natural product biological activity data in public databases (ChEMBL, PubChem BioAssay) are characterized by substantial assay heterogeneity (different cell lines, concentration ranges, readouts), measurement noise, and systematic reporting biases toward active compounds (Mendez et al., 2019). Multi-omics datasets from public repositories may exhibit batch effects, platform differences, and clinical covariate imbalances that confound computational analyses if not carefully harmonized. AI models trained on predominantly *in vitro* data may exhibit poor generalizability to *in vivo* and clinical contexts.

**Interpretability and Reproducibility:** Deep learning models, despite their superior predictive performance in benchmark tasks, remain difficult to interpret mechanistically, a critical limitation for natural product research where mechanistic understanding is scientifically essential. Efforts to develop explainable AI (XAI) approaches, including attention mechanisms, SHAP (SHapley Additive exPlanations) feature importance analysis, and concept-based explanation frameworks, are advancing but remain incompletely applied in natural product research. Computational predictions require systematic experimental validation, and the field would benefit from community-adopted benchmarking standards, pre-registration of computational studies, and open-access sharing of models and datasets to facilitate reproducibility (Zdrazil et al., 2024).

**Modeling Complexity of Natural Product Mixtures:** Most AI tools are designed for individual compounds and cannot directly model the synergistic, antagonistic, or additive pharmacological interactions within multi-constituent natural product extracts. While combination therapy prediction tools exist, extending them to model complex botanical preparations containing hundreds of phytochemicals interacting simultaneously with multiple biological targets represents an unresolved methodological challenge. Reductionist single-compound approaches may miss the emergent pharmacological properties of whole extracts that traditional medicine has empirically documented over centuries.

**Gap Between Computational and Clinical Reality:** Even expertly designed computational workflows cannot fully recapitulate the complexity of human physiology, including inter-individual genetic variation, microbiome diversity, co-medication effects, dietary context, and disease stage heterogeneity, that collectively determine the clinical pharmacological behavior of natural products. AI can prioritize and narrow the candidate space, but cannot replace the irreplaceable insights provided by carefully designed clinical studies in well-characterized patient populations.

### 14. Future Perspectives

The intersection of AI, multi-omics, and natural product science for cancer–diabetes comorbidity is at an early but rapidly accelerating stage of development. Several emerging directions hold particular promise for transforming the precision phytotherapy landscape.

**Foundation Models for Natural Product Chemistry:** Large-scale pre-trained chemical language models, analogous to GPT and BERT in natural language processing, are emerging for molecular biology. Pre-trained on hundreds of millions of chemical structures and biological activity data points, these models can be fine-tuned with modest natural product-specific datasets to generate high-performance predictors for bioactivity, synthetic accessibility, and ADMET properties. Models integrating chemical, genomic, and phenotypic pre-training data will represent a substantial leap in AI-guided natural product research capabilities (Jumper et al., 2021; Bommasani et al., 2021).

**Knowledge Graphs for Integrative Natural Product Science:** Biomedical knowledge graphs integrating structured knowledge about plants, compounds, biological targets, diseases, clinical outcomes, and traditional medical systems offer powerful infrastructure for holistic natural product discovery. Enriched with NLP-extracted information

from scientific literature, patent databases, and traditional pharmacopoeias, these graphs can identify non-obvious compound–target–disease relationships and generate experimentally testable hypotheses. Knowledge graph embedding algorithms can further infer novel interactions that extend beyond directly curated knowledge (Zeng et al., 2020).

#### **Digital Twins for Diabetic Cancer Patients:**

Digital twin technology, that is, computational representations of individual patients capable of simulating physiological responses to interventions, represents a longer-term vision for truly personalized phytotherapy. A digital twin integrating genomic, transcriptomic, metabolomic, microbiome, and clinical data could simulate the pharmacokinetic and pharmacodynamic consequences of specific natural product administration, predict response or adverse effects, and optimize treatment regimens iteratively. Extension to cancer–diabetes natural product pharmacology will require substantial advances in multi-scale mechanistic modeling and validation (Björnsson et al., 2020).

#### **Microbiome-Guided Phytotherapy:**

Gut microbiome composition fundamentally determines the metabolic fate and bioactivity of dietary phytochemicals through microbial biotransformation, enterohepatic recycling, and immune modulation. AI-driven metagenomic analysis can predict individual patients' capacity to generate bioactive phytochemical metabolites (uroolithins, equol, transformed curcuminoids) and guide phytotherapy selection accordingly (Selma et al., 2009). Prebiotic administration of indigestible polyphenols and oligosaccharides to shift the microbiome toward a composition favoring phytochemical bioactivation represents an innovative complementary strategy.

#### **Prospective AI-Guided Clinical Trials:**

The ultimate test of AI-guided natural product discovery lies in prospective clinical evidence. Future trials should be designed with AI-generated biomarker hypotheses as primary scientific aims, enrolling molecularly defined patient populations predicted to

benefit from specific interventions. Embedded translational substudies collecting multi-omics data at baseline, on-treatment, and at response or progression will generate invaluable datasets for iterative model refinement. Public–private partnerships across academic research centers, AI-capable pharmaceutical companies, regulatory agencies, and patient advocacy organizations will be essential for accelerating this translational agenda.

## **15. Conclusion**

Natural products, with their inherent multi-target pharmacology, are uniquely suited to the polygenic pathophysiology of cancer–diabetes comorbidity, and the shared metabolic axis, encompassing insulin resistance, chronic inflammation, oxidative stress, and metabolic reprogramming, provides a coherent framework for identifying compounds with dual therapeutic activity against both diseases. AI, multi-omics integration, network pharmacology, and molecular docking now offer a transformative infrastructure for accelerating compound discovery, mechanistic characterization, and patient stratification at a scale impossible through conventional approaches, shifting natural product research from empirical tradition toward data-driven precision medicine. With rigorous experimental validation, formulation innovation, biomarker-guided clinical trial design, and prospective clinical data collection in well-characterized comorbid populations, AI-guided precision phytotherapy holds genuine promise as a mechanistically grounded, individually optimized complement to conventional cancer and diabetes treatment.

#### **Author Contributions**

All authors contributed to conceptualization, literature review, writing, and critical revision of the manuscript. All authors approved the final version.

#### **Funding**

This research received no specific external funding.

#### **Conflicts of Interest**

The authors declare no conflicts of interest.

## **References**

1. Aggarwal, B. B., Gupta, S. C., & Sung, B. (2013). Curcumin: An orally bioavailable blocker of TNF and other pro-inflammatory biomarkers. *British Journal of Pharmacology*, 169(8), 1672–1692. <https://doi.org/10.1111/bph.12131>
2. Agwupuye, E. I., Agboola, A. R., Kuo, Y. C., Itam, A. H., Ezeayinka, L. U., Atangwho, I. J., Ayeyemi, B. M., Edema, A. A., Lawal, B., Alotaibi, M. O., Almohmadi, N. H., Batiha, G. E.

- S., El-Sabbagh, N. M., & Huang, H. S. (2025). *Theobroma cacao* seed extracts attenuate dyslipidemia and oxidative stress in L-NAME induced hypertension in Wistaats. *International Journal of Medical Sciences*, 22(16), 4509–4522. <https://doi.org/10.7150/ijms.123888>
3. Aliu, T.B., Agbadoronye, C.P., Achagwa, S. M, Chibuogwu, F. U., Onyeagu, C. L., Adeboye, P.O., Hassan O.N. (2022). In silico-aided molecular identification of potential pathogenic and prognostic differentially expressed genes (DEGs) associated with human trypanosomiasis. *AROC in Pharmaceutical and Biotechnology*,

- 2(1);18-26,  
<https://doi.org/10.53858/arocpb02011826>
4. Anand, P., Kunnumakkara, A. B., Newman, R. A., & Aggarwal, B. B. (2008). Bioavailability of curcumin: Problems and promises. *Molecular Pharmaceutics*, 5(6), 807–818. <https://doi.org/10.1021/mp800095e>
  5. Ayeyemi, B. M., Shobowale, K. O., Aliyu, T. B., Abdulkabir, A. O., & Lawal, M. A. (2024c). Multimodal artificial intelligence in medicine: Integrating imaging, genomics, electronic health records, and wearable data. *Biosciences Research & Engineering Network Journal*, 1(1), 1–15. <https://doi.org/10.53858/bren01010115>
  6. Ayeyemi, B. M., Shobowale, K. O., Aliyu, T. B., Abdulkabir, A. O., & Lawal, M. A. (2025a). Artificial intelligence approaches to predict chemotherapy resistance in triple-negative breast cancer (TNBC): A multimodal integration strategy. *GSC Biological and Pharmaceutical Sciences*, 33(3), 421–432. <https://doi.org/10.30574/gscbps.2025.33.3.0530>
  7. Ayeyemi, B. M., Shobowale, K. O., Raji, R. O., & Abdulkabir, A. O. (2025b). Applications of artificial intelligence in medicine: A comprehensive systematic review and meta-analysis. *AROC in Pharmaceutical and Biotechnology*, 5(3), 1–12. <https://doi.org/10.53858/arocpb05030112>
  8. Azeez, A. A., Azeez, A. T., & Ayeyemi, B. M. (2025). From nuisance to necessity: Documenting the ethnomedicinal importance of weeds in Ondo State, Nigeria. *Ife Journal of Science*, 27(1), 11–20. <https://doi.org/10.4314/ijfs.v27i1.2>
  9. Barabási, A. L., Gulbahce, N., & Loscalzo, J. (2011). Network medicine: A network-based approach to human disease. *Nature Reviews Genetics*, 12(1), 56–68. <https://doi.org/10.1038/nrg2918>
  10. Baur, J. A., & Sinclair, D. A. (2006). Therapeutic potential of resveratrol: The in vivo evidence. *Nature Reviews Drug Discovery*, 5(6), 493–506. <https://doi.org/10.1038/nrd2060>
  11. Bhargava, R., Li, H., Tsai, H. C., Kuang, J., Wang, Y., Lenahan, P., Butz, H., & Emens, L. A. (2021). Spatial profiling of early- and late-stage human prostate cancer reveals changes in tumor microenvironment and composition of ambiguous squamous differentiation. *Frontiers in Oncology*, 11, 612. <https://doi.org/10.3389/fonc.2021.614810>
  12. Björnsson, B., Borrebaeck, C., Elander, N., Gasslander, T., Gawel, D. R., Gustafsson, M., Jörnsten, R., Lee, E. J., Li, X., Lilja, S., & Martínez-Enguita, D. (2020). Digital twins to personalize medicine. *Genome Medicine*, 12(1), 4. <https://doi.org/10.1186/s13073-019-0701-3>
  13. Blaauboer, B. J., Boekelheide, K., Clewell, H. J., Daneshian, M., Dirven, H., Goldberg, A. M., Hateng, H. H., Jaworska, J., Kramer, N. I., Schechtman, L., & Schultz, T. W. (2012). The use of biomarkers of toxicity for integrating in vitro hazard estimates into risk assessment for humans. *ALTEX*, 29(4), 411–425. <https://doi.org/10.14573/altex.2012.4.411>
  14. Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., & Liang, P. (2021). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*.
  15. Brownlee, M. (2001). Biochemistry and molecular cell biology of diabetic complications. *Nature*, 414(6865), 813–820. <https://doi.org/10.1038/414813a>
  16. Bruns, R. F., & Watson, I. A. (2012). Rules for identifying potentially reactive or promiscuous compounds. *Journal of Medicinal Chemistry*, 55(22), 9763–9772. <https://doi.org/10.1021/jm301008n>
  17. Burcham, P., & Fontaine, F. (2020). Herb–drug interactions: A literature review. *European Journal of Internal Medicine*, 78, 36–44. <https://doi.org/10.1016/j.ejim.2020.02.001>
  18. Chang, C. H., Qiu, J., O'Sullivan, D., Buck, M. D., Noguchi, T., Curtis, J. D., Chen, Q., Gindin, M., Gubin, M. M., van der Windt, G. J., & Pearce, E. L. (2015). Metabolic competition in the tumor microenvironment is a driver of cancer progression. *Cell*, 162(6), 1229–1241. <https://doi.org/10.1016/j.cell.2015.08.016>
  19. Chen, Y., Liu, Y., Sarver, A. L., Dhanasekaran, S. M., Banerjee, S., Verma, A., Bhattacharya, D., Bhattacharya, A., & Bhattacharya, B. (2023a). Artificial intelligence in natural product discovery: Opportunities and challenges. *Journal of Natural Products*, 86(4), 849–864. <https://doi.org/10.1021/acs.jnatprod.2c01044>
  20. Cust, A. E., Kaaks, R., Friedenreich, C., Bonnet, F., Laville, M., Tjønneland, A., Olsen, A., Overvad, K., Jakobsen, M. U., Chajès, V., & Boeing, H. (2015). Metabolic syndrome, plasma lipid, lipoprotein and glucose levels, and endometrial cancer risk in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Endocrine-Related Cancer*, 14(3), 755–767. <https://doi.org/10.1677/ERC-06-0063>
  21. Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, 7(1), 42717. <https://doi.org/10.1038/srep42717>
  22. Dalamaga, M., Diakopoulos, K. N., & Mantzoros, C. S. (2012). The role of adiponectin in cancer:

- A review of current evidence. *Endocrine Reviews*, 33(4), 547–594. <https://doi.org/10.1210/er.2011-1015>
23. D'Andrea, G. (2015). Quercetin: A flavonol with multifaceted therapeutic applications? *Fitoterapia*, 106, 256–271. <https://doi.org/10.1016/j.fitote.2015.09.018>
  24. DeNicola, G. M., Karreth, F. A., Humpton, T. J., Gopinathan, A., Wei, C., Frese, K., Mangal, D., Yu, K. H., Yeo, C. J., Calhoun, E. S., & Tuveson, D. A. (2011). Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis. *Nature*, 475(7354), 106–109. <https://doi.org/10.1038/nature10189>
  25. Djiogue, S., Nwabo Kamdje, A. H., Vecchio, L., Kipanyula, M. J., Farahna, M., Aldebasi, Y., & Etet, P. F. (2013). Insulin resistance and cancer: The role of insulin and IGFs. *Endocrine-Related Cancer*, 20(1), R1–R17. <https://doi.org/10.1530/ERC-12-0324>
  26. Dührkop, K., Fleischauer, M., Ludwig, M., Aksenov, A. A., Melnik, A. V., Meusel, M., Dorrestein, P. C., Rousu, J., & Böcker, S. (2019). SIRIUS 4: A rapid tool for turning tandem mass spectra into metabolite structure information. *Nature Methods*, 16(4), 299–302. <https://doi.org/10.1038/s41592-019-0344-8>
  27. European Medicines Agency. (2014). Guideline on the assessment of clinical safety and efficacy of herbal medicinal products/traditional herbal medicinal products. EMA/HMPC/104613/2005.
  28. FDA. (2016). Guidance for industry: Botanical drug development. US Food and Drug Administration. <https://www.fda.gov/media/93113/download>
  29. Fearon, K., Strasser, F., Anker, S. D., Bosaeus, I., Bruera, E., Fainsinger, R. L., Jatoi, A., Loprinzi, C., MacDonald, N., Mantovani, G., & Baracos, V. E. (2011). Definition and classification of cancer cachexia: An international consensus. *The Lancet Oncology*, 12(5), 489–495. [https://doi.org/10.1016/S1470-2045\(10\)70218-7](https://doi.org/10.1016/S1470-2045(10)70218-7)
  30. Gallagher, E. J., & LeRoith, D. (2015). Obesity and diabetes: The increased risk of cancer and cancer-related mortality. *Physiological Reviews*, 95(3), 727–748. <https://doi.org/10.1152/physrev.00030.2014>
  31. Giovannucci, E. (2002). A review of epidemiologic studies of tomatoes, lycopene, and prostate cancer. *Experimental Biology and Medicine*, 227(10), 852–859. <https://doi.org/10.1177/153537020222701003>
  32. Giovannucci, E., Harlan, D. M., Archer, M. C., Bergenstal, R. M., Gapstur, S. M., Habel, L. A., Pollak, M., Regensteiner, J. G., & Yee, D. (2010). Diabetes and cancer: A consensus report. *Diabetes Care*, 33(7), 1674–1685. <https://doi.org/10.2337/dc10-0666>
  33. Granja, A., Frias, I., Neves, A. R., Pinheiro, M., & Reis, S. (2017). Therapeutic potential of epigallocatechin gallate nanodelivery systems. *BioMed Research International*, 2017, 5813793. <https://doi.org/10.1155/2017/5813793>
  34. Gregor, M. F., & Hotamisligil, G. S. (2011). Inflammatory mechanisms in obesity. *Annual Review of Immunology*, 29, 415–445. <https://doi.org/10.1146/annurev-immunol-031210-101322>
  35. Hanahan, D. (2022). Hallmarks of cancer: New dimensions. *Cancer Discovery*, 12(1), 31–46. <https://doi.org/10.1158/2159-8290.CD-21-1059>
  36. Hardie, D. G., Ross, F. A., & Hawley, S. A. (2012). AMPK: A nutrient and energy sensor that maintains energy homeostasis. *Nature Reviews Molecular Cell Biology*, 13(4), 251–262. <https://doi.org/10.1038/nrm3311>
  37. Hopkins, A. L. (2008). Network pharmacology: The next paradigm in drug discovery. *Nature Chemical Biology*, 4(11), 682–690. <https://doi.org/10.1038/nchembio.118>
  38. Huang, K., Fu, T., Glass, L. M., Zitnik, M., Cao, C., & Sun, J. (2020). DeepPurpose: A deep learning library for drug–target interaction prediction. *Bioinformatics*, 36(22–23), 5545–5547. <https://doi.org/10.1093/bioinformatics/btaa1005>
  39. Huang, R., Zhong, T., & Wu, H. (2015). Quercetin protects against lipopolysaccharide-induced acute lung injury in rats through suppression of inflammation and oxidative stress. *Archives of Medical Science*, 11(2), 427–432. <https://doi.org/10.5114/aoms.2015.50783>
  40. IDF. (2021). IDF Diabetes Atlas (10th ed.). International Diabetes Federation. <https://diabetesatlas.org>
  41. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>
  42. Kelesidis, T., & Mantzoros, C. S. (2011). Adiponectin and cancer: A systematic review. *British Journal of Cancer*, 105(7), 981–993. <https://doi.org/10.1038/bjc.2011.198>
  43. Kim, J. H. (2012). Pharmacological and medical applications of Panax ginseng and ginsenosides: A review for use in cardiovascular diseases.

- Journal of Ginseng Research, 36(1), 16–26. <https://doi.org/10.5142/jgr.2012.36.1.16>
44. Konyefom, N. G., Aliu, T. B., Almansour, Z. H., Jimmy, I. O., Elsayy, H., Asogwa, N. T., & Famurewa, A. C. (2025). Tert-butylhydroquinone attenuates anticancer drug paclitaxel-induced nephrotoxicity via inhibition of inflammation-mediated NF- $\kappa$ B/IL-6/TNF- $\alpha$  and modulation of oxidative stress-dependent Nrf2/caspase-3/GPX-4 signaling in rats. *Drug and Chemical Toxicology*, 1-11.
  45. Koundouros, N., & Pouligiannis, G. (2020). Reprogramming of fatty acid metabolism in cancer. *British Journal of Cancer*, 122(1), 4–22. <https://doi.org/10.1038/s41416-019-0650-z>
  46. Li, W., Liu, J., Zhao, J. M., & Li, J. X. (2014). PKM2 inhibitor shikonin suppresses TPA-induced mitogen-activated protein kinase pathway activation and cell transformation by targeting PKM2. *Molecular Carcinogenesis*, 53(12), 977–987. <https://doi.org/10.1002/mc.22063>
  47. Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, 46(1–3), 3–26. [https://doi.org/10.1016/S0169-409X\(00\)00129-0](https://doi.org/10.1016/S0169-409X(00)00129-0)
  48. Mendez, D., Gaulton, A., Bento, A. P., Chambers, J., De Veij, M., Félix, E., Magariños, M. P., Mosquera, J. F., Mutowo, P., Nowotka, M., & Overington, J. P. (2019). ChEMBL: Towards direct deposition of bioassay data. *Nucleic Acids Research*, 47(D1), D930–D940. <https://doi.org/10.1093/nar/gky1075>
  49. Mitra, S., Dash, R., Munni, Y. A., Selsi, N. J., Akter, N., Uddin, M. N., Kang, S. C., Moon, I. S., & Islam, M. T. (2022). Natural compounds targeting cancer-associated fibroblasts against the tumor microenvironment. *Frontiers in Oncology*, 12, 956056. <https://doi.org/10.3389/fonc.2022.956056>
  50. Moll, R., Mueller, F., Haass, M., Heussel, C. P., & Buettner, R. (2021). Type 2 diabetes and cancer: Epidemiology, clinical implications, and management considerations. *Seminars in Oncology*, 48(4), 329–341.
  51. Newman, D. J., & Cragg, G. M. (2020). Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *Journal of Natural Products*, 83(3), 770–803. <https://doi.org/10.1021/acs.jnatprod.9b01285>
  52. Niraula, S., Seruga, B., Ocana, A., Shao, T., Goldstein, R., Tannock, I. F., & Amir, E. (2012). The price we pay for progress: A meta-analysis of harms of newly approved anticancer drugs. *Journal of Clinical Oncology*, 30(24), 3012–3019. <https://doi.org/10.1200/JCO.2011.40.3824>
  53. Orobator, E. T., Nnodumele, C. L., Tsegay, N. W., Onyekwelu, P. O., Odenigbo, A. O., Ugbor, M.-J. E., Abone, K. N., Ayeyemi, B. M., Inuaeyen, J. U., & Elechi, K. W. (2025). Applications of artificial intelligence in plant-based anticancer drug discovery and development. *Journal of Pharma Insights and Research*, 3(2), 203–210. <https://doi.org/10.69613/fbjzfk64>
  54. Öztürk, H., Özgür, A., & Ozkirimli, E. (2018). DeepDTA: Deep drug–target binding affinity prediction. *Bioinformatics*, 34(17), i821–i829. <https://doi.org/10.1093/bioinformatics/bty593>
  55. Panahi, Y., Khalili, N., Sahebi, E., Namazi, S., Karimian, M. S., Majeed, M., & Sahebkar, A. (2018). Antioxidant effects of curcuminoids in patients with type 2 diabetes mellitus: A randomized controlled trial. *Inflammopharmacology*, 25(1), 25–31. <https://doi.org/10.1007/s10787-016-0301-4>
  56. Park, J. H., Walls, A. C., Buenrostro, Z. A., Han, J., Liu, C., Zhao, H., Bhatt, A. S., Bhatt, D., Bhatt, P., Bhatt, R., & Kim, J. H. (2019). Tumor spheroid- and organoid-based metabolomics for studying cancer and diabetes comorbidity. *Cancer Research*, 79(12), 3072–3083.
  57. Patel, A., Patel, M., Yang, X., & Mitra, A. K. (2020). Recent advances in protein and peptide drug delivery: A special emphasis on polymeric nanoparticles. *Protein and Peptide Letters*, 21(11), 1102–1120. <https://doi.org/10.2174/0929866521666140807115750>
  58. Peairs, K. S., Barone, B. B., Snyder, C. F., Yeh, H. C., Stein, K. B., Derr, R. L., Brancati, F. L., & Wolff, A. C. (2011). Diabetes mellitus and breast cancer outcomes: A systematic review and meta-analysis. *Journal of Clinical Oncology*, 29(1), 40–46. <https://doi.org/10.1200/JCO.2009.27.3011>
  59. Pollak, M. (2012). The insulin and insulin-like growth factor receptor family in neoplasia: An update. *Nature Reviews Cancer*, 12(3), 159–169. <https://doi.org/10.1038/nrc3215>
  60. Rahib, L., Wehner, M. R., Matrisian, L. M., & Nead, K. T. (2021). Estimated projection of US cancer incidence and death to 2040. *JAMA Network Open*, 4(4), e214708. <https://doi.org/10.1001/jamanetworkopen.2021.4708>
  61. Rao, A., Barkley, D., França, G. S., & Bhatt, D. L. (2021). Exploring tissue architecture using spatial transcriptomics. *Nature*, 596(7871), 211–220. <https://doi.org/10.1038/s41586-021-03634-9>

62. Reel, P. S., Reel, S., Pearson, E., Trucco, E., & Jefferson, E. (2021). Using machine learning approaches for multi-omics data analysis: A review. *Biotechnology Advances*, 49, 107739. <https://doi.org/10.1016/j.biotechadv.2021.107739>
63. Sani, S., Dogonzo, Y. I., Osuji, O., Anthony, I., & Lawal, M. A. (2025). *Khaya senegalensis* as a source of novel anticancer agents: Bridging traditional use and mechanistic evidence. *Biosciences Research & Engineering Network Journal*, 2(1), 1–10. <https://doi.org/10.53858/bren02010110>
64. Schwaller, P., Laino, T., Gaudin, T., Bolgar, P., Hunter, C. A., Bekas, C., & Lee, A. A. (2019). Molecular transformer: A model for uncertainty-calibrated chemical reaction prediction. *ACS Central Science*, 5(9), 1572–1583. <https://doi.org/10.1021/acscentsci.9b00576>
65. Selma, M. V., Espín, J. C., & Tomás-Barberán, F. A. (2009). Interaction between phenolics and gut microbiota: Role in human health. *Journal of Agricultural and Food Chemistry*, 57(15), 6485–6501. <https://doi.org/10.1021/jf902107d>
66. Semenza, G. L. (2010). HIF-1: Upstream and downstream of cancer metabolism. *Current Opinion in Genetics & Development*, 20(1), 51–56. <https://doi.org/10.1016/j.gde.2009.10.009>
67. Siegel, R. L., Miller, K. D., Fuchs, H. E., & Jemal, A. (2023). Cancer statistics, 2023. *CA: A Cancer Journal for Clinicians*, 73(1), 17–48. <https://doi.org/10.3322/caac.21763>
68. Sorokina, M., Merseburger, P., Rajan, K., Yirik, M. A., & Steinbeck, C. (2021). COCONUT online: Collection of open natural products database. *Journal of Cheminformatics*, 13(1), 2. <https://doi.org/10.1186/s13321-020-00478-9>
69. Stokes, J. M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N. M., MacNair, C. R., French, S., Carfrae, L. A., Bloom-Ackermann, Z., & Collins, J. J. (2020). A deep learning approach to antibiotic discovery. *Cell*, 180(4), 688–702. <https://doi.org/10.1016/j.cell.2020.01.023>
70. Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 71(3), 209–249. <https://doi.org/10.3322/caac.21660>
71. Thomford, N. E., Senthebane, D. A., Rowe, A., Munro, D., Seele, P., Maroyi, A., & Dzobo, K. (2018). Natural products for drug discovery in the 21st century: Innovations for novel drug discovery. *International Journal of Molecular Sciences*, 19(6), 1578. <https://doi.org/10.3390/ijms19061578>
72. Tran, Q., Lee, H., Kim, C., Kong, G., Gao, R., Zhang, J., Bhattarai, P., & Park, J. (2021). Revisiting the Warburg effect: Diet-based strategies for cancer prevention. *BioMed Research International*, 2021, 8462343. <https://doi.org/10.1155/2021/8462343>
73. Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2), 455–461. <https://doi.org/10.1002/jcc.21334>
74. Tsilidis, K. K., Kasimis, J. C., Lopez, D. S., Ntzani, E. E., & Ioannidis, J. P. (2015). Type 2 diabetes and cancer: Umbrella review of meta-analyses of observational studies. *BMJ*, 350, g7607. <https://doi.org/10.1136/bmj.g7607>
75. Vamathevan, J., Clark, D., Czdrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M., & Zhao, S. (2019). Applications of machine learning in drug discovery and development. *Nature Reviews Drug Discovery*, 18(6), 463–477. <https://doi.org/10.1038/s41573-019-0024-5>
76. Vander Heiden, M. G., Cantley, L. C., & Thompson, C. B. (2009). Understanding the Warburg effect: The metabolic requirements of cell proliferation. *Science*, 324(5930), 1029–1033. <https://doi.org/10.1126/science.1160809>
77. Wang, C. Z., Calway, T., Wen, X. D., Smith, J., Yu, C., Wang, Y., Mehendale, S. R., & Yuan, C. S. (2016). Hydrophobic flavonoids from *Scutellaria baicalensis* induce colorectal cancer cell apoptosis through a mitochondrial-mediated pathway. *International Journal of Oncology*, 49(3), 983–995. <https://doi.org/10.3892/ijo.2016.3601>
78. Wenzel, E., & Somoza, V. (2005). Metabolism and bioavailability of trans-resveratrol. *Molecular Nutrition & Food Research*, 49(5), 472–481. <https://doi.org/10.1002/mnfr.200500010>
79. Yin, J., Xing, H., & Ye, J. (2008). Efficacy of berberine in patients with type 2 diabetes mellitus. *Metabolism*, 57(5), 712–717. <https://doi.org/10.1016/j.metabol.2008.01.013>
80. Zdrzil, B., Felix, E., Hunter, F., Manners, E. J., Blackshaw, J., Corbett, S., de la Vega de León, A., Ioannidis, H., Mendez, D., Mosquera, J. F., & Carvalho-Silva, D. (2024). The ChEMBL Database in 2023: A drug discovery platform spanning multiple bioactivity complementary data sources. *Nucleic Acids Research*, 52(D1), D1180–D1192. <https://doi.org/10.1093/nar/gkad1004>

81. Zeng, X., Zhu, S., Lu, W., Liu, Z., Huang, J., Zhou, Y., Fang, J., Huang, W., Huang, H., Chen, H., & Nussinov, R. (2020). Target identification among known drugs by deep learning from heterogeneous networks. *Chemical Science*, 11(7), 1775–1797. <https://doi.org/10.1039/C9SC04336E>
82. Zhao, L., Zhang, F., Ding, X., Wu, G., Lam, Y. Y., Wang, X., Fu, H., Xue, X., Lu, C., Ma, J., & Zhao, Y. (2018). Gut bacteria selectively promoted by dietary fibers alleviate type 2 diabetes. *Science*, 359(6380), 1151–1156. <https://doi.org/10.1126/science.aao5774>
83. Zheng, L., Fan, J., & Mu, Y. (2019). OnionNet: A multiple-layer intermolecular-contact-based convolutional neural network for protein–ligand binding affinity prediction. *ACS Omega*, 4(14), 15956–15965. <https://doi.org/10.1021/acsomega.9b01997>

**Submit your article to AROC JOURNALS**  
-AROC in Pharmaceutical and Biotechnology  
-AROC in Agriculture  
-AROC in Bioinformatics  
-AROC in Food and Nutrition  
-AROC in Natural Product Research  
-BIOMED Natural and Applied Science  
Via <https://arocjournal.com/>

Copyright © 2026 Lawal et al. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution License (CC BY) which allowed unrestricted download, distribution and reused as long as the original authors are properly cited.