Navigating Multi-scale Cancer Systems Biology towards Model-driven Personalized Therapeutics

Mahnoor Naseer Gondal¹, Safee Ullah Chaudhary^{1,*}

¹ Biomedical Informatics Research Laboratory, Department of Biology, Syed Babar Ali School of Science and Engineering, Lahore University of Management Sciences, Lahore, 54792, Pakistan.

Running Title: Model-driven Personalized Oncology

* Correspondence to S.U.C.

E-mail: safeeullah@lums.edu.pk

Cell: +92 321 4255171

Abstract

Rapid advancements in high-throughput omics technologies and experimental protocols have led to the generation of vast amounts of biomolecular data on cancer that now populates several online databases and resources. Cancer systems biology models built on top of this data have the potential to provide specific insights into complex multifactorial aberrations underpinning tumor initiation, development, and metastasis. Furthermore, the annotation of these single- or multi-scale models with patient data can additionally assist in designing personalized therapeutic interventions as well as aid in clinical decision-making. Here, we have systematically reviewed the emergence and evolution of (i) repositories with scale-specific and multiscale biomolecular cancer data, (ii) systems biology models developed using this data, (iii) associated simulation software for development of personalized cancer therapeutics, and (iv) translational attempts to pipeline multi-scale panomics data for data-driven *in silico* clinical oncology. The review concludes by highlighting that the absence of a generic, zero-code, panomics-based multi-scale modeling pipeline and associated software framework, impedes the development and seamless deployment of personalized *in silico* multi-scale models in clinical settings.

Keywords: multi-scale cancer modeling, personalized cancer therapeutics, cancer systems biology, data-driven oncology, *in silico* cancer systems oncology, predictive systems oncology

Introduction

In 1971, President Richard Nixon declared his euphemistic “*war on cancer*” through the promulgation of the National Cancer Act (1–4). Five decades later, with billions spent on cancer research and development, a definitive and affordable treatment for the disease still evades humankind (5). Numerous “breakthrough” treatments have gone on to exhibit adverse side effects (6,7) that lower patients' quality of life (QoL) or have reported degrading efficacies(8). At the heart of this problem lies our limited understanding of the bewildering multifactorial biomolecular complexity as well as patient-centricity of cancer.

Recent advances in biomolecular cancer research have helped factored the system-level oncological manifestations into mutations across genetic, transcriptomic, proteomic, and metabolomic scales (9–12) that also act in concert (13,14). Crosstalk between multiscale pathways comprising of these oncogenic mutations can further exacerbate the etiology of the disease (10,15–17). The combination of mutational diversity and interplay between the constituent pathways adds genetic heterogeneity and phenotypic plasticity in cancer cells (18–20). Hanahan and Weinburg (21,22) summarized this heterogeneity and plasticity into “*Hallmarks of Cancer*” – a set of progressively acquired traits during the development of cancer.

Experimental techniques such as high throughput next-generation sequencing and mass spectrometry-based proteomics are now providing specific spatiotemporal cues on patient-specific biomolecular aberrations involved in cancer development and growth. The voluminous high-throughput patient data coupled with the remarkable complexity of the disease has also given impetus to data integrative *in silico* cancer modeling and therapeutic evaluation approaches (23). Specifically, scale-specific molecular insights into key regulators underpinning each hallmark of cancers are now helping unravel the complex dynamics of the disease (24) besides creating avenues for personalized therapeutics (25,26). In this review, we will evaluate the emergence, evolution, and integration of multiscale cancer data towards building coherent and biologically plausible *in silico* models and their integrative analysis for employment in personalized cancer treatment in clinical settings.

Scale-specific Biomolecular Data and its Applications in Cancer

Rapid advancements in molecular biology research, particularly in high-throughput genomics (27), transcriptomics (28), and proteomics (29) have resulted in the generation of big data on spatiotemporal measurements of scale-specific biomolecules (Figure 1A) in physiological as well as pathological contexts (30,31). This vast and complex spatiotemporal data is expected to exceed 40 exabytes by 2025 (32), and is currently populating several online databases and repositories. These databases can be broadly categorized into seven salient database sub-types: biomolecules (33–35), pathways (36–38), networks (39–41), environment (42,43), cell lines (44–47), histopathological images (48–50), and mutations, and drug (51–56) databases, which are discussed below.

Biomolecular Databases

Biomolecular and clinical data generated from large-scale omics approaches for cancer research can be divided into four sub-categories: (i) genome, (ii) transcriptome, (iii) proteome, and (iv) metabolome (57).

(i) Genome-scale Databases: The foremost endeavor to collect and organize large-scale genomics data into coherent and accessible repositories led to the establishment of GenBank in 1986(58) (Figure 1B, Supplementary Table 1). This open-access resource now forms one of the largest and most comprehensive public databases for nucleotide sequences from large-scale sequencing projects comprising over 300,000 species (59). In a salient study employing GenBank, Diez *et al.* (60) screened breast and ovarian cancer families with mutations in BRCA1 and BRCA2 genes. Medrek *et al.* (61) employed microarray profile sets from GenBank to analyze gene levels for CD163 and CD68 in different breast cancer patient groups. To date, GenBank remains the largest and most comprehensive nucleotide database; however, its data heterogeneity poses a significant challenge in its employment in the development of personalized cancer therapeutics. Towards an improved data stratification and retrieval of genome-scale data, in 2002, Hubbard *et al.* (62) launched the Ensemble genome database. Ensemble provides a comprehensive resource for human genome sequences capable of automatic annotation and organization of large-scale sequencing data. Amongst various genome-wide studies utilizing Ensemble, Easton *et al.* (63) employed the database to identify novel breast cancer susceptibility loci.

(ii) Transcriptome-scale Databases: Although gene-level information can facilitate the development of personalized cancer models, however, gene expression varies from cell to cell. As a result, different cancer patients have divergent genetic signatures and transcript-level gene expression(64). Therefore, high-throughput transcriptomic data has the potential to provide valuable insights into the transcriptomic complexity in cancer cells and can be useful in investigating cell state, physiology, and relevant biological events (65,66)(Figure 1B, Supplementary Table 1). Towards developing a transcript information resource, in 2000, Edgar *et al.* (35,67) launched the Gene Expression Omnibus (GEO) initiative. GEO acts as a tertiary resource providing coherent high throughput transcriptomic and functional genomics data. The platform now hosts over 3800 genomic datasets and is expanding exponentially and was employed by Chakraborty *et al.* (68) to annotate chemo-resistant cell line models which helped investigate chemoresistance and glycolysis in ovarian cancers. Furthermore, patient-specificity (13,14) and mutational diversity (23) in cancer manifests across spatiotemporal scales. Hence, the availability of patient-specific data for each type of cancer has the potential to furnish valuable insights into the biomolecular foundation of the disease. In an attempt to provide cancer type-specific mutation data, Wellcome Trust's Sanger Institute (69) developed Catalogue of Somatic Mutations in Cancer (COSMIC) (69) comprising of 10,000 somatic mutations from 66,634 clinical samples. Through the employment of COSMIC, Schubbert *et al.* (70) decoded Ras activity in cancers and developmental disorders, while Weir *et al.* (71) characterized the genome of lung adenocarcinomas. The curation of patient-specific gene and protein expression data then led to the development of The Cancer Genome Atlas (TCGA) (72). TCGA also captures the copy number variations and DNA methylation profiles for different cancer subtypes. TCGA's potential (53,73) is well exhibited by Leiserson *et al.*'s (74) investigation that helped identify combinations of rare somatic mutations across pathways and protein complexes. Davis *et al.* (75) further evaluated

the genomic landscape of chromophobe renal cell carcinomas (ChRCCs) to elicit molecular patterns as clues for determining the origin of cancer cells. To facilitate data management across different cancer projects as well as to ensure data uniformity towards developing data-driven models, the International Cancer Genome Consortium (ICGC) was launched in 2010 (76). ICGC adopts a federated data storage architecture that enables it to host a collection of scale-specific data from TCGA and 24 other projects (77). Burn *et al.* (78) estimated the distribution of cytosine in liver tumor data using ICGC, while Supek *et al.* (79) compared mutation rates between different human cancers. Although numerous databases have been established to store large-scale genomic data, however, insights from an integrated analysis of genomic data across databases have the potential to provide precise biomolecular cues into complex processes and evolution in cancer. Such capability was provided by cBio Cancer Genomics Portal (cBioPortal) (80) in 2012, with multidimensional dataset retrieval, and exploration from multiple databases. The platform additionally provides data visualization tools, pathway exploration, statistical analysis, and selective data download features for seamless utilization of large-scale genomics data across genes, patient samples, projects, and databases (81). Numerous studies have effectively employed cBioPortal (82–84), in particular, Jiao *et al.* (85) evaluated the prognostic value of TP53 and its correlation with EGFR mutations in advanced non-small-cell lung cancer (NSCLC). Hou *et al.* (86) also used cBioPortal to deduce targetable genotypes which are present in young patients with lung adenocarcinomas.

(iii) Proteome-scale Databases: Gene transcriptomics is limited in providing a deterministic proteome profile (87–89). Particularly, transcripts produced in a cell can be degraded, translated inefficiently, or modified due to post-translational modification (90,91) resulting in no or a very small amount of functional protein (66). This relatively low correlation between transcriptome and proteome data was highlighted in 2019, by Bathke *et al.* (92) where it was shown that an increase in transcripts synthesis cannot be directly associated with an increase in functional response in a cell. To facilitate functional analysis, there is a need to utilize proteomic-level data, which can help to capture a more accurate quantitative assessment of complex biomolecular regulations for functional studies (Figure 1B, Supplementary Table 1). Following the successful completion of the Human Genome Project (HGP) (1998), in 2003, a group of Swedish researchers reported the Human Protein Atlas (HPA) (93,94) with an aim to map the entire set of human proteins in cells, tissues, and organs for normal as well as cancerous state (95). HPA employs large-scale omics-based technologies to localize and quantify protein expression patterns. The database has successfully managed to host comprehensive information on human proteins from cells, tissues, pathology, brain, and blood regions-related studies. HPA data can be employed for various purposes such as investigating the spatial distribution of proteins in different tissue and comparing normal and cancerous protein expression patterns across samples, etc (93). In a salient study employing HPA, Gámez-Pozo *et al.* (96) localized the expression pattern of proteins to aid in the profiling of human lung cancer subtypes, while Imberg-Kazdan *et al.* (97) identified novel regulators of androgen receptor function present in prostate cancer. Another significant stride towards proteome-level information came with the establishment of the Human Proteome Project (HPP) (98,99) in 2008 (100) by the Human Proteome Organization (HUPO). HPP consolidated mass spectrometry-based proteomics data, and bioinformatics pipelines, with the aim to organize and map the entire human proteome. To date, numerous studies have utilized

HPP towards identifying the complex protein machinery involved in cancer cell fate outcomes (101–106). Amongst the foremost attempts, in 2001, Sebastian *et al.* (107) employed the HPP platform to deduce the complex regulatory region of the human CYP19 gene (*'armatose'*), one of the contributors to breast cancer regulation. HPP project was later segmented into “biology and disease-oriented HPP” (B/D HPP) (108) and chromosome-centric HPP (C-HPP) (109). Gupta *et al.* (110) carried out an extensive analysis of existing experimental and bioinformatics databases to annotate and decipher proteins associated with glioma on chromosome 12, while, Wang *et al.* (111) performed a qualitative and quantitative assessment of human chromosome 20 genes in cancer tissue and cells using C-HPP resources.

(iv) Metabolome-scale Databases: Metabolic reprogramming is one of the earliest manifestations during tumorigenesis (112) and, therefore, potentiates identification of biomarkers involved in cancer onset, its prognosis as well as treatment. Large-scale efforts to collect metabolomics data led to the development of several online databases (113–115) (Figure 1B, Supplementary Table 1) including the Golm Metabolome Database (GMD) (113), in 2004. GMD provides a comprehensive resource on metabolic profiles, customized mass spectral libraries, along with spectral information for use in metabolite identification. GMD was employed in 2011 by Wedge *et al.* (116) to identify and compare metabolic profiles in serum and plasma samples for small-cell lung cancer patients towards determining optimal agent for onwards analysis. Whereas, in 2013, Pasikanti *et al.* (117) utilized GMD to identify biomarker metabolites present in bladder cancer. To allow for large-scale metabolic data stratification and retrieval, in 2007, Wishart *et al.* published the Human Metabolome Database (HMDB) (114). HMDB contains organism-specific information on metabolites across various biospecimens and their accompanying environments. It is now the world’s largest metabolomics database with around 114,100 metabolites that have been characterized and annotated. HMDB was employed by Sugimoto *et al.* (118) to identify environmental compounds specific to oral, breast, and pancreatic cancer profiles, while Agren *et al.* (119) constructed metabolic network models for 69 human cells and 16 cancer types. Although HMDB supports data deposition and dissemination, however, integrated exploratory analysis is not available. The Metabolomics Workbench (115), reported in 2016, provides information on metabolomics metadata and experimental data across species, along with an integrated set of exploratory analysis tools. The platform also acts as a resource to integrate, deposit, track, analyze, as well as disseminate large-scale heterogeneous metabolomics data from a variety of studies. In a case study built using this platform, Hattori *et al.* (120) studied cancer progression by metabolic reprogramming in myeloid leukaemia, whereas, Playdon (121) investigated dietary biomarkers in breast cancer etiology.

Biomolecular Pathway Databases

Investigations restricted to single specific biomolecular scales have limited translational potential as cancer dysregulation is driven by tightly coupled biomolecular pathways constituted by biomolecules from a variety of spatiotemporal scales (discussed above). Such biomolecular pathways represent organized cascades of interactions integrating different spatiotemporal scales towards reaching specific phenotypic cell fate outcomes. Numerous scale-specific and multi-omics biomolecular pathway databases now exist to help retrieve, store and analyze existing pathway information towards understanding cellular communication in light of complex

cancer regulation (38,122,123). One of the earliest attempts at integrating genomics data for pathway construction was in 1995, with the establishment of the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (122) (Figure 1C, Supplementary Table 2). Over time, KEGG has significantly expanded to include high-throughput multi-omics data (124). As a result, the resource is divided into fifteen sub-groups including KEGG Genome (for genome-level pathways), KEGG Compound (for small molecules level pathways), KEGG Gene (for gene and protein pathways), KEGG Reaction (for biochemical reaction and metabolic pathways), etc. Li *et al.* (125) used the KEGG database to perform pathway enrichment analysis for predicting the function of circular RNA (circRNA) dysregulation in cancer, while Feng *et al.* (126) identified genes associated with poor prognosis in ovarian cancer. To furnish information on pathways limited to gene-protein and protein-protein interactions for high-throughput functional analysis studies, PANTHER (protein annotation through evolutionary relationship) database was established in 2010 (38,123). PANTHER hosts information on ontological gene and protein-protein interaction pathways by leveraging GenBank and Human Gene Mutation Database (HGMD) (127), etc. Turcan *et al.* (128) employed PANTHER to perform network pathway enrichment for biological processes in differentially expressed genes to investigate IDH1 mutations in glioma hypermethylation phenotype. In order to store metabolic pathway information in a cell, Karp *et al.* (129) developed MetaCyc, a comprehensive reference database of metabolic pathways. MetaCyc is currently available as a web-based resource with metabolic pathway information which can be employed to investigate metabolic reengineering in cancers, carry out biochemistry-based studies, and explore cancer cell metabolism, etc. Tang and Aittokalli (130) demonstrated the utility of MetaCyc database by annotating a human metabolic network to evaluate multi-target cancer treatments.

Biomolecular Network Databases

The regulatory complexity of cancer is compounded by the cross-talk between multi-scale biomolecular pathways resulting in the formation of interaction networks (Figure 1D, Supplementary Table 3). One of the earliest biomolecular network databases, the Biomolecular Interaction Network Database (BIND) (131) was established in 2001, with an aim to organize biomolecular interactions between genes, transcripts, proteins, metabolites, as well as small molecules. Chen *et al.* (132) employed BIND to construct a biological interaction network (BIN) towards investigating tyrosine kinase regulation in breast cancer development. Although BIND provides a comprehensive resource of predefined interacting pathways, however, it does not contain ‘indirect’ interaction information. In contrast, the Molecular INTeraction database (MINT) (41), developed in 2002, curates existing literature to develop networks with both direct as well as indirect interactions from large-scale projects with information from genes, transcripts, proteins, promoter regions, etc. MINT can store data on “functional” interaction as well such as enzymatic properties and modifications present in a biomolecular regulatory network. The database was employed by Sun *et al.* (133) to construct a human interactome network to evaluate cancer patients' interactome, whereas, Vinayagam *et al.* (134) constructed an HIV (human immunodeficiency virus) network that helped identify novel cancer genes across genomic datasets. The Database of Interacting Proteins (DIP) (135) was developed to mine existing literature and experimental studies on biomolecules and their pathways to construct

interacting protein networks. Numerous studies on cancer research have been reported using DIP (136–138). Goh *et al.* (136) constructed a protein-protein interaction network for investigating liver cancer using network information from DIP while Zhao *et al.* (137) identified microRNA-regulated autophagy pathways in plants. To further consolidate and integrate protein interaction data across pathways as well as organisms, the Search Tool for the Retrieval of Interacting Genes/Proteins – STRING database was developed in 2005 (139). STRING provides a comprehensive text-mining and computational prediction platform which is accessible through an intuitive web interface (140). STRING database also provides interaction weights for edges between biomolecules to show an estimated likelihood for each interaction in the network (140). Mlecnik *et al.* (141) employed the STRING database to study T-cells homing factors in colorectal cancer. Later, Bindea *et al.* (142) assembled a gene-gene correlation network to investigate natural immunity to cancer in humans.

Environment Databases

Each pathway within a biomolecular network requires input cues from the extracellular environment for onward downstream signal transduction (143–145). In the case of cancer, the biomolecular milieu constituting the tumor microenvironment (TME) acts as a niche for tumor development, metastasis, as well as therapy response (146) (Figure 1E, Supplementary Table 4). Efforts to curate information from the environmental factors such as metabolites, matrisome, and other environmental compounds led to the development of MatrixDB (42) in 2011, which curated matrix-based information on interactions between extracellular proteins and polysaccharides. MatrixDB additionally links databases with information on genes encoding extracellular proteins such as Human Protein Atlas (147) and UniGene (148) as well as host transcripts information. Celik *et al.* (149) employed MatrixDB data to evaluate Epithelial-mesenchymal transition (EMT) inducers in the environment in nine ovarian cancer datasets. With an aim to host studies on extracellular matrix (ECM) proteins from normal as well as disease-inflicted tissue samples, MatrisomeDB (43) was established in 2020. The database contains 17 studies on 15 physiologically healthy murine and human tissue, and 6 cancer types from different stages (including breast cancers, colon cancers, ovarian cancers, lung cancers, and melanomas) along with other diseases. Levi-Galibov *et al.* (150) employed the MatrisomeDB to investigate the progression of chronic intestinal inflammation to colon cancer.

Cell Line Databases

In vitro cell lines derived from cancer patients have become an essential tool for clinical and translational research (151). These cell lines are defined based on gene expression profiles and morphological features which have been cataloged in various databases such as the Cancer Cell Line Encyclopedia (CCLE) (46) (Figure 1F, Supplementary Table 5). CCLE contains mutation data of 947 different human cancer cell lines coupled with pharmacological profiles of 24 anti-cancer drugs (46) for evaluating therapeutic effectiveness and sensitivity. Li *et al.* (152) employed CCLE data to investigate cancer cell line metabolism, while Hanniford *et al.* (153) demonstrated epigenetic silencing of RNA during invasion and metastasis in melanoma. Other cell line databases include Cell Line Data Base (CLDB) (47), and The COSMIC Cell Lines Project (44), and CellMinerCDB (45). The CellMinerCDB (2018) curates data from National Cancer Institute (NCI)

(154), BROAD institute (155), Sanger institute (156), and Massachusetts General Hospital (MGH) (157) and provides a platform for pharmacological and genomic analysis.

Histopathological Image Databases

Additionally, histopathological image datasets derived from the microscopic examination of tumor biopsy samples furnish information on cellular structure, function, chemistry, morphology, etc. Numerous histopathological image-based databases have been developed to store, manage and retrieve such information (Figure 1G, Supplementary Table 6). Amongst these databases, The Cancer Imaging Archive (TCIA) (48), reported in 2013, provides a multi-component architecture with various types of images including region-specific (e.g., Breast), cancer-type specific (e.g., TCGA-GBM, TCGA-BRCA), radiology, and anatomy images (e.g., Prostate-MRI). The cancer image collection in TCIA is captured using a variety of modalities including radiation treatment, X-ray, mammography, and computed tomography (CT), etc (158). Zhu *et al.* (159) exhibited the value of TCIA by predicting the risk for breast cancer recurrence, while Sun *et al.* (160) employed image data to perform a cohort study to validate a radiomics-based biomarker in cancer patients. In 2013, Image data from TCIA was integrated with The Cancer Digital Slide Archive (CDSA) (49) which hosts imaging and histopathological data and provides more than 20,00 images from 22 different cancers. The integrated datasets in CDSA demonstrate the clinical association of genomic data with pathology imaging. The whole-slide images of individual patients are also reported in CDSA, thereby linking tumor morphology with the patient's genomic and clinical data. Khosravi *et al.* (161) performed a deep convolution study, using CDSA, to distinguish heterogeneous digital pathology images across different types of cancers. To associate a relationship between a patient's genetic information and histology images Genotype-Tissue Expression (GTEx) (50) was reported in 2014. The GTEx project was created from datasets from 1000 patients containing tissue-specific information on gene expression and regulation.

Mutation and Drug Databases

Pharmacological investigations have elucidated the mechanism and efficacies of numerous cancer drugs, in clinical and preclinical studies (162,163). Databases with drug-target information can be employed in precision oncology towards designing efficacious patient-centric therapeutic strategies (Figure 1H, Supplementary Table 7). Such databases include DrugBank (164), established in 2006, which contains information from over 4100 drug entries, 800 FDA-approved small molecules, and 14,000 protein or drug target sequences. DrugBank combines drug data with drug-target information this, in turn, can provide potential cues for various applications in cancer biology including *in silico* drug target discovery, drug design, drug interaction prediction, etc. In a study employing DrugBank, Augustyn *et al.* (165) evaluated potential therapeutic targets of ASCL2 genes in lung cancers, while Han *et al.* (166) determined synergistic combinations of drug targets in K562 chronic myeloid leukemia (CML) cells to dissect functional genetic interaction (GI) networks. Further to evaluate drugs in light of the patient's genomic signature, the PanDrugs (56) database was established in 2019. The platform currently hosts data from 24 sources and 56297 drug-target information along with 9092 unique compounds and 4804 genes. Fernández-Navarro *et al.* (167) prioritized personalized drug treatments using PanDrugs, for Acute T-cell lymphoblastic leukemia (T-ALL) patients.

Altogether, the availability of voluminous high-resolution biomolecular data has enabled the development of a quantitative understanding of aberrant mechanisms underpinning hallmarks of cancer as well as create avenues for personalized therapeutic insights.

Figure 1. Overview of complex biomolecular regulation in cancer and scale-specific databases.

(A) The complexity between genomic, transcriptomic, proteomic, metabolomic, cell-level, and environmental levels in a cancerous cell. Four examples of biomolecular signaling pathways are listed e.g., Hedgehog, Notch, (Wingless) Wnt, TGF β , and AKT pathway. Stimuli from the extracellular environment signal the downstream pathway activation, in the cell, towards alternating the regulations in the proteomic, metabolomic, transcriptomic, and genomic scales, bringing out a system-level outcome in cancers. Lists **(B)** biomolecule (nucleotide, genes, proteins, and metabolites) databases such as GenBank, GEO, TCGA, HPP, HMDB, etc. **(C)** Pathways databases such as PathDB, KEGG, STRING, etc. **(D)** Networks databases such as BioGRID, DIP, BIND, etc. **(E)** Environment databases e.g., HMDB, MatrixDB, MatrisomeDB, etc. **(F)**. Cell lines databases such as CCLE, CLDB, Cellosaurus, etc. **(G)** Histopathological image database, for instance, TCIA, GTEx, TMAD, etc, and **(H)** Mutation and drug databases such as DrugBank, KEGG Drug, OncoKB, etc.

Data-driven Integrative Modeling in Cancer Systems Biology

The need to prognosticate system-level outcomes in light of oncogenic dysregulation (168,169) has led researchers to develop integrative data-driven computational models (170–177). Such models can help decode emergent mechanisms underpinning tumorigenesis as well as aid in the development of patient-centered therapeutic strategies (170–177). These *in silico* models can be broadly grouped into four salient sub-scales as biomolecular (178–180), environment (181–183), cell level (184,185), and multi-scale cancer models (186–188) (Figure 2).

Biomolecular-scale Models

In silico biomolecular models of cancer can be classified into (i) genome-scale, (ii) transcriptome-scale, (iii) proteome-scale, (iv) metabolome-scale models.

(i) Genome-scale Models: Amongst initial attempts at developing cancer gene regulation models, in 1987, Leppert *et al.* (178) developed a computational model to study the genetic locus of familial polyposis coli and its involvement in colonic polyposis and colorectal cancer. The study was also validated by Mehl *et al.* (179) in 1991, which further elucidated the formation and development of familial polyposis coli genes in colorectal cancer patients. In 2014, Stratmann *et al.* (180) developed a personalized genome-scale 3D lung cancer model to study epithelial-mesenchymal transition (EMT) by TGF β -based stimulation, while in 2016, Margolin *et al.* (189) developed a blood-based diagnostic model to help detect DNA hypermethylation of potential pan-cancer marker ZNF154. In 2017, Jahangiri *et al.* (190) employed an *in silico* pipeline to evaluate Staphylococcal Enterotoxin B for DNA-based vaccine for cancer therapy. A similar genome-scale model of DNA damage and repair was proposed by Smith *et al.* (191) to evaluate proton treatment in cancer (Figure 2A).

(ii) Transcriptome-scale Models: *In silico* models developed using transcriptomic expression data can assist in comparing gene expressions in cancer for investigating genetic heterogeneity in tumorigenesis and cancer development (192) as well as towards precision therapy (193) (Figure 2B). In 2003, Huang *et al.* (194) developed a mathematical model using a large-scale

transcriptional dataset of breast cancer patients to elucidate patterns of metagenes for nodal metastases and relapses. In another large-scale cancer transcriptomics study, 3,000 patient samples from nine different cancer types were used to decode the genomic evolution of cancer by Cheng *et al.* (195). In 2013, Chen *et al.* (196) employed an *in silico* pipeline which helped identify 183 new tumor-associated gene candidates with the potential to be involved in the development of hepatocellular carcinoma (HCC), while, in 2014, Agren *et al.* (197) reported a personalized transcriptomic data-based model to identify anticancer drugs for HCC. In 2019, Béal *et al.* (198) reported a logical network modeling pipeline for personalized cancer medicine using individual breast cancer patients' data using existing gene expression databases. The pipeline was validated in 2021 (199), using *in silico* personalized logical models to investigate melanomas and colorectal cancers in response to BRAF treatments. In a similar study conducted in 2019, Rodriguez *et al.* (200) developed a mathematical model for breast cancer using transcriptional regulation data to predict hypervariability in a large dynamic dataset and reveal the basis of expression heterogeneity in breast cancer.

(iii) Proteome-scale Models: To capture the quantitative aspects of biomolecular regulations and for functional studies (88,89), data-driven proteomic-based cancer models are essential (Figure 2C). Such models can be particularly helpful in diagnostic as well as prognostic purposes as well as for monitoring response to treatment (88,89). In a study employing proteome-level information, in 2011, Baloria *et al.* (201) carried out an *in silico* proteome-based characterization of the human epidermal growth factor receptor 2 (HER-2) to evaluate its immunogenicity in an *in silico* DNA vaccine. Akhoun *et al.* (202) simplified this approach with the development of a new prophylactic *in silico* DNA vaccine using IL-12 as an adjuvant. In 2017, Fang *et al.* (203) further employed proteome level data towards predicting *in silico* drug-target interactions for targeted cancer therapeutics, while in 2018, Azevedo *et al.* (204) designed novel glycomarkers in bladder cancer. Recently, in 2020, Lee *et al.* (205) have reported an integrated proteome model of macrophage migration in a complex tumor microenvironment. However, proteome-level models are limited in their ability to provide a complete analysis of the biomolecules present in a cell since they lack information on low molecular weight biomolecular compounds such as metabolites (112).

(iv) Metabolome-scale Models: Metabolic data-driven cancer models can be especially useful in understanding cancer cell metabolism, mitochondrial dysfunction, metabolic pathway alteration, etc (Figure 2D). In 2007, Ma *et al.* (206) developed the Edinburgh Human Metabolic Network (EHMN) model with more than 3000 metabolic reactions alongside 2000 metabolic genes for employment in metabolite-related studies and functional analysis. In 2011, Folger *et al.*'s (207) employed the EHMN model to propose a large-scale flux balance analysis (FBA) model for investigating metabolic alterations in different cancer types and for predicting potential drug targets. In 2014, Aurich *et al.* (208) reported a workflow to characterize cellular metabolic traits using extracellular metabolic data from lymphoblastic leukemia cell lines (Molt-4) towards investigating cancer cell metabolism. Yurkovich *et al.* (209) augmented this workflow in 2017 and reported eight biomarkers for accurately predicting quantitative metabolite concentration in human red blood cells. Alakwaa *et al.* (210) employed a mathematical model to predict the status of Estrogen Receptor in breast cancer metabolomics dataset, while in 2018, Azadi *et al.* (211)

used an integrative *in silico* pipeline to evaluate the anti-cancerous effects of *Syzygium aromaticum* employing data from the Human Metabolome Database (114).

Environment Models

Integrative mathematical models of environmental cues and extracellular matrix can help researchers abstract tumor microenvironments. Such data-driven models can be used to study angiogenesis (212), cell adhesion (213), and vasculature (214) (Figure 2E). Amongst initial attempts at developing cancer environment-based *in silico* models, in 1972, Greenpan *et al.* (181) developed a solid carcinoma *in silico* model to evaluate cancer cell behavior in limited diffusion settings. In 1976, the model was expanded (182) to investigate tumor growth in asymmetric conditions. In 1996, Chaplain developed a mathematical model to elucidate avascular growth, angiogenesis, and vascular growth in solid tumors (183). Anderson and Chaplain (215), in 1998, expanded this strategy and reported a continuous and discrete model for tumor-induced angiogenesis. This modeling approach was further augmented in 2005 by Anderson (216) to a hybrid mathematical model of a solid tumor to study cellular adhesion in tumor cell invasion. Organ-specific metastases and associated survival ratios in small cell lung cancer patients have been modeled and evaluated (217) using similar models.

Cell Models

To model cell population-level behavior in cancer, researchers are increasingly developing innovative cell line *in silico* models which can complement *in vivo* wet-lab experiments, while overcoming wet-lab limitations (218) (Figure 2F). Such models are employed to investigate cell-to-cell interactions and evaluate physical features of the synthetic extracellular matrix (ECM) (218), etc (184,219). Amongst initial attempts at developing *in silico* cell line models, in 1989, Shackney *et al.* (184) reported an *in silico* cancer cell model to study tumor evolution. Results showed an association between discrete aneuploidy peaks with the activation of growth-promoting genes. In 2007, Aubert *et al.* (219) developed an *in silico* glioma cell migration model and validated cell migration preferences for homotypic and heterotypic gap junctions with experimental results. Gerlee and Nelander (185) expanded this work, in 2012, to investigate the effect of phenotype switching on glioblastoma growth. Although cell-based cancer models helped provide scale-specific insights into cancer, however, they remain limited in investigating the spatiotemporal tissue diversity and heterogeneity in cancer patients.

Multi-scale Models

Recently, data-driven multi-scale models are becoming increasingly popular in cancer (220) (Figure 2G). One of the earliest attempts at developing multi-scale cancer models was in 1985 when Balding developed a mathematical model to demonstrate tumor-induced capillary growth (212). In 2000, Swanson *et al.* (221) proposed a quantitative model to investigate glioma cells. In a similar study, Zhang *et al.* (222) generated a 3D, multiscale agent-based model of the brain to study the role (EGFR)-mediated activation of signaling protein phospholipase role in a cell's decision to either proliferate or migrate. In 2010, Wang *et al.* (187) also took a multi-scale agent-based modeling approach to identify therapeutic targets in concurrent EGFR-TGF β signaling pathway in non-small cell lung cancer (NSCLC). Later in 2011, (188) they employed the approach

to also identify critical molecular components in NSCLC. Similarly, Perfahl *et al.* (223) developed a multi-scale vascular tumor growth model to investigate spatiotemporal regulations in cancer and response to therapy. In 2007, Anderson *et al.* (224) proposed a mathematical model for studying cancer growth, evolution, and invasion. This model was later built upon by Chaudhary *et al.* (173) with a multi-scale modeling strategy to investigate tumorigenesis induced by mitochondrial incapacitation in cell death, in 2011. In 2017, Vavourakis *et al.* (225) developed a multi-scale model to investigate tumor angiogenesis and growth. In 2017, Norton *et al.* (226) used an agent-based computational model of three negative breast cancer to study the effects of migration of CCR5+ cancer cells, stem cell proliferation, and hypoxia on the system. They later (227) reported an agent-based and hybrid model to investigate tumor immune microenvironment, in 2019. In the same year, Karolak *et al.* (228) modeled *in silico* breast cancer organoid morphologies (229) to help elucidate efficacies amongst drug treatment based on the morphophenotypic classification. Similarly, Berrouet *et al.* (230) employed a multi-scale mathematical model to evaluate the effect of drug concentration on monolayers and spheroid cultures.

Summarily, data-integrative computational models have now assumed the forefront in decoding the complex biomolecular regulations involved in cancer and are increasingly been employed for the development of personalized preclinical models as well as therapeutics design (231).

Figure 2. Evolution timeline of *in silico* scale-specific and multi-scale data-drive cancer models. Timeline of salient *in silico* scale-specific and multi-scale cancer models, along with PubMed yearly report (1990-2020) to display the evolutionary trends seen in the development of (A) genome-scale cancer models, (B) Transcript-level cancer models, (C) Proteome-scale models, (D) Metabolome scale models, (E) Environment-based models, (F) Cell-level models, and (G) Multi-scale cancer models.

Software Platforms for Modeling in Cancer Systems Biology

Over the past decade, *in silico* modeling of cancer has gained significant popularity in systems biology research (170–177). In particular, data-drive computational models are now acting as an enabling technology for precision medicine and personalized treatment of cancer. To date, several single and multi-scale software platforms have been reported to model biomolecular (180,232–239), environmental (240), cell-level (241–243), and multi-scale (244–246) information into coherent *in silico* cancer models (Figure 3).

Biomolecular Modeling Platforms

Software platforms aimed at modeling biomolecular entities, abstract information from published literature as well as high-throughput technologies, and model them using a Boolean modeling approach or a differential equations modeling strategy.

Boolean Modeling Software: Boolean network modeling technique was first introduced by Kauffman (247,248), in 1969. This approach has been widely adopted as a tool to model gene, transcript, protein, and metabolite regulatory networks. Numerous mathematical and computational cancer models have been developed using this representation (180,232–236). To facilitate the Boolean model development and analysis process, several platforms have been devised (249–251) (Supplementary Table 8). The applicability of these platforms can be further categorized into qualitative or quantitative Boolean networks modeling.

(i) Qualitative Boolean Modeling: Biomolecular qualitative Boolean models are a widely employed approach in cancer systems biology research to cater for cases where there is insufficient quantitative information, and/or lack of mechanistic understanding. Thus far, numerous platforms have been reported to help researchers develop qualitative Boolean network models. Amongst them, FluxAnalyzer (252), reported by Klamt *et al.* in 2003, was developed to undertake metabolic pathway construction, flux optimization, topological feature detection, flux analysis, etc. To expand the scope of the platform and include cell signaling, gene as well as protein regulatory networks, in 2007, Klamt *et al.* expanded FluxAnalyzer and developed CellNetAnalyzer (237). Tian *et al.* (253) employed CellNetAnalyzer to develop a p53 network model for evaluating DNA damage in cancer, while Hetmanski *et al.* (254) designed MAPK-driven feedback loop in Rho-A-driven cancer cell invasion. Although CellNetAnalyzer remains a widely used logical modeling software, however, its programmability and MATLAB dependency hinders its clinical employment towards the development of personalized cancer models. Towards addressing this challenge, in 2008, Albert *et al.* reported BooleanNet (250), an open-source, freely available Boolean modeling software for large-scale simulations of dynamic

biological systems. Saadatpoort *et al.* (255) employed BooleanNet's general asynchronous (GA) method to deduce therapeutic targets for granular lymphocyte leukemia. Similarly, in 2008, Kachalo *et al.* reported NET-SYNTHESIS (238); a platform for undertaking network synthesis, inference, and simplification. Steinway *et al.* (256) employed both BooleanNet and NET-SYNTHESIS platforms to model epithelial-to-mesenchymal transition (EMT) in light of TGF β cell signaling, in hepatocellular carcinoma patients towards elucidating potential therapeutic targets. In particular, BooleanNet was used to undertake model simulation and NET-SYNTHESIS for carrying out network interference and simplification. Another logical modeling platform, GIMsim (257), published by Naldi *et al.*, in 2009, also employed asynchronous state transition graphs to perform qualitative logical modeling which is especially useful for networks with large state space. This platform was employed by Flobak *et al.* (258) to map cell fate decisions in gastric adenocarcinoma cell-line towards evaluating drug synergies for treatment purposes, while Remy *et al.* (259) studied mutually exclusive and co-occurring genetic alterations in bladder cancer. GIMsim also implemented multi-valued logical functions, useful in simulating qualitative dynamical behavior in cancer research. However, the platform was unable to program automatic theoretical predictions, moreover, it only employed qualitative analysis approaches and could not be used to accurately map cell fates based on quantitative biomolecular expression data. Taken together, classical qualitative Boolean modeling approaches remain limited in developing predictive cancer models that could leverage quantitative biomolecular expression data generated from next-generation proteomics and related-sequencing projects.

(ii) Quantitative Boolean Modeling: Platforms aimed to integrate quantitative expression data from existing literature and databases towards carrying out network annotation and onwards analysis can be particularly useful in developing personalized cancer models. One such platform, the Markovian Boolean Stochastic Simulator (MaBoss), was established by Stoll *et al.* (260) in 2017, for stochastic and semi-quantitative Boolean network model development, mutations, and drug evaluation, sensitivity analysis based on experimental data, and eliciting model predictions. In 2019, Béal *et al.* (198) employed MaBoss to develop a logical model to evaluate breast cancer in light of individual patients' genomic signature for personalized cancer medicine. This model was later expanded, in 2021, to investigate BRAF treatments in melanomas and colorectal cancer patients (199). Similarly, Kondratova *et al.* (261) used MaBoss to model an immune checkpoint network to evaluate the synergistic effects of combined checkpoint inhibitors in different types of cancers. In a similar attempt at developing quantitative Boolean networks, BoolNet (262) was reported in 2010 by Müssel and Kestler. BoolNet allows its users to reconstruct networks from time-series data, perform robustness and perturbation analysis and visualize the resultant cell fates attractor. BoolNet was employed by Steinway *et al.* (263) to construct a metabolic network towards evaluating gut microbiome in normal and disease conditions, whereas, Cohen *et al.* (264) studied tumor cell invasion and migration. BoolNet, however, lacks a graphical user interface, and results from the analysis cannot be visualized interactively, which hindered its employment. To address this issue, Shah *et al.* (265), in 2018, developed an Attractor Landscape Analysis Toolbox for Cell Fate Discovery and Reprogramming (ATLANTIS). ATLANTIS has an intuitive graphical user interface and interactive result visualization feature, for ease in utilization. The platform can be employed to perform deterministic as well as probabilistic analysis and was

validated through literature-based case studies on the yeast cell cycle (266), breast cancer (267), and colorectal cancer (268).

Differential Equations Modeling Software: Although Boolean models have proven to be a powerful tool in modeling complex biomolecular signaling networks, however, these models are unable to describe continuous concentration, and cannot be used to quantify the time-dependent behavior of biological systems, necessitating the need to switch to quantitative differential equations (269). As a result, numerous stand-alone and web-based tools have been developed to build continuous network models to help describe the temporal evolution of biomolecules towards elucidating more accurate cell fate outcomes from quantitative expression data (Supplementary Table 8). Amongst initial attempts at developing such software, GEPASI (GEneral PATHway Simulator) was reported by Pedro Mendes *et al.* (239), in 1993. GEPASI is a stand-alone simulator that facilitates formulating mathematical models of biochemical reaction networks. GEPASI can also be used to perform parametric sensitivity analysis using an automatic pipeline that evaluates networks in light of exhaustive combinatorial input parameters. Ricci *et al.* (270) employed GEPASI to investigate the mechanism of action of anticancer drugs, while Marín-Hernández *et al.* (271) constructed kinetic models of glycolysis in cancer. In 2006, Hoops *et al.* reported a successor of GEPASI; COPASI (COmplex PATHway Simulator) (272), a user-friendly independent biochemical simulator that can handles larger networks for faster simulation results, through parallel computing. Orton *et al.* (273) employed COPASI to model cancerous mutations in EGFR/ERK pathway, while cellular senescence was evaluated by Pezze *et al.* (274) for targeted therapeutic interventions. Towards establishing a user-friendly software with an intuitive graphical user interface (GUI), another desktop application, CellDesigner was published by Funahashi *et al.* (275). CellDesigner application can be extended to include various simulation and analysis packages through integration with systems biology workbench (SBW) (276). In a case study using CellDesigner, Calzone *et al.* (277) developed a network of retinoblastoma protein (RB/RB1) and evaluated its influence in the cell cycle, while Grieco *et al.* (278) investigated the impact of Mitogen-Activated Protein Kinase (MAPK) network on cancer cell fate outcomes.

Environmental Modeling Software

To facilitate the development of environmental models that can help investigate inter-, intra-, and extracellular interactions between cellular network models and their dynamic environment, several software and platforms have been reported (240,279). These platforms employ discrete, continuous, and hybrid approaches to develop models of cellular microenvironments towards setting up specific biological contexts such as normoxia, hypoxia, Warburg effect, etc (Supplementary Table 9). In 2014, Starrauß *et al.* (240) published Morpheus, a platform for modeling complex tumor microenvironment. Morpheus leverages a cellular pots modeling approach to integrate and stimulate cell-based biomolecular systems for modeling intra- and extra cellular dynamics. In a case study using Morpheus, Felix *et al.* (280) evaluated pancreatic ductal adenocarcinoma's adaptive and innate immune response levels, while Meyer *et al.* (281) investigated the dynamics of biliary fluid in the liver lobule. Although Morpheus is a widely employed modeling software, however, its diffusion solver is limited in its capacity to model large 3D domains. Towards modeling fast simulations for larger cellular systems, the Finite Volume Method for biological problems (BioFVM) software (279) was reported by Ghaffarizadeh *et al.*, in

2016. BioFVM is an efficient transport solver for single as well as multi-cell biological problems such as excretion, decomposition, diffusion, and consumption of substrates, etc (282). In a case study using BioFVM, Ozik *et al.* (283) evaluated tumor-immune interactions, while Wang *et al.* (284) elucidated the impact of tumor-parenchyma on the progression of liver metastasis. BioFVM, however, relies on its users to have programming-based knowledge to develop their models, which limits its translational potential. Towards minimizing programming requirement, SALSA (ScAffoLd SimulAtor) was reported by Cortesi *et al.* (218) in 2020. SALSA is general-purpose software that employs a minimum programming requirement, a significant advantage over its predecessors. The platform, in particular, has been useful in studying cellular diffusion in 3D cultures. This recent tool was validated in 2021, with a case study that evaluated and predicted therapeutic agents in 3D cell cultures (285).

Cell-level Modeling Software

Towards modeling cancer cell-specific behaviors such as cellular adhesion, membrane transport, loss of cell polarity, etc several software have been reported which can help develop *in silico* cancer cellular models (Supplementary Table 10). The foremost endeavor to develop software for cell-level modeling and simulation, led to the establishment of E-Cell (241), in 1999. E-Cell can be employed to model biochemical regulations and genetic processes using biomolecular regulatory networks in cells. In a case study using E-Cell, Edwards *et al.* (286) predicted the metabolic capabilities of *Escherichia coli*, and validate the results using existing literature, while Orton *et al.* (287) modeled the receptor-tyrosine-kinase-activated MAPK pathway. This software was expanded in 2001, with the development of Virtual Cell (V-Cell) (242,243,288), a web-based general-purpose modeling platform. V-Cell has an intuitive graphical and mathematical interface that allows ease in the design and simulation of whole cells, along with sub-cellular biomolecular networks and the external environment. Neves *et al.* (289) employed V-Cell to investigate the flow of spatial information in cAMP/PKA/B-Raf/MAPK1,2 networks, while Calebiro *et al.* (290) modeled cell signaling by internalized gprotein-coupled receptors. Similarly, to increase the efficiency in the design and modeling of synthetic regulatory networks in cells, in 2009, Chandran *et al.* reported TinkerCell (291) with computer-aided design (CAD) functionality which enabled faster simulation and associated analysis. The platform employed a modular approach for constructing networks and provides built-in features for ease in network construction, robustness analysis, and evaluating networks using existing databases. The evolutionary trend for TinkerCell's platform adaptability and flexibility is highlighted in Figure 4, which shows a gradual shift from being a model specific platform to a domain modelling one. In a study employing TinkerCell, Renicke *et al.* (292) constructed a generic photosensitive degron (psd) model to investigate protein degradation and cellular function, while Chandran *et al.* (293) reported computer-aided biological circuits.

Multi-scale Modeling Software

Multi-scale cancer modeling approaches bring together scale-specific models towards undertaking an integrative analysis of heterogeneous experimental data by building coherent and biologically plausible models (294–297). Several multi-scale modeling platforms have been reported to help develop multiscale cancer models (244–246) (Supplementary Table 11).

Amongst these, the REcursive Porous Agent Simulation Toolkit (Repast) (298,299) published in 2003, provides a free and open-source tool for modeling and simulating agent-based models, with high-performance computing (HPC) capability. Repast toolkit was employed in 2007, by Folcik *et al.* (300) to develop an agent-based model which was then used to study interactions between cells and the immune system. Similarly, Mehdizadeh *et al.* (301) used Repast to model angiogenesis in porous biomaterial scaffolds. Although Repast can be used for simulating several types of evolutionary trends between agents, there is no established guideline for selecting a mechanism to model such trends, hindering its use by naïve users. Moreover, Repast does not have a GUI or a software development kit (SDK) interface for implementing subcellular mechanisms e.g., gene, protein, and metabolic networks. In contrast, CompuCell (295), published in 2004 by Izaguirre *et al.*, provides an elaborative GUI to model cell-scale or tissue-scale simulations by integrating biomolecular networks, intra- and extracellular environment, and cell to environment interactions. Mahoney *et al.* (302) employed CompuCell to develop an angiogenesis-based model in cancer for investigating novel cancer therapies. This model was later optimized in 2010 (303) for evaluating potential new approaches for therapeutic intervention for angiogenesis. Although CompuCell provides an intuitive framework modeling paradigm, however, the platform core is not conducive to multi-scale cancer modeling. The focus of the software is primarily multi-agent simulations rather than multi-scale cancer modeling. This poses a significant challenge in the utilization of the software. In 2013, CHASTE (296) (Cancer Heart and Soft Tissue Environment) was launched, which provides a computational simulation pipeline for mathematical modeling of complex multi-scale models. Users can employ CHASTE for a wide range of problems involving on and off-lattice modeling workflows. CHASTE has also previously been employed to model colorectal cancer crypts. Nonetheless, CHASTE does not have a GUI and can only be executed by command line text commands. Furthermore, it requires recompilation on part of the modeler to use the code updates performed by the group. To further improve the multi-scale modeling approach for investigating cancer, in 2013, Chaudhary *et al.* (297) published ELECANS (Electronic Cancer System). ELECANS had an intuitive but rigid GUI along with a programmable SDK besides the lack of a high-performance simulation engine. ELECANS was employed to model the mitochondrial processes in cancer towards elucidating the hidden mechanisms involved in the Warburg effect (304). Although ELECANS provided a feature-rich environment for constructing multi-scale models, however, the platform lacked a biomolecular network integration pipeline, and it also placed a heavy programming requirement on its users. In contrast, in 2018, PhysiCell (305) was reported for 2D and 3D multi-cell off-lattice agent-based simulations. PhysiCell is coupled with BioFVM (279)'s finite volume method to model multi-scale cancer systems (283). In a salient example, Wang *et al.* (284) employed PhysiCell to model liver metastatic progression. In 2019, PhysiCell's agent-based modeling features and MaBoss's Boolean cell signaling network feature were coupled together to develop an integrated platform, PhysiBoss (306). As a result, PhysiBoss provided an agent-based modeling environment to study physical dimension and cell signaling networks in a cancer model. In 2020, Colin *et al.* (307) employed PhysiBoss's source code to model diffusion in oocytes during prophase 1 and meiosis 1, while Getz *et al.* (308) proposed a framework using PhysiBoss to develop a multi-scale model

of SARS-CoV-2 dynamics in lung tissue. Recently, in 2021, Gondal *et al.* (309) reported Theatre for *in silico* Systems Oncology (TISON), a web-based multi-scale “zero-code” modeling and simulation platform for *in silico* oncology. TISON aims to develop single or multi-scale models for developing personalized cancer therapeutics. To exemplify the use case for TISON, Gondal *et al.* employed TISON to model colorectal tumorigenesis in *Drosophila melanogaster*’s midgut towards evaluating efficacious combinatorial therapies for individual colorectal cancer patients (236).

Summarily, multi-scale modeling software have enabled the development of biologically plausible cancer models to varying degrees. These platforms, however, fall short of providing a generic and high-throughput environment that could be conveniently translated in clinical settings.

Scales		Network					Environment			Cell Line			Multi-scale					
Software		CellDesigner	COPASI	CellNetAnalyzer	BoolNet	GIMSim	CellWare	Morpheus	BioFVM	E-Cell	Virtual Cell	ThinkCell	Compucell	CHASTE	ELECANS	PhysiCell	PhysiBoSS	TISON
Year		2003	2006	2008	2010	2012	2003	2014	2015	1999	2008	2009	2004	2013	2013	2018	2019	2021
Features List	Boolean/Qualitative Networks			•	•	•											•	•
	Boolean/Quantitative Networks				•													•
	Quantitative Networks	•	•				•	•		•	•							
	Stochastic Semantics		•		•					•	•		•				•	•
	ODE		•				•	•			•	•	•	•				•
	Agent-based Modeling												•	•	•	•	•	
	2D Tissue							•		•	•		•	•	•	•	•	•
	3D Tissue								•	•	•	•	•	•	•	•	•	•
	Biomolecular Diffusion Continuous									•	•		•	•				•
	Biomolecular Diffusion Discrete							•	•	•	•			•	•	•	•	•
	Biomolecular Diffusion Hybrid						•				•							•
	Parallelized		•	•	•			•	•		•		•				•	•
	GUI	•	•	•		•	•	•	•	•	•	•	•		•		•	•
	Web-based Platform	•	•			•		•		•	•	•						•
	Language	Java	C++	MATLAB	R	Java	Java	C++	C++	C++	C++/Java	Python/C	C++	C#	C++	C++	C++	C#
	Reference PMID's	doi:10.1016/S1478-5382(03)002370-9	17032683	17408509	20378558	22144167	14871872	24443380	26656933	10068694	19045830	19874625	14764549	23516352	23390137	29474446	30169736	doi:10.1101/2021.05.04.442539

Figure 3. Feature-by-feature comparison of networks, environments, cell lines, and multi-scale modeling software in chronological order.

Figure 4. Evolution of scale-specific and multi-scale software. Evolution of multiscale modeling software for abstracting and simulating the spatiotemporal biomolecular complexity. Highlighting the need for a generic, data-driven, zero-code software requirement.

Pipelining Panomics Data towards *in silico* Clinical Systems Oncology

In silico multi-scale cancer models, annotated with patient-specific biomolecular and clinical data, can help decode complex mechanisms underpinning tumorigenesis and assist in clinical decision-making (310,311). Clinically-driven *in silico* multi-scale cancer models simulate *in vivo* tumor growth and response to therapies across biocomplexity scales, within a clinical environment, towards evaluating efficacious treatment combinations. To facilitate the development of multi-scale cancer models, several large-scale program projects have been launched (Supplementary Table 12) such as Advancing Clinico-Genomic Trials on Cancer (ACGT) (312), Clinically Oriented Translational Cancer Multilevel Modeling (ContraCancrum) (313), Personalized Medicine (p-medicine) (314), Transatlantic Tumor Model Repositories (TUMOR) (315), and Computational Horizons In Cancer (CHIC) (316), amongst others (Figure 5). Here, we review and evaluate five salient projects for multi-scale cancer modeling towards their clinical deployment.

ACGT Project

The ACGT project (312), launched in 2007, proposed to develop Clinico-Genomic infrastructure for organizing clinical and genomic data towards investigating personalized therapeutics regimens for an individual cancer patient (317). The ACGT platform provides an open-source and open-access infrastructure designed to support the development of “oncosimulators” to help clinicians accurately compare results from different clinical trials and enhance their efficiency towards optimizing cancer treatment (318,319). The ACGT framework employs molecular and clinical data generated from different sources including whole genome sequencing, histopathological, imaging, molecular, and clinical data, etc. to develop simulators for mimicking clinical trials. Personalized panomics data employed to develop oncosimulators in ACGT is extracted from real patients which enables the oncosimulators to be clinically relevant for predictive purposes. Additionally, ACGT provides data retrieval, storage, integrative, anonymization, and analysis as well as results presentation capabilities. Using the platform, in 2008, Kolokotroni *et al.* (320) developed a personalized, spatiotemporal oncosimulator model of breast cancer to mimic a clinical trial based on protocols outlined in the Trail of Principle (TOP), towards evaluating the model response to chemotherapeutic treatment in neoadjuvant settings (321). Similarly, in 2009, Graf *et al.* (322,323) modeled nephroblastoma oncosimulator, a childhood cancer of the kidney, based on a clinical trial run by the International Society of Paediatric Oncology (SIOP) for simulating tumor response to therapeutic regimens in the clinical trials. The results generated from the TOP and SIOP trials enabled the ACGT oncosimulators to adapt in light of real clinical conditions and the software to be validated against multi-scale patient data. The focus of ACGT oncosimulators, however, is limited to existing clinical trials for predicting efficacious treatment combinations.

The ContraCancrum Project

Towards establishing a platform for the development of composite multi-scale models for simulating malignant tumor models, in 2008, ContraCancrum project (313) was initiated. ContraCancrum aimed to develop a multi-scale computational framework for translating personalized cancer models into clinical settings towards simulating malignant tumor

development and response to therapeutic regimens. For that, the Individualized Medicine Simulation Environment (IMENSE) platform (324,325) was established, to undertake the oncosimulator development process. The platform was employed for molecular and clinical data storage, retrieval, integration, and analysis. Oncosimulators, developed under the ContraCancrum project, employ patient data across biologically and clinically relevant scales including molecular, environmental, cells, and tissues level. These oncosimulators can be used to optimize personalized cancer therapeutics for assisting clinician decision-making process. To date, several oncosimulators have been developed, under the ContraCancrum project (31,313,323,325–332). In particular, the initial validation of the ContraCancrum workflow was performed using two case studies; glioblastoma multiforme (GBM) (313,326,332) and non-small cell lung cancer (NSCLC) (313,330,331). In 2010, Folarin and Stamatakis (333) developed a glioblastoma oncosimulator using personalized molecular patient data to evaluate treatment response under the effect of a chemotherapeutic drug (temozolomide) (324). Similarly, Roniotis *et al.* (332) developed a multi-scale finite elements workflow to model glioblastoma growth, while Giatili *et al.* (326) outlined explicit boundary condition treatment in glioblastoma using an *in silico* tumor growth model. The results from the case studies were validated by comparing *in silico* prediction with pre- and post-operative imaging and clinical data (313,328). Moreover, The ContraCancrum project hosts more than 100 lung patients' tumor and blood samples (313). This data is employed to develop clinically validated *in silico* multilevel cancer models for NSCLC using patient-specific data (330). In 2010, using a biochemical oncosimulator, Wan *et al.* (330) investigated the binding affinities for AEE788 and Gefitinib tyrosine kinase inhibitor against mutated epidermal growth factor receptor (EGFR) for NSCLC treatment. Similarly, Wang *et al.* (331) developed a 2D agent-based NSCLC model to investigate proliferation markers and evaluated ERK as a suitable target for targeted therapy. Although ContraCancrum's project has created numerous avenues for multilevel cancer model development, however, the platform is limited to only specific types of cancer for which data is internally available on the platform.

p-medicine Project

Towards improving clinical deployment capabilities of oncosimulators, another pilot project: the personalized medicine (p-medicine) project was launched in 2011 (314). The main aim of p-medicine was to create biomedical tools facilitating translation of current medicine to P5 medicine (predictive, personalized, preventive, participatory, and psycho-cognitive) (334). For that, the p-medicine portal provides a web-based environment that hosts specifically-purpose tools for personalized panomics data integration, management, and model development. The portal has an intuitive graphical user interface (GUI) with an integrated workbench application for integrating information from clinical practices, histopathological imaging, treatment and omics data, etc. Computational models developed under p-medicine workflow, are quantitatively adapted to clinical settings since they are derived using real multi-scale data. Several multi-scale cancer simulation models (oncosimulators) (314,334) have been devised, using the p-medicine workflow. Amongst these, in 2012, Georgiadi *et al.* (335) developed a four-dimensional nephroblastoma treatment model and evaluated its employment in clinical decisions making. Towards evaluating personalized therapeutic combinations, in 2014, Blazewicz *et al.* (336) developed a p-medicine parallelized oncosimulator which evaluated nephroblastoma tumor response to therapy. The parallelization enhanced model usability and accuracy for eventual

translation into clinical settings for supporting clinical decisions. In 2016, Argyri *et al.* (337) developed a breast cancer oncosimulator to evaluate vascular tumor growth in light of single-agent bevacizumab therapy (anti-angiogenic treatment), while in 2014, Stamatakis *et al.* (338) evaluated breast cancer treatment under an anthracycline drug for chemotherapy (epirubicin). In 2012, Ouzounoglou *et al.* (339) developed a personalized acute lymphoblastic leukaemia (ALL) oncosimulator for evaluating the efficacy of prednisolone (a steroid medication). This study was further augmented, in 2015, by Kolokotroni *et al.* (340) to investigate the potential cytotoxic side effects of prednisolone. Later in 2017, Ouzounoglou *et al.* (341) expanded their ALL oncosimulator model to a hybrid oncosimulator for predicting pre-phase treatments for ALL patients. The validation of these models was undertaken using clinical trial data in pre and post-treatment (342). Although, p-medicine presents a state-of-the-art *in silico* multi-scale cancer modeling environment, however, the project is limited in its application for determining individual biomarkers for potential novel therapy identification. Moreover, the project is limited to its niche set of tools and models that cannot be integrated with other similar model repositories and platforms.

TUMOR Project

In 2012, a transatlantic USA-EU partnership was initiated with the launch of Transatlantic Tumor Model Repositories (TUMOR) (315). TUMOR provides an integrated, interoperable transatlantic research environment for developing a clinically-driven cancer model repository. The repository aimed to integrate computational cancer models developed by different research groups. TUMOR's transatlantic aim was to couple models from ACGT and ContraCancrum projects with those at the Center for the Development of a Virtual Tumor (CViT) (343) as well as other relevant centers. TUMOR is predicted to serve as an international clinical translation, interoperable and validation platform for *in silico* oncology hosting multi-scale cancer models from different cancer model repositories and platforms such as E-Cell (344), CellML (345), FieldML (346), and BioModels (39,347). To support model data interoperability between platforms, TumorML (Tumor model repositories Markup Language) (347) was developed towards facilitating inter-data operability in the TUMOR project. The TUMOR environment also offers a wide range of additional services supporting predictive oncology and individualized optimization of cancer treatment. For example, the platform allows remote access of predictive cancer models in hospitals and to clinicians for the development of quantitative cancer research and personalized cancer therapy model. TUMOR environment also incorporates deterministic as well as stochastic models through COPASI simulator (246). An automatic validation pipeline is also embedded for the execution and deployment of these models in clinical settings (315). TUMOR's ability to couple and integrate models from different scales, approaches, and repositories towards increasing model accuracy in predictive oncology was exemplified with Wang *et al.*'s (348) NSCLC model. In 2007, Wang *et al.* developed a 2D multi-scale NSCLC model for evaluating tumor expansion dynamics in NSCLC patients, in CViT. This model was exported in TumorML and made available in the TUMOR repository to help evaluate growth factor influence in aggressive cancer (349). Similarly, in 2014 Sakalis *et al.* (350) employed the TUMOR platform to interlink and couple three independent glioblastoma-specific cancer models (EGFR signaling (351), cancer metabolism (352), Oncosimulator (350,353)), reported by different research groups. The

resultant model was used to investigate the impact of radiation and temozolomide (chemotherapy) on glioblastoma multiforme to evaluate treatment effectiveness.

CHIC Project

Another transatlantic project Computational Horizons In Cancer (CHIC) project (316) launched in 2013, aimed to provide an oncosimulator modeling platform for *in silico* oncology. CHIC was initiated to develop and implement predictive oncology and individualized multi-scale cancer modeling tools towards assisting quantitative cancer research and personalized cancer therapy. The workflow and tools established under the ambient of the CHIC project allowed the development of robust, interoperable, and collaborative *in silico* models in cancer and normal conditions (316). The CHIC project also proposed a pipeline to translate the model towards supporting clinicians to make optimal personalized treatment plans for individual patients. To this end, several models are created using CHIC such as non-small cell lung cancer (354), glioblastoma (355), leukemia model (356), etc. These models furnish a quantitative understanding of tumorigenesis towards providing avenues for promoting individual cancer patient treatment combinations. One such model reported by Kolokotroni *et al.* (354) evaluated the efficacy of cisplatin-based therapy for NSCLC patients using *in silico* multi-scale cancer model. While, in 2015, Antonopoulos and Stamatakis (355) modeled the infiltration of glioblastoma cells in normal brain regions using a novel treatment. In 2017, Stamatakis and Giatili (357) extended glioblastoma oncosimulator for modeling tumor growth using reaction-diffusion numerical handling based on multi-scale panomics data. They further proposed a clinical pipeline to translate the model into clinical settings towards supporting clinicians to make optimal personalized treatment plans for individual patients. Ouzounoglou *et al.* (356) developed an *in silico* multi-scale leukemia oncosimulator model towards modeling deregulations in the G1/S pathway to investigate the altered function of retinoblastoma in ALL patients. However, a clinical translation of these models is currently in the works (358).

Figure 5. Salient projects pipelining multi-scale panomics data into clinical settings – a timeline. Timeline highlighting salient project platforms for developing realistic and clinically-driven multi-scale cancer models, along with their associated leading case studies.

Future Directions and Conclusion

In this work, we have evaluated the use of data-driven multi-scale cancer models in deciphering complex biomolecular underpinnings in cancer research towards developing personalized therapeutics interventions for clinical decision-making. Specifically, we have discussed the chronological evolution of online cancer data repositories that host high-resolution datasets from multiple spatiotemporal scales. Next, we evaluate how this data drives single- and multi-scale systems biology models towards decoding complex cancer regulation in patients. We then track the development of various modeling software and their applications in enhancing the translational role of cancer systems biology models in clinics. We conclude that the contemporary multi-scale modeling software line-up remains limited in their clinical employment due to the lack of a generic, zero-code, panomic-based framework for translating research models into clinical settings. Such a framework would help annotate *in silico* cancer models developed using single and multi-scale databases (69,359–365). The framework should also provide an environment for developing the extra-cellular matrix of a cancer cell which can then be integrated into cellular models. Existing environment (42,43) and cell line databases (45–47) need to be integrated to design environmental models along with biologically plausible cell line structures. These cell lines could then be assembled into three-dimensional geometries to create multi-scale *in silico* organoids. The pipeline should also furnish capabilities such as a convenient import workflow for clinical data integration through histopathological image data repositories (48–50) for designing biologically accurate organoid structures based on each cancer patient's underlying cellular morphology. Once the personalized multi-scale model has been constructed, the pipeline would allow investigation into the temporal evolution of the multi-scale organoid under personalized inputs and user-designed biomolecular entities. Data generated from the multi-scale model simulation can be analyzed to elicit biomolecular cues for each cancer patient as well as determine its role.

Taken together, a translational *in silico* systems oncology pipeline is the need of the hour and will help develop and deliver personalized treatments of cancer as well as substantively inform clinical decision-making processes.

Conflict of Interest

The authors declare that they have no conflict of interest.

Funding

This work was supported by the National ICT-R&D Fund (SRG-209), RF-NCBC-015, NGIRI-2020-4771, HEC (21-30SRGP/R&D/HEC/ 2014, 20-2269/NRPU/R&D/ HEC/12/ 4792 and 20-3629/NRPU/R&D/HEC/14/ 585), TWAS (RG 14-319 RG/ITC/AS_C) and LUMS (STG-BIO-1008, FIF-BIO-2052, FIF-BIO-0255, SRP-185-BIO, SRP-058-BIO and FIF-477-1819-BIO) grants.

Data Availability Statement

The supplementary tables for the manuscript titled “*Navigating Multi-scale Cancer Systems Biology towards Model-driven Personalized Therapeutics*” is available online.

References

1. National Cancer Act of 1971 - National Cancer Institute.
2. Brennan R, Federico S, Dyer MA. The war on cancer: have we won the battle but lost the war? *Oncotarget*. 2010;1(2):77.
3. DeVita VT. The “War on Cancer” and its impact. *Nat Clin Pract Oncol* [Internet]. 2004;1(2):55. Available from: <https://doi.org/10.1038/ncponc0036>
4. National Cancer Act of 1971 [Internet]. Available from: <https://www.cancer.gov/about-nci/overview/history/national-cancer-act-1971>
5. Wong AH-H, Deng C-X. Precision Medicine for Personalized Cancer Therapy. Vol. 11, *International journal of biological sciences*. 2015. p. 1410–2.
6. Giacomini KM, Yee SW, Ratain MJ, Weinshilboum RM, Kamatani N, Nakamura Y. Pharmacogenomics and patient care: one size does not fit all. *Sci Transl Med*. 2012 Sep;4(153):153ps18.
7. Krzyszczyk P, Acevedo A, Davidoff EJ, Timmins LM, Marrero-Berrios I, Patel M, et al. The growing role of precision and personalized medicine for cancer treatment. *TECHNOLOGY*. 2018 Sep;06(03n04):79–100.
8. Binotto M, Reinert T, Werutsky G, Zaffaroni F, Schwartzmann G. Health-related quality of life before and during chemotherapy in patients with early-stage breast cancer. *Ecancermedicalscience*. 2020;14:1007.
9. Adler AS, Lin M, Horlings H, Nuyten DSA, van de Vijver MJ, Chang HY. Genetic regulators of large-scale transcriptional signatures in cancer. *Nat Genet*. 2006 Apr;38(4):421–30.
10. Shi Y, Wei J, Chen Z, Yuan Y, Li X, Zhang Y, et al. Integrative Analysis Reveals Comprehensive Altered Metabolic Genes Linking with Tumor Epigenetics Modification in Pan-Cancer. Nakayasu ES, editor. *Biomed Res Int* [Internet]. 2019;2019:6706354. Available from: <https://doi.org/10.1155/2019/6706354>
11. Chen Y, Ni J, Gao Y, Zhang J, Liu X, Chen Y, et al. Integrated proteomics and metabolomics reveals the comprehensive characterization of antitumor mechanism underlying Shikonin on colon cancer patient-derived xenograft model. *Sci Rep* [Internet]. 2020;10(1):14092. Available from: <https://doi.org/10.1038/s41598-020-71116-5>
12. Angermueller C, Clark SJ, Lee HJ, Macaulay IC, Teng MJ, Hu TX, et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods* [Internet].

- 2016;13(3):229–32. Available from: <https://doi.org/10.1038/nmeth.3728>
13. Yoshida A, Kohyama S, Fujiwara K, Nishikawa S, Doi N. Regulation of spatiotemporal patterning in artificial cells by a defined protein expression system. *Chem Sci*. 2019;11064–72.
14. Rejniak KA, Anderson ARA. Hybrid models of tumor growth. *Wiley Interdiscip Rev Syst Biol Med*. 2011;3(1):115–25.
15. Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, et al. Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* [Internet]. 2018;173(2):321–337.e10. Available from: <http://www.sciencedirect.com/science/article/pii/S0092867418303593>
16. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med*. 2017 Jun;23(6):703–13.
17. Al Hajri Q, Dash S, Feng W, Garner HR, Anandkrishnan R. Identifying multi-hit carcinogenic gene combinations: Scaling up a weighted set cover algorithm using compressed binary matrix representation on a GPU. *Sci Rep* [Internet]. 2020;10(1):2022. Available from: <https://doi.org/10.1038/s41598-020-58785-y>
18. Roerink SF, Sasaki N, Lee-Six H, Young MD, Alexandrov LB, Behjati S, et al. Intra-tumour diversification in colorectal cancer at the single-cell level. *Nature*. 2018 Apr;556(7702):457–62.
19. National Cancer Institute [Internet]. Available from: <http://www.cancer.gov/types/by-body-location>
20. Song Q, Merajver SD, Li JZ. Cancer classification in the genomic era: five contemporary problems. *Hum Genomics*. 2015 Oct;9:27.
21. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000;100(1):57–70.
22. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646–74.
23. Saha R, Chowdhury A, Maranas CD. Recent advances in the reconstruction of metabolic models and integration of omics data. *Curr Opin Biotechnol* [Internet]. 2014;29(1):39–45. Available from: <http://dx.doi.org/10.1016/j.copbio.2014.02.011>
24. Deisboeck TS, Berens ME, Kansal AR, Torquato S, Stemmer-Rachamimov AO, Chiocca EA. Pattern of self-organization in tumour systems: Complex growth dynamics in a novel brain tumour spheroid model. *Cell Prolif*. 2001;34(2):115–34.
25. Lorenzo G, Scott MA, Tew K, Hughes TJR, Zhang YJ, Liu L, et al. Tissue-scale, personalized modeling and simulation of prostate cancer growth. *Proc Natl Acad Sci* [Internet]. 2016 Nov 29;113(48):E7663 LP-E7671. Available from: <http://www.pnas.org/content/113/48/E7663.abstract>

26. Pingle SC, Sultana Z, Pastorino S, Jiang P, Mukthavaram R, Chao Y, et al. In silico modeling predicts drug sensitivity of patient-derived cancer cells. *J Transl Med* [Internet]. 2014;12(1):128. Available from: <https://doi.org/10.1186/1479-5876-12-128>
27. Ridley M, Matthews P. *Genome*. WF Howes; 2000.
28. Velculescu VE, Zhang L, Zhou W, Vogelstein J, Basrai MA, Bassett DEJ, et al. Characterization of the yeast transcriptome. *Cell*. 1997 Jan;88(2):243–51.
29. Anderson NL, Anderson NG. Proteome and proteomics: new technologies, new concepts, and new words. *Electrophoresis*. 1998 Aug;19(11):1853–61.
30. Manzoni C, Kia DA, Vandrovcova J, Hardy J, Wood NW, Lewis PA, et al. Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences. *Brief Bioinform* [Internet]. 2016;19(2):286–302. Available from: <https://doi.org/10.1093/bib/bbw114>
31. Stamatakos GS, Dionysiou DD, Graf NM, Sofra NA, Desmedt C, Hoppe A, et al. The “Oncosimulator”: a multilevel, clinically oriented simulation system of tumor growth and organism response to therapeutic schemes. Towards the clinical evaluation of in silico oncology. *Conf Proc IEEE Eng Med Biol Soc*. 2007;6629–32.
32. Stephens ZD, Lee SY, Faghri F, Campbell RH, Zhai C, Efron MJ, et al. Big Data: Astronomical or Genomical? *PLoS Biol*. 2015 Jul;13(7):e1002195.
33. Krupp M, Marquardt JU, Sahin U, Galle PR, Castle J, Teufel A. RNA-Seq Atlas--a reference database for gene expression profiling in normal tissue by next-generation sequencing. *Bioinformatics*. 2012 Apr;28(8):1184–5.
34. Park S-J, Yoon B-H, Kim S-K, Kim S-Y. GENT2: an updated gene expression database for normal and tumor tissues. *BMC Med Genomics*. 2019;12(5):101.
35. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*. 2002 Jan;30(1):207–10.
36. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*. 1999 Jan;27(1):29–34.
37. Nikitin A, Egorov S, Daraselia N, Mazo I. Pathway studio--the analysis and navigation of molecular networks. *Bioinformatics*. 2003 Nov;19(16):2155–7.
38. Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, et al. PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res*. 2003 Sep;13(9):2129–41.
39. Le Novère N, Bornstein B, Broicher A, Courtot M, Donizelli M, Dharuri H, et al. BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Res*. 2006 Jan;34(Database issue):D689-91.

40. Hermjakob H, Montecchi-Palazzi L, Lewington C, Mudali S, Kerrien S, Orchard S, et al. IntAct: an open source molecular interaction database. *Nucleic Acids Res.* 2004 Jan;32(Database issue):D452-5.
41. Zanzoni A, Montecchi-Palazzi L, Quondam M, Ausiello G, Helmer-Citterich M, Cesareni G. MINT: a Molecular INTERaction database. *FEBS Lett.* 2002 Feb;513(1):135-40.
42. Chautard E, Fatoux-Ardore M, Ballut L, Thierry-Mieg N, Ricard-Blum S. MatrixDB, the extracellular matrix interaction database. *Nucleic Acids Res.* 2011 Jan;39(Database issue):D235-40.
43. Shao X, Taha IN, Clauser KR, Gao YT, Naba A. MatrisomeDB: the ECM-protein knowledge database. *Nucleic Acids Res.* 2020 Jan;48(D1):D1136-44.
44. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* 2019 Jan;47(D1):D941-7.
45. Rajapakse VN, Luna A, Yamade M, Loman L, Varma S, Sunshine M, et al. CellMinerCDB for Integrative Cross-Database Genomics and Pharmacogenomics Analyses of Cancer Cell Lines. *iScience.* 2018 Dec;10:247-64.
46. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature.* 2012;483(7391):603-307.
47. Romano P, Manniello A, Aresu O, Armento M, Cesaro M, Parodi B. Cell Line Data Base: structure and recent improvements towards molecular authentication of human cell lines. *Nucleic Acids Res.* 2009 Jan;37(Database issue):D925-32.
48. Prior FW, Clark K, Commean P, Freymann J, Jaffe C, Kirby J, et al. TCIA: An information resource to enable open science. *Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Int Conf.* 2013;2013:1282-5.
49. Gutman DA, Cobb J, Somanna D, Park Y, Wang F, Kurc T, et al. Cancer Digital Slide Archive: an informatics resource to support integrated in silico analysis of TCGA pathology data. *J Am Med Inform Assoc.* 2013;20(6):1091-8.
50. Carithers LJ, Moore HM. The Genotype-Tissue Expression (GTEx) Project. *Biopreserv Biobank.* 2015;13(5):307-8.
51. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res.* 2010 Jan;38(Database issue):D355-60.
52. Hamosh A, Scott AF, Amberger J, Bocchini C, Valle D, McKusick VA. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res.* 2002 Jan;30(1):52-5.
53. Chakravarty D, Gao J, Phillips SM, Kundra R, Zhang H, Wang J, et al. OncoKB: A Precision Oncology Knowledge Base. *JCO Precis Oncol.* 2017 Jul;2017.

54. Wishart DS, Knox C, Guo AC, Cheng D, Shrivastava S, Tzur D, et al. DrugBank: A knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.* 2008;
55. Smirnov P, Kofia V, Maru A, Freeman M, Ho C, El-Hachem N, et al. PharmacoDB: an integrative database for mining in vitro anticancer drug screening studies. *Nucleic Acids Res.* 2018;46(D1):D994–1002.
56. Piñeiro-Yáñez E, Reboiro-Jato M, Gómez-López G, Perales-Patón J, Troulé K, Rodríguez JM, et al. PanDrugs: a novel method to prioritize anticancer drug treatments according to individual genomic data. *Genome Med.* 2018;10(1):1–11.
57. Pinu FR, Beale DJ, Paten AM, Kouremenos K, Swarup S, Schirra HJ, et al. Systems biology and multi-omics integration: Viewpoints from the metabolomics research community. *Metabolites.* 2019;9(4):1–31.
58. Bilofsky HS, Burks C, Fickett JW, Goad WB, Wayne FIL, Rindone P, et al. The GenBank genetic sequence databank. 1986;14(1).
59. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, et al. GenBank. *Nucleic Acids Res [Internet].* 2012;41(D1):D36–42. Available from: <https://doi.org/10.1093/nar/gks1195>
60. Díez O, Osorio A, Durán M, Martínez-Ferrandis JI, De la Hoya M, Salazar R, et al. Analysis of BRCA1 and BRCA2 genes in Spanish breast/ovarian cancer patients: A high proportion of mutations unique to Spain and evidence of founder effects. *Hum Mutat.* 2003;22(4):301–12.
61. Medrek C, Pontén F, Jirström K, Leandersson K. The presence of tumor associated macrophages in tumor stroma as a prognostic marker for breast cancer patients. *BMC Cancer [Internet].* 2012;12(1):306. Available from: <https://doi.org/10.1186/1471-2407-12-306>
62. Hubbard T, Barker D, Birney E, Cameron G, Chen Y, Clark L, et al. The Ensembl genome database project. *Nucleic Acids Res.* 2002 Jan;30(1):38–41.
63. Easton DF, Pooley KA, Dunning AM, Pharoah PDP, Thompson D, Ballinger DG, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature [Internet].* 2007;447(7148):1087–93. Available from: <https://doi.org/10.1038/nature05887>
64. Phillips T. Regulation of transcription and gene expression in eukaryotes. *Nat Educ.* 2008;1(1):199.
65. Sequenc NG, Pareek CS. Next Generation : Sequencing & Applications Transcriptome Analysis on RNA-seq Data. 2015;1–2.
66. Zhao D, Tang Y, Xia X, Sun J, Meng J, Shang J, et al. Integration of Transcriptome, Proteome, and Metabolome Provides Insights into How Calcium Enhances the Mechanical Strength of Herbaceous Peony Inflorescence Stems. *Cells.* 2019 Jan;8(2).

67. Clough E, Barrett T. The Gene Expression Omnibus Database. *Methods Mol Biol*. 2016;1418:93–110.
68. Chakraborty PK, Mustafi SB, Xiong X, Dwivedi SKD, Nesin V, Saha S, et al. MICU1 drives glycolysis and chemoresistance in ovarian cancer. *Nat Commun* [Internet]. 2017;8(1):14634. Available from: <https://doi.org/10.1038/ncomms14634>
69. Bamford S, Dawson E, Forbes S, Clements J, Pettett R, Dogan A, et al. The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br J Cancer*. 2004;91:355–8.
70. Schubbert S, Shannon K, Bollag G. Hyperactive Ras in developmental disorders and cancer. *Nat Rev Cancer* [Internet]. 2007 Apr 1;7:295. Available from: <https://doi.org/10.1038/nrc2109>
71. Weir BA, Woo MS, Getz G, Perner S, Ding L, Beroukheim R, et al. Characterizing the cancer genome in lung adenocarcinoma. *Nature* [Internet]. 2007;450(7171):893–8. Available from: <https://doi.org/10.1038/nature06358>
72. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. Vol. 1A, *Współczesna Onkologia*. 2015. p. A68–77.
73. Tamborero D, Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Kandoth C, Reimand J, et al. Comprehensive identification of mutational cancer driver genes across 12 tumor types. *Sci Rep*. 2013;3(1):1–10.
74. Leiserson MDM, Vandin F, Wu H-T, Dobson JR, Eldridge J V, Thomas JL, et al. Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nat Genet* [Internet]. 2015;47(2):106–14. Available from: <https://doi.org/10.1038/ng.3168>
75. Davis CF, Ricketts CJ, Wang M, Yang L, Cherniack AD, Shen H, et al. The Somatic Genomic Landscape of Chromophobe Renal Cell Carcinoma. *Cancer Cell* [Internet]. 2014;26(3):319–30. Available from: <http://www.sciencedirect.com/science/article/pii/S1535610814003043>
76. Zhang J, Baran J, Cros A, Guberman JM, Haider S, Hsu J, et al. International Cancer Genome Consortium Data Portal—a one-stop shop for cancer genomics data. *Database*. 2011;2011.
77. Jones DTW, Jäger N, Kool M, Zichner T, Hutter B, Sultan M, et al. Dissecting the genomic complexity underlying medulloblastoma. *Nature*. 2012 Aug;488(7409):100–5.
78. Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, Leonard B, et al. APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* [Internet]. 2013;494(7437):366–70. Available from: <https://doi.org/10.1038/nature11881>
79. Supek F, Miñana B, Valcárcel J, Gabaldón T, Lehner B. Synonymous Mutations Frequently Act as Driver Mutations in Human Cancers. *Cell* [Internet]. 2014;156(6):1324–35.

Available from: <http://www.sciencedirect.com/science/article/pii/S0092867414001457>

80. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2012 May;2(5):401–4.
81. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal.* 2013 Apr;6(269):pl1.
82. Robinson D, Allen EM Van, Wu Y-M, Schultz N, Lonigro RJ, Mosquera J-M, et al. Integrative Clinical Genomics of Advanced Prostate Cancer. *Cell* [Internet]. 2015;161(5):1215–28. Available from: <http://www.sciencedirect.com/science/article/pii/S0092867415005486>
83. Zaretsky JM, Garcia-Diaz A, Shin DS, Escuin-Ordinas H, Hugo W, Hu-Lieskovan S, et al. Mutations associated with acquired resistance to PD-1 blockade in melanoma. *N Engl J Med.* 2016;375(9):819–29.
84. Bailey P, Chang DK, Nones K, Johns AL, Patch A-M, Gingras M-C, et al. Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* [Internet]. 2016;531(7592):47–52. Available from: <https://doi.org/10.1038/nature16965>
85. Jiao X-D, Qin B-D, You P, Cai J, Zang Y-S. The prognostic value of TP53 and its correlation with EGFR mutation in advanced non-small cell lung cancer, an analysis based on cBioPortal data base. *Lung Cancer* [Internet]. 2018;123:70–5. Available from: <https://www.sciencedirect.com/science/article/pii/S0169500218304628>
86. Hou H, Zhang C, Qi X, Zhou L, Liu D, Lv H, et al. Distinctive targetable genotypes of younger patients with lung adenocarcinoma: a cBioPortal for cancer genomics data base analysis. *Cancer Biol Ther.* 2020;21(1):26–33.
87. Chen G, Gharib TG, Huang C-C, Taylor JMG, Misek DE, Kardia SLR, et al. Discordant protein and mRNA expression in lung adenocarcinomas. *Mol Cell Proteomics.* 2002 Apr;1(4):304–13.
88. Cox J, Mann M. Is proteomics the new genomics? *Cell.* 2007;130(3):395–8.
89. Aslam B, Basit M, Nisar MA, Khurshid M, Rasool MH. Proteomics: Technologies and Their Applications. *J Chromatogr Sci* [Internet]. 2017 Feb 1;55(2):182–96. Available from: <https://doi.org/10.1093/chromsci/bmw167>
90. Williams CW, Elmendorf HG. Identification and analysis of the RNA degrading complexes and machinery of *Giardia lamblia* using an in silico approach. *BMC Genomics* [Internet]. 2011;12(1):586. Available from: <https://doi.org/10.1186/1471-2164-12-586>
91. Gill G. Post-translational modification by the small ubiquitin-related modifier SUMO has big effects on transcription factor activity. *Curr Opin Genet Dev.* 2003;13(2):108–13.
92. Bathke J, Konzer A, Remes B, McIntosh M, Klug G. Comparative analyses of the variation

- of the transcriptome and proteome of *Rhodobacter sphaeroides* throughout growth. BMC Genomics [Internet]. 2019;20(1):358. Available from: <https://doi.org/10.1186/s12864-019-5749-3>
93. Marx V. Proteomics: An atlas of expression. Nature. 2014;509(7502):645–9.
94. Persson A, Hober S, Uhlén M. A human protein atlas based on antibody proteomics. Curr Opin Mol Ther. 2006;8(3):185–90.
95. Uhlén M, Björling E, Agaton C, Szigartyo CAK, Amini B, Andersen E, et al. A human protein atlas for normal and cancer tissues based on antibody proteomics. Mol Cell Proteomics. 2005;4(12):1920–32.
96. Gámez-Pozo A, Sánchez-Navarro I, Nistal M, Calvo E, Madero R, Díaz E, et al. MALDI Profiling of Human Lung Cancer Subtypes. PLoS One [Internet]. 2009 Nov 5;4(11):e7731. Available from: <https://doi.org/10.1371/journal.pone.0007731>
97. Imberg-Kazdan K, Ha S, Greenfield A, Poultney CS, Bonneau R, Logan SK, et al. A genome-wide RNA interference screen identifies new regulators of androgen receptor function in prostate cancer cells. Genome Res. 2013;23(4):581–91.
98. Collins FS. New Goals for the U.S. Human Genome Project: 1998-2003. Science (80-). 1998;282(5389):682–9.
99. Legrain P, Aebersold R, Archakov A, Bairoch A, Bala K, Beretta L, et al. The human proteome project: current state and future direction. Mol Cell Proteomics. 2011 Jul;10(7):M111.009993.
100. HUPO ' s Human Proteome Project : the next big thing ? 2008;7(6):2192.
101. Wang Q, Wen B, Yan G, Wei J, Xie L, Xu S, et al. Qualitative and Quantitative Expression Status of the Human Chromosome 20 Genes in Cancer Tissues and the Representative Cell Lines. J Proteome Res [Internet]. 2013 Jan 4;12(1):151–61. Available from: <https://doi.org/10.1021/pr3008336>
102. Hüttenhain R, Soste M, Selevsek N, Röst H, Sethi A, Carapito C, et al. Reproducible Quantification of Cancer-Associated Proteins in Body Fluids Using Targeted Proteomics. Sci Transl Med [Internet]. 2012 Jul 11;4(142):142ra94 LP-142ra94. Available from: <http://stm.sciencemag.org/content/4/142/142ra94.abstract>
103. Zhang B, Wang J, Wang X, Zhu J, Liu Q, Shi Z, et al. Proteogenomic characterization of human colon and rectal cancer. Nature. 2014;513(7518):382–7.
104. Huang Z, Ma L, Huang C, Li Q, Nice EC. Proteomic profiling of human plasma for cancer biomarker discovery. Proteomics [Internet]. 2017 Mar 1;17(6):1600240. Available from: <https://doi.org/10.1002/pmic.201600240>
105. Mertins P, Mani DR, Ruggles K V, Gillette MA, Clauser KR, Wang P, et al. Proteogenomics connects somatic mutations to signalling in breast cancer. Nature. 2016 Jun;534(7605):55–62.

106. Zhang H, Liu T, Zhang Z, Payne SH, Zhang B, McDermott JE, et al. Integrated Proteogenomic Characterization of Human High-Grade Serous Ovarian Cancer. *Cell*. 2016;166(3):755–65.
107. Sebastian S, Bulun SE. A Highly Complex Organization of the Regulatory Region of the Human CYP19 (Aromatase) Gene Revealed by the Human Genome Project. *J Clin Endocrinol Metab* [Internet]. 2001 Oct 1;86(10):4600–2. Available from: <https://doi.org/10.1210/jcem.86.10.7947>
108. Aebersold R, Bader GD, Edwards AM, Van Eyk JE, Kussmann M, Qin J, et al. The biology/disease-driven human proteome project (B/D-HPP): Enabling protein research for the life sciences community. Vol. 12, *Journal of Proteome Research*. 2013. p. 23–7.
109. Paik Y-K, Jeong S-K, Omenn GS, Uhlen M, Hanash S, Cho SY, et al. The Chromosome-Centric Human Proteome Project for cataloging proteins encoded in the genome. *Nat Biotechnol*. 2012;30(3):221–3.
110. Gupta MK, Jayaram S, Madugundu AK, Chavan S, Advani J, Pandey A, et al. Chromosome-centric human proteome project: Deciphering proteins associated with glioma and neurodegenerative disorders on chromosome 12. *J Proteome Res*. 2014;13(7):3178–90.
111. Wang Q, Wen B, Yan G, Wei J, Xie L, Xu S, et al. Qualitative and Quantitative Expression Status of the Human Chromosome 20 Genes in Cancer Tissues and the Representative Cell Lines. *J Proteome Res*. 2013 Jan;12(1):151–61.
112. Fan W, Ge G, Liu Y, Wang W, Liu L, Jia Y. Proteomics integrated with metabolomics: analysis of the internal causes of nutrient changes in alfalfa at different growth stages. *BMC Plant Biol* [Internet]. 2018;18(1):78. Available from: <https://doi.org/10.1186/s12870-018-1291-8>
113. Kopka J, Schauer N, Krueger S, Birkemeyer C, Usadel B, Bergmüller E, et al. GMD@CSB.DB: the Golm Metabolome Database. *Bioinformatics* [Internet]. 2005 Apr 15;21(8):1635–8. Available from: <https://doi.org/10.1093/bioinformatics/bti236>
114. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, et al. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res*. 2018 Jan;46(D1):D608–17.
115. Sud M, Fahy E, Cotter D, Azam K, Vadivelu I, Burant C, et al. Metabolomics Workbench: An international repository for metabolomics data and metadata, metabolite standards, protocols, tutorials and training, and analysis tools. *Nucleic Acids Res*. 2016 Jan;44(D1):D463–70.
116. Wedge DC, Allwood JW, Dunn W, Vaughan AA, Simpson K, Brown M, et al. Is serum or plasma more appropriate for intersubject comparisons in metabolomic studies? An assessment in patients with small-cell lung cancer. *Anal Chem*. 2011;83(17):6689–97.
117. Pasikanti KK, Esuvaranathan K, Hong Y, Ho PC, Mahendran R, Raman Nee Mani L, et al. Urinary metabotyping of bladder cancer using two-dimensional gas chromatography time-of-flight mass spectrometry. *J Proteome Res*. 2013;12(9):3865–73.

118. Sugimoto M, Wong DT, Hirayama A, Soga T, Tomita M. Capillary electrophoresis mass spectrometry-based saliva metabolomics identified oral, breast and pancreatic cancer-specific profiles. *Metabolomics* [Internet]. 2010;6(1):78–95. Available from: <https://doi.org/10.1007/s11306-009-0178-y>
119. Agren R, Bordel S, Mardinoglu A, Pornputtapong N, Nookaew I, Nielsen J. Reconstruction of Genome-Scale Active Metabolic Networks for 69 Human Cell Types and 16 Cancer Types Using INIT. *PLOS Comput Biol* [Internet]. 2012 May 17;8(5):e1002518. Available from: <https://doi.org/10.1371/journal.pcbi.1002518>
120. Hattori A, Tsunoda M, Konuma T, Kobayashi M, Nagy T, Glushka J, et al. Cancer progression by reprogrammed BCAA metabolism in myeloid leukaemia. *Nature* [Internet]. 2017;545(7655):500–4. Available from: <http://dx.doi.org/10.1038/nature22314>
121. Mary Christine Playdon. Dietary Biomarker Discovery using Metabolomics and Application to Understanding Breast Cancer Etiology. *J Chem Inf Model*. 2013;53(9):1689–99.
122. Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*. 2000 Jan;28(1):27–30.
123. Mi H, Muruganujan A, Casagrande JT, Thomas PD. Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc* [Internet]. 2013;8(8):1551–66. Available from: <https://doi.org/10.1038/nprot.2013.092>
124. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res*. 2017 Jan;45(D1):D353–61.
125. Li X-N, Wang Z-J, Ye C-X, Zhao B-C, Li Z-L, Yang Y. RNA sequencing reveals the expression profiles of circRNA and indicates that circDDX17 acts as a tumor suppressor in colorectal cancer. *J Exp Clin Cancer Res* [Internet]. 2018;37(1):325. Available from: <https://doi.org/10.1186/s13046-018-1006-x>
126. Feng H, Gu Z-Y, Li Q, Liu Q-H, Yang X-Y, Zhang J-J. Identification of significant genes with poor prognosis in ovarian cancer via bioinformatical analysis. *J Ovarian Res* [Internet]. 2019;12(1):35. Available from: <https://doi.org/10.1186/s13048-019-0508-2>
127. Stenson PD, Ball E V, Mort M, Phillips AD, Shaw K, Cooper DN. The Human Gene Mutation Database (HGMD) and its exploitation in the fields of personalized genomics and molecular evolution. *Curr Protoc Bioinforma*. 2012 Sep;Chapter 1:Unit1.13.
128. Turcan S, Rohle D, Goenka A, Walsh LA, Fang F, Yilmaz E, et al. IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature* [Internet]. 2012;483(7390):479–83. Available from: <https://doi.org/10.1038/nature10866>
129. Karp PD, Riley M, Paley SM, Pellegrini-Toole A. The MetaCyc Database. *Nucleic Acids Res*. 2002 Jan;30(1):59–61.

130. Tang J, Aittokallio T. Network Pharmacology Strategies Toward Multi-Target Anticancer Therapies: From Computational Models to Experimental Design Principles. *Curr Pharm Des.* 2014;20(1):23–36.
131. Bader GD, Donaldson I, Wolting C, Ouellette BF, Pawson T, Hogue CW. BIND--The Biomolecular Interaction Network Database. *Nucleic Acids Res.* 2001 Jan;29(1):242–5.
132. Chen Y, Choong L-Y, Lin Q, Philp R, Wong C-H, Ang B-K, et al. Differential Expression of Novel Tyrosine Kinase Substrates during Breast Cancer Development *. *Mol Cell Proteomics* [Internet]. 2007;6(12):2072–87. Available from: <https://www.sciencedirect.com/science/article/pii/S1535947620319009>
133. Jia P, Sun J, Guo AY, Zhao Z. SZGR: a comprehensive schizophrenia gene resource. *Mol Psychiatry* [Internet]. 2010;15(5):453–62. Available from: <https://doi.org/10.1038/mp.2009.93>
134. Vinayagam A, Gibson TE, Lee HJ, Yilmazel B, Roesel C, Hu Y, et al. Controllability analysis of the directed human protein interaction network identifies disease genes and drug targets. *Proc Natl Acad Sci U S A.* 2016;113(18):4976–81.
135. Xenarios I, Rice DW, Salwinski L, Baron MK, Marcotte EM, Eisenberg D. DIP: the database of interacting proteins. *Nucleic Acids Res.* 2000 Jan;28(1):289–91.
136. Goh WW Bin, Lee YH, Zubaidah RM, Jin J, Dong D, Lin Q, et al. Network-based pipeline for analyzing MS data: An application toward liver cancer. *J Proteome Res.* 2011;10(5):2261–72.
137. Fu LL, Zhao X, Xu HL, Wen X, Wang SY, Liu B, et al. Identification of microRNA-regulated autophagic pathways in plant lectin-induced cancer cell death. *Cell Prolif.* 2012;45(5):477–85.
138. Graeber TG, Eisenberg D. Bioinformatic identification of potential autocrine signaling loops in cancers from gene expression profiles. *Nat Genet* [Internet]. 2001;29(3):295–300. Available from: <https://doi.org/10.1038/ng755>
139. Von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, et al. STRING: known and predicted protein–protein associations, integrated and transferred across organisms. *Nucleic Acids Res.* 2005;33(suppl_1):D433–7.
140. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 2017 Jan;45(D1):D362–8.
141. Mlecnik B, Tosolini M, Charoentong P, Kirilovsky A, Bindea G, Berger A, et al. Biomolecular Network Reconstruction Identifies T-Cell Homing Factors Associated With Survival in Colorectal Cancer. *Gastroenterology* [Internet]. 2010;138(4):1429–40. Available from: <http://dx.doi.org/10.1053/j.gastro.2009.10.057>
142. Bindea G, Mlecnik B, Fridman W-H, Pagès F, Galon J. Natural immunity to cancer in

- humans. *Curr Opin Immunol* [Internet]. 2010;22(2):215–22. Available from: <https://www.sciencedirect.com/science/article/pii/S0952791510000385>
143. Azpeitia E, Balanzario EP, Wagner A. Signaling pathways have an inherent need for noise to acquire information. *BMC Bioinformatics* [Internet]. 2020;21(1):462. Available from: <https://doi.org/10.1186/s12859-020-03778-x>
144. Peng R, Xu L, Wang H, Lyu Y, Wang D, Bi C, et al. DNA-based artificial molecular signaling system that mimics basic elements of reception and response. *Nat Commun* [Internet]. 2020;11(1):978. Available from: <https://doi.org/10.1038/s41467-020-14739-6>
145. Suderman R, Deeds EJ. Intrinsic limits of information transmission in biochemical signalling motifs. *Interface Focus*. 2018;8(6).
146. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. *Nat Med*. 2013 Nov;19(11):1423–37.
147. Thul PJ, Lindskog C. The human protein atlas: A spatial map of the human proteome. *Protein Sci*. 2018;27(1):233–44.
148. Sehuler GD. Pieces of use puzzle: Expressed sequence tags and the catalog of human genes. *J Mol Med*. 1997;75(10):694–8.
149. Celik S, Logsdon BA, Battle S, Drescher CW, Rendi M, Hawkins RD, et al. Extracting a low-dimensional description of multiple gene expression datasets reveals a potential driver for tumor-associated stroma in ovarian cancer. *Genome Med* [Internet]. 2016;8(1):66. Available from: <https://doi.org/10.1186/s13073-016-0319-7>
150. Levi-Galibov O, Lavon H, Wassermann-Dozorets R, Pevsner-Fischer M, Mayer S, Wershof E, et al. Heat Shock Factor 1-dependent extracellular matrix remodeling mediates the transition from chronic intestinal inflammation to colon cancer. *Nat Commun* [Internet]. 2020;11(1):6245. Available from: <https://doi.org/10.1038/s41467-020-20054-x>
151. Mirabelli P, Coppola L, Salvatore M. Cancer Cell Lines Are Useful Model Systems for Medical Research. *Cancers (Basel)*. 2019 Aug;11(8).
152. Li H, Ning S, Ghandi M, Kryukov G V., Gopal S, Deik A, et al. The landscape of cancer cell line metabolism. *Nat Med* [Internet]. 2019;25(5):850–60. Available from: <http://dx.doi.org/10.1038/s41591-019-0404-8>
153. Hanniford D, Ulloa-Morales A, Karz A, Berzoti-Coelho MG, Moubarak RS, Sánchez-Sendra B, et al. Epigenetic Silencing of CDR1as Drives IGF2BP3-Mediated Melanoma Invasion and Metastasis. *Cancer Cell* [Internet]. 2020;37(1):55–70.e15. Available from: <https://www.sciencedirect.com/science/article/pii/S153561081930577X>
154. National Cancer Institute [Internet]. Available from: <https://www.cancer.gov/>
155. BROAD institute [Internet]. Available from: <https://www.broadinstitute.org/>
156. Wellcome Sanger Institute [Internet]. Available from: <https://www.sanger.ac.uk/>

157. MGH Institute of Health Professionals [Internet]. Available from: <https://www.mghihp.edu/>
158. Prior F, Smith K, Sharma A, Kirby J, Tarbox L, Clark K, et al. The public cancer radiology imaging collections of The Cancer Imaging Archive. *Sci data*. 2017 Sep;4:170124.
159. Li H, Zhu Y, Burnside ES, Drukker K, Hoadley KA, Fan C, et al. MR imaging radiomics signatures for predicting the risk of breast cancer recurrence as given by research versions of MammaPrint, oncotype DX, and PAM50 gene assays. *Radiology*. 2016;281(2):382–91.
160. Sun R, Limkin EJ, Vakalopoulou M, Dercle L, Champiat S, Han SR, et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol* [Internet]. 2018;19(9):1180–91. Available from: <https://www.sciencedirect.com/science/article/pii/S1470204518304133>
161. Khosravi P, Kazemi E, Imielinski M, Elemento O, Hajirasouliha I. Deep Convolutional Neural Networks Enable Discrimination of Heterogeneous Digital Pathology Images. *EBioMedicine* [Internet]. 2018;27:317–28. Available from: <https://www.sciencedirect.com/science/article/pii/S2352396417305078>
162. Guo S, Jiang X, Mao B, Li Q-X. The design, analysis and application of mouse clinical trials in oncology drug development. *BMC Cancer* [Internet]. 2019;19(1):718. Available from: <https://doi.org/10.1186/s12885-019-5907-7>
163. Strickland JC, Lile JA, Rush CR, Stoops WW. Comparing exponential and exponentiated models of drug demand in cocaine users. *Exp Clin Psychopharmacol*. 2016 Dec;24(6):447–55.
164. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res*. 2006 Jan;34(Database issue):D668-72.
165. Augustyn A, Borromeo M, Wang T, Fujimoto J, Shao C, Dospoy PD, et al. ASCL1 is a lineage oncogene providing therapeutic targets for high-grade neuroendocrine lung cancers. *Proc Natl Acad Sci U S A*. 2014;111(41):14788–93.
166. Han K, Jeng EE, Hess GT, Morgens DW, Li A, Bassik MC. Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. *Nat Biotechnol*. 2017;35(5):463–74.
167. Fernández-Navarro P, López-Nieva P, Piñeiro-Yañez E, Carreño-Tarragona G, Martínez-López J, Sánchez Pérez R, et al. The use of PanDrugs to prioritize anticancer drug treatments in a case of T-ALL based on individual genomic data. *BMC Cancer* [Internet]. 2019;19(1):1005. Available from: <https://doi.org/10.1186/s12885-019-6209-9>
168. Powathil GG, Swat M, Chaplain MAJ. Systems oncology: Towards patient-specific treatment regimes informed by multiscale mathematical modelling. *Semin Cancer Biol*

- [Internet]. 2015;30:13–20. Available from:
<http://www.sciencedirect.com/science/article/pii/S1044579X14000273>
169. Valladares-Ayerbes M, Haz-Conde M, Blanco-Calvo M. Systems oncology: toward the clinical application of cancer systems biology. *Futur Oncol*. 2015;11(4):553–5.
 170. Anderson ARA, Rejniak KA, Gerlee P, Quaranta V. Microenvironment driven invasion: A multiscale multimodel investigation. *J Math Biol*. 2009;58(4–5):579–624.
 171. Stolarska MA, Yangjin KIM, Othmer HG. Multi-scale models of cell and tissue dynamics. *Philos Trans R Soc A Math Phys Eng Sci*. 2009;367(1902):3525–53.
 172. Silva A, Anderson A, Gatenby R. A multiscale model of the bone marrow and hematopoiesis. *Math Biosci Eng*. 2011;8(2):643–58.
 173. Chaudhary SU, Shin S, Won J, Cho K, Member S. Multiscale Modeling of Tumorigenesis Induced by Mitochondrial Incapacitation in Cell Death. *IEEE Trans Biomed Eng*. 2011;58(2):3028–32.
 174. Barbarroux L, Michel P, Adimy M, Crauste F. Multi-scale modeling of the CD8 immune response. *AIP Conf Proc*. 2016;1738.
 175. Szabó A, Merks RMH. Blood vessel tortuosity selects against evolution of aggressive tumor cells in confined tissue environments: A modeling approach. Vol. 13, *PLoS Computational Biology*. 2017. 1–32 p.
 176. Kumar S, Das A, Sen S. Multicompartment cell-based modeling of confined migration: regulation by cell intrinsic and extrinsic factors. *Mol Biol Cell*. 2018;29(13):1599–610.
 177. Unni P. *Mathematical Modeling, Analysis, and Simulation of Tumor*. 2019;2019.
 178. Leppert M, Dobbs M, Scambler P, O’Connell P, Nakamura Y, Stauffer D, et al. The gene for familial polyposis coli maps to the long arm of chromosome 5. *Science* (80-) [Internet]. 1987 Dec 4;238(4832):1411 LP – 1413. Available from:
<http://science.sciencemag.org/content/238/4832/1411.abstract>
 179. Mehl LE. A mathematical computer simulation model for the development of colonic polyps and colon cancer. *J Surg Oncol*. 1991;47(4):243–52.
 180. Stratmann AT, Fecher D, Wangorsch G, Göttlich C, Walles T, Walles H, et al. Establishment of a human 3D lung cancer model based on a biological tissue matrix combined with a Boolean in silico model. *Mol Oncol*. 2014 Mar;8(2):351–65.
 181. Greenspan HP. Models for the Growth of a Solid Tumor by Diffusion. *Stud Appl Math* [Internet]. 1972 Dec 1;51(4):317–40. Available from:
<https://doi.org/10.1002/sapm1972514317>
 182. Greenspan HP. On the growth and stability of cell cultures and solid tumors. *J Theor Biol*. 1976;56(1):229–42.
 183. Chaplain MAJ. Avascular growth, angiogenesis and vascular growth in solid tumours: The

- mathematical modelling of the stages of tumour development. *Math Comput Model.* 1996;23(6):47–87.
184. Shackney SE, Smith CA, Miller BW, Burholt DR, Murtha K, Pollice AA, et al. Model for the Genetic Evolution of Human Solid Tumors. *Cancer Res.* 1989;49(12):3344–54.
 185. Gerlee P, Nelander S. The Impact of Phenotypic Switching on Glioblastoma Growth and Invasion. *PLOS Comput Biol* [Internet]. 2012 Jun 14;8(6):e1002556. Available from: <https://doi.org/10.1371/journal.pcbi.1002556>
 186. Setty Y. In-silico models of stem cell and developmental systems. *Theor Biol Med Model.* 2014;11:1.
 187. Wang Z, Bordas V, Sagotsky J, Deisboeck TS. Identifying therapeutic targets in a combined EGFR–TGFβR signalling cascade using a multiscale agent-based cancer model. *Math Med Biol a J IMA.* 2012;29(1):95–108.
 188. Wang Z, Bordas V, Deisboeck T. Identification of Critical Molecular Components in a Multiscale Cancer Model Based on the Integration of Monte Carlo, Resampling, and ANOVA [Internet]. Vol. 2, *Frontiers in Physiology*. 2011. p. 35. Available from: <https://www.frontiersin.org/article/10.3389/fphys.2011.00035>
 189. Margolin G, Petrykowska HM, Jameel N, Bell DW, Young AC, Elnitski L. Robust Detection of DNA Hypermethylation of ZNF154 as a Pan-Cancer Locus with in Silico Modeling for Blood-Based Diagnostic Development. *J Mol Diagnostics* [Internet]. 2016;18(2):283–98. Available from: <https://www.sciencedirect.com/science/article/pii/S1525157815002743>
 190. Jahangiri A, Amani J, Halabian R, Imani fooladi AA. In Silico Analyses of Staphylococcal Enterotoxin B as a DNA Vaccine for Cancer Therapy. *Int J Pept Res Ther* [Internet]. 2018;24(1):131–42. Available from: <https://doi.org/10.1007/s10989-017-9595-3>
 191. Smith EAK, Henthorn NT, Warmenhoven JW, Ingram SP, Aitkenhead AH, Richardson JC, et al. In Silico Models of DNA Damage and Repair in Proton Treatment Planning: A Proof of Concept. *Sci Rep* [Internet]. 2019;9(1):19870. Available from: <https://doi.org/10.1038/s41598-019-56258-5>
 192. Adams J. Transcriptome: connecting the genome to gene function. *Nat Educ.* 2008;1(1):195.
 193. Kirilov K, Borovska P. Conceptual model of integrated approach for in silico knowledge data discovery for breast cancer diagnostics and precision therapy. In: *AIP Conference Proceedings*. AIP Publishing LLC; 2019. p. 20003.
 194. Huang E, Cheng SH, Dressman H, Pittman J, Tsou MH, Horng CF, et al. Gene expression predictors of breast cancer outcomes. *Lancet.* 2003;361(9369):1590–6.
 195. Cheng F, Liu C, Lin CC, Zhao J, Jia P, Li WH, et al. A Gene Gravity Model for the Evolution of Cancer Genomes: A Study of 3,000 Cancer Genomes across 9 Cancer Types. *PLoS Comput Biol.* 2015;11(9):1–25.

196. Chen J-S, Hung W-S, Chan H-H, Tsai S-J, Sun HS. In silico identification of oncogenic potential of fyn-related kinase in hepatocellular carcinoma. *Bioinformatics* [Internet]. 2013 Feb 15;29(4):420–7. Available from: <https://doi.org/10.1093/bioinformatics/bts715>
197. Agren R, Mardinoglu A, Asplund A, Kampf C, Uhlen M, Nielsen J. Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling. *Mol Syst Biol*. 2014 Mar;10(3):721.
198. Béal J, Montagud A, Traynard P, Barillot E, Calzone L. Personalization of Logical Models With Multi-Omics Data Allows Clinical Stratification of Patients. *Front Physiol* [Internet]. 2019;9:1965. Available from: <https://www.frontiersin.org/article/10.3389/fphys.2018.01965>
199. Béal J, Pantolini L, Noël V, Barillot E, Calzone L. Personalized logical models to investigate cancer response to BRAF treatments in melanomas and colorectal cancers. *PLOS Comput Biol* [Internet]. 2021 Jan 28;17(1):e1007900. Available from: <https://doi.org/10.1371/journal.pcbi.1007900>
200. Rodriguez J, Ren G, Day CR, Zhao K, Chow CC, Larson DR. Intrinsic Dynamics of a Human Gene Reveal the Basis of Expression Heterogeneity. *Cell*. 2019 Jan;176(1–2):213–226.e18.
201. Baloria U, Akhoun BA, Gupta SK, Sharma S, Verma V. In silico proteomic characterization of human epidermal growth factor receptor 2 (HER-2) for the mapping of high affinity antigenic determinants against breast cancer. *Amino Acids*. 2012;42(4):1349–60.
202. Akhoun BA, Slathia PS, Sharma P, Gupta SK, Verma V. In silico identification of novel protective VSG antigens expressed by *Trypanosoma brucei* and an effort for designing a highly immunogenic DNA vaccine using IL-12 as adjuvant. *Microb Pathog* [Internet]. 2011;51(1):77–87. Available from: <https://www.sciencedirect.com/science/article/pii/S0882401011000222>
203. Fang J, Wu Z, Cai C, Wang Q, Tang Y, Cheng F. Quantitative and Systems Pharmacology. 1. In Silico Prediction of Drug–Target Interactions of Natural Products Enables New Targeted Cancer Therapy. *J Chem Inf Model* [Internet]. 2017 Nov 27;57(11):2657–71. Available from: <https://doi.org/10.1021/acs.jcim.7b00216>
204. Azevedo R, Silva AMN, Reis CA, Santos LL, Ferreira JA. In silico approaches for unveiling novel glycomarkers in cancer. *J Proteomics* [Internet]. 2018;171:95–106. Available from: <https://www.sciencedirect.com/science/article/pii/S1874391917302671>
205. Lee SWL, Seager RJ, Litvak F, Spill F, Sieow JL, Leong PH, et al. Integrated in silico and 3D in vitro model of macrophage migration in response to physical and chemical factors in the tumor microenvironment. *Integr Biol* [Internet]. 2020 Apr 20;12(4):90–108. Available from: <https://doi.org/10.1093/intbio/zyaa007>
206. Ma H, Sorokin A, Mazein A, Selkov A, Selkov E, Demin O, et al. The Edinburgh human metabolic network reconstruction and its functional analysis. *Mol Syst Biol*. 2007;3:135.
207. Folger O, Jerby L, Frezza C, Gottlieb E, Ruppin E, Shlomi T. Predicting selective drug

- targets in cancer through metabolic networks. *Mol Syst Biol* [Internet]. 2011 Jan 1;7(1):501. Available from: <https://doi.org/10.1038/msb.2011.35>
208. Aurich MK, Paglia G, Rolfsson Ó, Hrafnisdóttir S, Magnúsdóttir M, Stefaniak MM, et al. Prediction of intracellular metabolic states from extracellular metabolomic data. *Metabolomics*. 2015;11(3):603–19.
209. Yurkovich JT, Yang L, Palsson BO. Biomarkers are used to predict quantitative metabolite concentration profiles in human red blood cells. *PLOS Comput Biol* [Internet]. 2017 Mar 6;13(3):e1005424. Available from: <https://doi.org/10.1371/journal.pcbi.1005424>
210. Alakwaa FM, Chaudhary K, Garmire LX. Deep Learning Accurately Predicts Estrogen Receptor Status in Breast Cancer Metabolomics Data. *J Proteome Res*. 2018;17(1):337–47.
211. Azadi M, Alemi F, Sadeghi S, Mohammadi M, Rahimi NA, Mirzaie S, et al. An integrative in silico mathematical modelling study of the anti-cancer effect of clove extract (*Syzygium aromaticum*) combined with in vitro metabolomics study using ¹Hnmr spectroscopy. *Iran J Biotechnol*. 2020;18(3):45–54.
212. Balding D, McElwain DL. A mathematical model of tumour-induced capillary growth. *J Theor Biol*. 1985 May;114(1):53–73.
213. Anderson ARA. A hybrid mathematical model of solid tumour invasion: the importance of cell adhesion. *Math Med Biol A J IMA* [Internet]. 2005;22(2):163–86. Available from: <https://doi.org/10.1093/imammb/dqi005>
214. McDougall SR, Anderson ARA, Chaplain MAJ, Sherratt JA. Mathematical modelling of flow through vascular networks: implications for tumour-induced angiogenesis and chemotherapy strategies. *Bull Math Biol*. 2002 Jul;64(4):673–702.
215. Anderson ARA, Chaplain MAJ. Continuous and discrete mathematical models of tumor-induced angiogenesis. *Bull Math Biol*. 1998;60(5):857–99.
216. Anderson ARA. A hybrid mathematical model of solid tumour invasion: the importance of cell adhesion. *Math Med Biol A J IMA*. 2005;22(2):163–86.
217. Nakazawa K, Kurishima K, Tamura T, Kagohashi K, Ishikawa H, Satoh H, et al. Specific organ metastases and survival in small cell lung cancer. *Oncol Lett*. 2012 Oct;4(4):617–20.
218. Cortesi M, Liverani C, Mercatali L, Ibrahim T, Giordano E. An in-silico study of cancer cell survival and spatial distribution within a 3D microenvironment. *Sci Rep* [Internet]. 2020;10(1):12976. Available from: <https://doi.org/10.1038/s41598-020-69862-7>
219. Aubert M, Badoual M, Christov C, Grammaticos B. A model for glioma cell migration on collagen and astrocytes. *J R Soc Interface*. 2008;5(18):75–83.
220. Harris LA, Beik S, Ozawa PMM, Jimenez L, Weaver AM. Modeling heterogeneous tumor growth dynamics and cell–cell interactions at single-cell and cell-population resolution.

- Curr Opin Syst Biol [Internet]. 2019;17:24–34. Available from:
<https://www.sciencedirect.com/science/article/pii/S2452310019300411>
221. Swanson KR, Alvord ECJ, Murray JD. A quantitative model for differential motility of gliomas in grey and white matter. *Cell Prolif*. 2000 Oct;33(5):317–29.
222. Zhang L, Athale CA, Deisboeck TS. Development of a three-dimensional multiscale agent-based tumor model: Simulating gene-protein interaction profiles, cell phenotypes and multicellular patterns in brain cancer. *J Theor Biol*. 2007;244(1):96–107.
223. Perfahl H, Byrne HM, Chen T, Estrella V, Alarcón T, Lapin A, et al. Multiscale Modelling of Vascular Tumour Growth in 3D: The Roles of Domain Size and Boundary Conditions. *PLoS One* [Internet]. 2011 Apr 13;6(4):e14790. Available from:
<https://doi.org/10.1371/journal.pone.0014790>
224. Anderson ARA, Rejniak KA, Gerlee P, Quaranta V. Modelling of Cancer Growth, Evolution and Invasion: Bridging Scales and Models. *Math Model Nat Phenom*. 2007;2(3):1–29.
225. Vavourakis V, Wijeratne PA, Shipley R, Loizidou M, Stylianopoulos T, Hawkes DJ. A Validated Multiscale In-Silico Model for Mechano-sensitive Tumour Angiogenesis and Growth. *PLOS Comput Biol* [Internet]. 2017 Jan 26;13(1):e1005259. Available from:
<https://doi.org/10.1371/journal.pcbi.1005259>
226. Norton K-A, Wallace T, Pandey NB, Popel AS. An agent-based model of triple-negative breast cancer: the interplay between chemokine receptor CCR5 expression, cancer stem cells, and hypoxia. *BMC Syst Biol*. 2017;11(1):68.
227. Norton KA, Gong C, Jamalian S, Popel AS. Multiscale agent-based and hybrid modeling of the tumor immune microenvironment. *Processes*. 2019;7(1):1–23.
228. Karolak A, Poonja S, Rejniak KA. Morphophenotypic classification of tumor organoids as an indicator of drug exposure and penetration potential. *PLOS Comput Biol* [Internet]. 2019 Jul 16;15(7):e1007214. Available from:
<https://doi.org/10.1371/journal.pcbi.1007214>
229. Kenny PA, Lee GY, Myers CA, Neve RM, Semeiks JR, Spellman PT, et al. The morphologies of breast cancer cell lines in three-dimensional assays correlate with their profiles of gene expression. *Mol Oncol*. 2007;1(1):84–96.
230. Berrouet C, Dorilas N, Rejniak KA, Tuncer N. Comparison of Drug Inhibitory Effects (IC50) in Monolayer and Spheroid Cultures. *Bull Math Biol*. 2020;82(6):68.
231. Stamatakis GS, Graf N, Radhakrishnan R. Multiscale Cancer Modeling and In Silico Oncology: Emerging Computational Frontiers in Basic and Translational Cancer Research. *J Bioeng Biomed Sci* [Internet]. 2013/05/25. 2013;3(2):e114. Available from:
<https://pubmed.ncbi.nlm.nih.gov/30740263>
232. Nagaraj SH, Reverter A. A Boolean-based systems biology approach to predict novel genes associated with cancer: Application to colorectal cancer. *BMC Syst Biol* [Internet].

- 2011;5(1):35. Available from: <https://doi.org/10.1186/1752-0509-5-35>
233. Fumiã HF, Martins ML. Boolean Network Model for Cancer Pathways: Predicting Carcinogenesis and Targeted Therapy Outcomes. PLoS One [Internet]. 2013 Jul 26;8(7):e69008. Available from: <https://doi.org/10.1371/journal.pone.0069008>
234. Yang L, Meng Y, Bao C, Liu W, Ma C, Li A, et al. Robustness and Backbone Motif of a Cancer Network Regulated by miR-17-92 Cluster during the G1/S Transition. PLoS One [Internet]. 2013 Mar 1;8(3):e57009. Available from: <https://doi.org/10.1371/journal.pone.0057009>
235. der Heyde S Von, Bender C, Henjes F, Sonntag J, Korf U, Beißbarth T. Boolean ErbB network reconstructions and perturbation simulations reveal individual drug response in different breast cancer cell lines. BMC Syst Biol [Internet]. 2014;8(1):75. Available from: <https://doi.org/10.1186/1752-0509-8-75>
236. Gondal MN, Butt RN, Shah OS, Nasir Z, Hussain R, Khawar H, et al. In silico Drosophila Patient Model Reveals Optimal Combinatorial Therapies for Colorectal Cancer. bioRxiv [Internet]. 2020; Available from: <https://www.biorxiv.org/content/early/2020/09/01/2020.08.31.274829>
237. Klamt S, Saez-Rodriguez J, Gilles ED. Structural and functional analysis of cellular networks with CellNetAnalyzer. BMC Syst Biol. 2007;1(1):2.
238. Kachalo S, Zhang R, Sontag E, Albert R, DasGupta B. NET-SYNTHESIS: a software for synthesis, inference and simplification of signal transduction networks. Bioinformatics [Internet]. 2008 Jan 15;24(2):293–5. Available from: <https://doi.org/10.1093/bioinformatics/btm571>
239. Mendes P. Dynamics , Steady States and Control of Biochemical and Other Systems. Cabios. 1993;9(5):563–71.
240. Starruß J, De Back W, Brusch L, Deutsch A. Morpheus: a user-friendly modeling environment for multiscale and multicellular systems biology. Bioinformatics. 2014;30(9):1331–2.
241. Tomita M, Hashimoto K, Takahashi K, Shimizu TS, Matsuzaki Y, Miyoshi F, et al. E-CELL: Software environment for whole-cell simulation. Bioinformatics. 1999;
242. Schaff J, Slepchenko BM, Morgan FR, Cowan a E, Moraru II, Wagner J, et al. Computational cell biology: The Virtual Cell software environment. Mol Biol Cell. 2000;11(10):690.
243. Moraru II, Schaff JC, Slepchenko BM, Blinov ML, Morgan F, Lakshminarayana A, et al. Virtual Cell modelling and simulation software environment. In: IET Systems Biology. 2008.
244. Ion I. Moraru, James C. Schaff, Boris M. Slepchenko, Michael Blinov, Frank Morgan, Anuradha Lakshminarayana, Fei Gao, Ye Li and LML. The Virtual Cell Modeling and

- Simulation Software Environment. PMC. 2008;
245. Eissing T, Kuepfer L, Becker C, Block M, Coboeken K, Gaub T, et al. A computational systems biology software platform for multiscale modeling and simulation: Integrating whole-body physiology, disease biology, and molecular reaction networks. *Front Physiol.* 2011;FEB(February):1–10.
 246. Hoops S, Gauges R, Lee C, Pahle J, Simus N, Singhal M, et al. COPASI - A COmplex PATHway Simulator. *Bioinformatics.* 2006;22(24):3067–74.
 247. Kauffman SA. Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol.* 1969 Mar;22(3):437–67.
 248. Glass L, Kauffman SA. The logical analysis of continuous, non-linear biochemical control networks. *J Theor Biol.* 1973 Apr;39(1):103–29.
 249. Wegner K, Knabe J, Robinson M, Egri-Nagy A, Schilstra M. The NetBuilder' project: development of a tool for constructing, simulating, evolving, and analysing complex regulatory networks. *BMC Syst Biol.* 2007;1(Suppl 1):P72.
 250. Albert I, Thakar J, Li S, Zhang R, Albert R. Boolean network simulations for life scientists. *Source Code Biol Med.* 2008;3.
 251. Zheng J, Zhang D, Przytycki PF, Zielinski R, Capala J, Przytycka TM. Simboolnet-A cytoscape plugin for dynamic simulation of signaling networks. Vol. 26, *Bioinformatics.* 2009. p. 141–2.
 252. Klamt S, Stelling J, Ginkel M, Gilles ED. FluxAnalyzer: exploring structure, pathways, and flux distributions in metabolic networks on interactive flux maps. *Bioinformatics.* 2003 Jan;19(2):261–9.
 253. Tian K, Rajendran R, Doddananjaiah M, Krstic-Demonacos M, Schwartz J-M. Dynamics of DNA Damage Induced Pathways to Cancer. *PLoS One [Internet].* 2013 Sep 4;8(9):e72303. Available from: <https://doi.org/10.1371/journal.pone.0072303>
 254. Hetmanski JHR, Zindy E, Schwartz J-M, Caswell PT. A MAPK-Driven Feedback Loop Suppresses Rac Activity to Promote RhoA-Driven Cancer Cell Invasion. *PLOS Comput Biol [Internet].* 2016 May 3;12(5):e1004909. Available from: <https://doi.org/10.1371/journal.pcbi.1004909>
 255. Saadatpour A, Wang R-S, Liao A, Liu X, Loughran TP, Albert I, et al. Dynamical and Structural Analysis of a T Cell Survival Network Identifies Novel Candidate Therapeutic Targets for Large Granular Lymphocyte Leukemia. *PLOS Comput Biol [Internet].* 2011 Nov 10;7(11):e1002267. Available from: <https://doi.org/10.1371/journal.pcbi.1002267>
 256. Steinway SN, Zanudo JGT, Ding W, Rountree CB, Feith DJ, Loughran TP, et al. Network modeling of TGF β signaling in hepatocellular carcinoma epithelial-to-mesenchymal transition reveals joint sonic hedgehog and Wnt pathway activation. *Cancer Res.* 2014;74(21):5963–77.

257. Naldi A, Berenguier D, Fauré A, Lopez F, Thieffry D, Chaouiya C. Logical modelling of regulatory networks with GINsim 2.3. *Biosystems* [Internet]. 2009;97(2):134–9. Available from: <https://www.sciencedirect.com/science/article/pii/S0303264709000665>
258. Flobak Å, Baudot A, Remy E, Thommesen L, Thieffry D, Kuiper M, et al. Discovery of Drug Synergies in Gastric Cancer Cells Predicted by Logical Modeling. *PLOS Comput Biol* [Internet]. 2015 Aug 28;11(8):e1004426. Available from: <https://doi.org/10.1371/journal.pcbi.1004426>
259. Remy E, Rebouissou S, Chaouiya C, Zinovyev A, Radvanyi F, Calzone L. A modeling approach to explain mutually exclusive and co-occurring genetic alterations in bladder tumorigenesis. *Cancer Res*. 2015;75(19):4042–52.
260. Stoll G, Caron B, Viara E, Dugourd A, Zinovyev A, Naldi A, et al. MaBoSS 2.0: an environment for stochastic Boolean modeling. *Bioinformatics*. 2017 Jul;33(14):2226–8.
261. Kondratova M, Barillot E, Zinovyev A, Calzone L. Modelling of Immune Checkpoint Network Explains Synergistic Effects of Combined Immune Checkpoint Inhibitor Therapy and the Impact of Cytokines in Patient Response. Vol. 12, *Cancers* . 2020.
262. Müssel C, Hopfensitz M, Kestler HA. BoolNet—an R package for generation, reconstruction and analysis of Boolean networks. *Bioinformatics*. 2010;26(10):1378–80.
263. Steinway SN, Biggs MB, Loughran Jr TP, Papin JA, Albert R. Inference of Network Dynamics and Metabolic Interactions in the Gut Microbiome. *PLOS Comput Biol* [Internet]. 2015 Jun 23;11(6):e1004338. Available from: <https://doi.org/10.1371/journal.pcbi.1004338>
264. Cohen DPA, Martignetti L, Robine S, Barillot E, Zinovyev A, Calzone L. Mathematical Modelling of Molecular Pathways Enabling Tumour Cell Invasion and Migration. *PLOS Comput Biol* [Internet]. 2015 Nov 3;11(11):e1004571. Available from: <https://doi.org/10.1371/journal.pcbi.1004571>
265. Shah OS, Chaudhary MFA, Awan HA, Fatima F, Arshad Z, Amina B, et al. ATLANTIS - Attractor Landscape Analysis Toolbox for Cell Fate Discovery and Reprogramming. *Sci Rep* [Internet]. 2018;8(1):1–11. Available from: <http://dx.doi.org/10.1038/s41598-018-22031-3>
266. Han B, Wang J. Quantifying robustness and dissipation cost of yeast cell cycle network: the funneled energy landscape perspectives. *Biophys J*. 2007;92(11):3755–63.
267. Choi M, Shi J, Jung SH, Chen X, Cho K-H. Attractor Landscape Analysis Reveals Feedback Loops in the p53 Network That Control the Cellular Response to DNA Damage. *Sci Signal* [Internet]. 2012;5(251):ra83--ra83. Available from: <https://stke.sciencemag.org/content/5/251/ra83>
268. Cho S-H, Park S-M, Lee H-S, Lee H-Y, Cho K-H. Attractor landscape analysis of colorectal tumorigenesis and its reversion. *BMC Syst Biol*. 2016;10(1):96.

269. Wittmann DM, Krumsiek J, Saez-Rodriguez J, Lauffenburger DA, Klamt S, Theis FJ. Transforming Boolean models to continuous models: methodology and application to T-cell receptor signaling. *BMC Syst Biol* [Internet]. 2009;3(1):98. Available from: <https://doi.org/10.1186/1752-0509-3-98>
270. Ricci G, De Maria F, Antonini G, Turella P, Bullo A, Stella L, et al. 7-Nitro-2,1,3-benzoxadiazole Derivatives, a New Class of Suicide Inhibitors for Glutathione S-Transferases: MECHANISM OF ACTION OF POTENTIAL ANTICANCER DRUGS*. *J Biol Chem* [Internet]. 2005;280(28):26397–405. Available from: <https://www.sciencedirect.com/science/article/pii/S0021925820568324>
271. Marín-Hernández A, Gallardo-Pérez JC, Rodríguez-Enríquez S, Encalada R, Moreno-Sánchez R, Saavedra E. Modeling cancer glycolysis. *Biochim Biophys Acta - Bioenerg* [Internet]. 2011;1807(6):755–67. Available from: <https://www.sciencedirect.com/science/article/pii/S0005272810007644>
272. Hoops S, Gauges R, Lee C, Pahle J, Simus N, Singhal M, et al. COPASI - A COMplex PATHway Simulator. *Bioinformatics*. 2006;22(24):3067–74.
273. Orton RJ, Adriaens ME, Gormand A, Sturm OE, Kolch W, Gilbert DR. Computational modelling of cancerous mutations in the EGFR/ERK signalling pathway. *BMC Syst Biol* [Internet]. 2009;3(1):100. Available from: <https://doi.org/10.1186/1752-0509-3-100>
274. Dalle Pezze P, Nelson G, Otten EG, Korolchuk VI, Kirkwood TBL, von Zglinicki T, et al. Dynamic Modelling of Pathways to Cellular Senescence Reveals Strategies for Targeted Interventions. *PLOS Comput Biol* [Internet]. 2014 Aug 28;10(8):e1003728. Available from: <https://doi.org/10.1371/journal.pcbi.1003728>
275. Funahashi A, Morohashi M, Kitano H, Tanimura N. CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. *Biosilico*. 2003;1(5):159–62.
276. Sauro HM, Hucka M, Finney A, Wellock C, Bolouri H, Doyle J, et al. Next generation simulation tools: the Systems Biology Workbench and BioSPICE integration. *OMICS*. 2003;7(4):355–72.
277. Calzone L, Gelay A, Zinovyev A, Radvanyi F, Barillot E. A comprehensive modular map of molecular interactions in RB/E2F pathway. *Mol Syst Biol* [Internet]. 2008 Jan 1;4(1):174. Available from: <https://doi.org/10.1038/msb.2008.7>
278. Grieco L, Calzone L, Bernard-Pierrot I, Radvanyi F, Kahn-Perlès B, Thieffry D. Integrative Modelling of the Influence of MAPK Network on Cancer Cell Fate Decision. *PLOS Comput Biol* [Internet]. 2013 Oct 24;9(10):e1003286. Available from: <https://doi.org/10.1371/journal.pcbi.1003286>
279. Ghaffarizadeh A, Friedman SH, Macklin P. BioFVM: An efficient, parallelized diffusive transport solver for 3-D biological simulations. *Bioinformatics*. 2016;32(8):1256–8.
280. Felix TF, Lopez Lapa RM, de Carvalho M, Bertoni N, Tokar T, Oliveira RA, et al. MicroRNA modulated networks of adaptive and innate immune response in pancreatic ductal

- adenocarcinoma. PLoS One [Internet]. 2019 May 31;14(5):e0217421. Available from: <https://doi.org/10.1371/journal.pone.0217421>
281. Meyer K, Ostrenko O, Bourantas G, Morales-Navarrete H, Porat-Shliom N, Segovia-Miranda F, et al. A Predictive 3D Multi-Scale Model of Biliary Fluid Dynamics in the Liver Lobule. *Cell Syst*. 2017;4(3):277-290.e9.
282. Ozik J, Collier N, Heiland R, An G, Macklin P. Learning-accelerated discovery of immune-tumour interactions. *Mol Syst Des Eng*. 2019;4(4):747–60.
283. Ozik J, Collier N, Wozniak JM, Macal C, Cockrell C, Friedman SH, et al. High-throughput cancer hypothesis testing with an integrated PhysiCell-EMEWS workflow. *BMC Bioinformatics* [Internet]. 2018;19(18):483. Available from: <https://doi.org/10.1186/s12859-018-2510-x>
284. Wang Y, Brodin E, Nishii K, Frieboes HB, Mumenthaler SM, Sparks JL, et al. Impact of tumor-parenchyma biomechanics on liver metastatic progression: a multi-model approach. *Sci Rep* [Internet]. 2021;11(1):1710. Available from: <https://doi.org/10.1038/s41598-020-78780-7>
285. Cortesi M, Liverani C, Mercatali L, Ibrahim T, Giordano E. Development and validation of an in-silico tool for the study of therapeutic agents in 3D cell cultures. *Comput Biol Med* [Internet]. 2021;130:104211. Available from: <https://www.sciencedirect.com/science/article/pii/S0010482521000056>
286. Edwards JS, Ibarra RU, Palsson BO. In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. *Nat Biotechnol* [Internet]. 2001;19(2):125–30. Available from: <https://doi.org/10.1038/84379>
287. Orton RJ, Sturm OE, Vyshemirsky V, Calder M, Gilbert DR, Kolch W. Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway. *Biochem J* [Internet]. 2005 Nov 22;392(2):249–61. Available from: <https://doi.org/10.1042/BJ20050908>
288. Loew LM, Schaff JC. The Virtual Cell: a software environment for computational cell biology. *TRENDS Biotechnol*. 2001;19(10):401–6.
289. Neves SR, Tsokas P, Sarkar A, Grace EA, Rangamani P, Taubenfeld SM, et al. Cell Shape and Negative Links in Regulatory Motifs Together Control Spatial Information Flow in Signaling Networks. *Cell* [Internet]. 2008;133(4):666–80. Available from: <https://www.sciencedirect.com/science/article/pii/S0092867408005175>
290. Calebiro D, Nikolaev VO, Gagliani MC, de Filippis T, Dees C, Tacchetti C, et al. Persistent cAMP-Signals Triggered by Internalized G-Protein–Coupled Receptors. *PLOS Biol* [Internet]. 2009 Aug 18;7(8):e1000172. Available from: <https://doi.org/10.1371/journal.pbio.1000172>
291. Chandran D, Bergmann FT, Sauro HM. TinkerCell: modular CAD tool for synthetic biology. *J Biol Eng*. 2009 Oct;3:19.

292. Renicke C, Schuster D, Usherenko S, Essen L-O, Taxis C. A LOV2 Domain-Based Optogenetic Tool to Control Protein Degradation and Cellular Function. *Chem Biol* [Internet]. 2013;20(4):619–26. Available from: <https://www.sciencedirect.com/science/article/pii/S1074552113000884>
293. Chandran D, Bergmann FT, Sauro HM. Computer-aided design of biological circuits using TinkerCell. *Bioeng Bugs*. 2010;1(4):276–83.
294. Collier N. RePast : An Extensible Framework for Agent Simulation. *Univ Chicagos Soc Sci Res*. 2003;36:371–5.
295. Izaguirre JA, Chaturvedi R, Huang C, Cickovski T, Coffland J, Thomas G, et al. CompuCell, a multi-model framework for simulation of morphogenesis. *Bioinformatics*. 2004;20(7):1129–37.
296. Mirams GR, Arthurs CJ, Bernabeu MO, Bordas R, Cooper J, Corrias A, et al. Chaste: An Open Source C++ Library for Computational Physiology and Biology. *PLoS Comput Biol*. 2013;9(3).
297. Chaudhary SU, Shin S-Y, Lee D, Song J-H, Cho K-H. ELECANS--an integrated model development environment for multiscale cancer systems biology. *Bioinformatics*. 2013;29(7):957–9.
298. Collier N, Howe TR. Proceedings of the first annual North American Association for Computational Social and Organizational Science conference. 2003;
299. North MJ, Collier NT, Ozik J, Tatara ER, Macal CM, Bragen M, et al. Complex adaptive systems modeling with Repast Symphony. *Complex Adapt Syst Model*. 2013;1(1):3.
300. Folcik VA, An GC, Orosz CG. The Basic Immune Simulator: An agent-based model to study the interactions between innate and adaptive immunity. *Theor Biol Med Model* [Internet]. 2007;4(1):39. Available from: <https://doi.org/10.1186/1742-4682-4-39>
301. Mehdizadeh H, Sumo S, Bayrak ES, Brey EM, Cinar A. Three-dimensional modeling of angiogenesis in porous biomaterial scaffolds. *Biomaterials* [Internet]. 2013;34(12):2875–87. Available from: <https://www.sciencedirect.com/science/article/pii/S0142961213000288>
302. Mahoney AW, Smith BG, Flann NS, Podgorski GJ. Discovering novel cancer therapies: A computational modeling and search approach. In: 2008 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology. 2008. p. 233–40.
303. Mahoney AW, Podgorski GJ, Flann NS. Multiobjective Optimization Based-Approach for Discovering Novel Cancer Therapies. *IEEE/ACM Trans Comput Biol Bioinforma*. 2012;9(1):169–84.
304. Cho K-H, Chaudhary SU, Sormark JE, Won JK, Shin SY. Multi-scale Modeling and Game-Theoretic Analysis of Mitochondrial Process Elucidates the Hidden Mechanism of

- Warburg Effect in Tumorigenesis. In: 11th Int Conf on Systems Biology (ICSB2010). 2010.
305. Ghaffarizadeh A, Heiland R, Friedman SH, Mumenthaler SM, Macklin P. PhysiCell: An open source physics-based cell simulator for 3-D multicellular systems. PLoS Comput Biol. 2018 Feb;14(2):e1005991.
 306. Letort G, Montagud A, Stoll G, Heiland R, Barillot E, MacKlin P, et al. PhysiBoSS: A multi-scale agent-based modelling framework integrating physical dimension and cell signalling. Bioinformatics. 2019;
 307. Colin A, Letort G, Razin N, Almonacid M, Ahmed W, Betz T, et al. Active diffusion in oocytes nonspecifically centers large objects during prophase I and meiosis I. J Cell Biol [Internet]. 2020 Jan 17;219(3). Available from: <https://doi.org/10.1083/jcb.201908195>
 308. Getz M, Wang Y, An G, Becker A, Cockrell C, Collier N. Rapid community-driven development of a SARS- CoV-2 tissue simulator. 2020;1–55.
 309. Gondal MN, Sultan MU, Arif A, Rehman A, Awan HA, Arshad Z, et al. TISON: a next-generation multi-scale modeling theatre for in silico systems oncology. bioRxiv [Internet]. 2021 Jan 1;2021.05.04.442539. Available from: <http://biorxiv.org/content/early/2021/05/05/2021.05.04.442539.abstract>
 310. Stamatakis GS, Dionysiou D. 6th International Advanced Research Workshop on In Silico Oncology and Cancer Investigation–The CHIC Project Workshop (IARWISOCI).
 311. Stamatakis GS, Dionysiou DD. 5th International Advanced Research Workshop on In Silico Oncology and Cancer Investigation–The TUMOR Project Workshop (IARWISOCI).
 312. Martin L, Anguita A, Graf N, Tsiknakis M, Brochhausen M, Rüping S, et al. ACGT: advancing clinico-genomic trials on cancer–four years of experience. In: User Centred Networked Health Care. IOS Press; 2011. p. 734–8.
 313. Marias K, Sakkalis V, Roniotis A, Farmaki C, Stamatakis G, Dionysiou D, et al. Clinically Oriented Translational Cancer Multilevel Modeling: The ContraCancrum Project. In Springer, Berlin, Heidelberg; 2009. p. 2124–7.
 314. Rossi S, Christ-Neumann M, Rüping S, Buffa F, Wegener D, McVie G, et al. p-Medicine: From data sharing and integration via VPH models to personalized medicine. Ecanermedicalscience. 2011;5:218.
 315. Sakkalis V, Marias K, Mckeever S, Graf N. The TUMOR project: Integrating cancer model repositories for supporting predictive oncology Collaborative Open Plant Omics-e-infrastructure for plant sciences View project PhenoMeNal-Large Scale Computing for Medical Metabolomics View project SEE PROFILE. 2014;
 316. Stamatakis G, Dionysiou D, Misichroni F, Graf N, Gool S van, Bohle R, et al. Computational horizons in cancer (CHIC): Developing meta- and hyper-multiscale models and repositories for in Silico Oncology - A brief technical outline of the project. In: Proceedings of the 2014 6th International Advanced Research Workshop on In Silico

- Oncology and Cancer Investigation - The CHIC Project Workshop (IARWISOCI). 2014. p. 1–5.
317. Karlsson J, Garcia M, Martín-Requena V, Trelles O. Web-Service Based Analysis of Gene-expression Data for Cancer Patients. This Vol is Dedic to cancer patients. :17.
318. Sfakianakis S, Graf N, Hoppe A, Rüping S, Wegener D, Koumakis L. Building a system for advancing clinico-genomic trials on cancer. CEUR Workshop Proc. 2009;475(February 2014):36–47.
319. Stamatakis G, Dionysiou D, Georgiadi E, Kolokotroni E, Giatili S, Hoppe A, et al. The ACGT Oncosimulator: from conceptualization to development via multiscale cancer modeling. This Vol is Dedic to cancer patients. 2010;54.
320. Kolokotroni EA, Stamatakis GS, Dionysiou DD, Georgiadi EC, Desmedt C, Graf NM. Translating multiscale cancer models into clinical trials: Simulating breast cancer tumor dynamics within the framework of the “Trial of Principle” clinical trial and the ACGT project. In: 2008 8th IEEE International Conference on Bioinformatics and BioEngineering. 2008. p. 1–8.
321. Durbecq V, Paesmans M, Cardoso F, Desmedt C, Di Leo A, Chan S, et al. Topoisomerase-II α expression as a predictive marker in a population of advanced breast cancer patients randomly treated either with single-agent doxorubicin or single-agent docetaxel. Mol Cancer Ther. 2004;3(10):1207–14.
322. Georgiadi EC, Stamatakis GS, Graf NM, Kolokotroni EA, Dionysiou DD, Hoppe A, et al. Multilevel cancer modeling in the clinical environment: Simulating the behavior of Wilms tumor in the context of the SIOP 2001/GPOH clinical trial and the ACGT project. In: 2008 8th IEEE International Conference on Bioinformatics and BioEngineering. IEEE; 2008. p. 1–8.
323. Graf N, Hoppe A, Georgiadi E, Belleman R, Desmedt C, Dionysiou D, et al. ‘In silico’ oncology for clinical decision making in the context of nephroblastoma. Klin Paediatr. 2009;221(03):141–9.
324. Zasada SJ, Wang T, Haidar A, Liu E, Graf N, Clapworthy G, et al. IMENSE: An e-infrastructure environment for patient specific multiscale data integration, modelling and clinical treatment. J Comput Sci [Internet]. 2012;3(5):314–27. Available from: <https://www.sciencedirect.com/science/article/pii/S1877750311000639>
325. Marias K, Dionysiou D, Sakkalis V, Graf N, Bohle RM, Coveney P V, et al. Clinically driven design of multi-scale cancer models: the ContraCancrum project paradigm. Interface Focus [Internet]. 2011 Jun 6;1(3):450–61. Available from: <https://doi.org/10.1098/rsfs.2010.0037>
326. Giatili SG, Uzunoglu NK, Stamatakis GS. An Explicit Boundary Condition Treatment of a Diffusion–Based Glioblastoma Tumor Growth Model. This Vol is Dedic to cancer patients. 2010;86.

327. Stamatakos GS. In Silico Oncology: A Hypermatrix–Operator Formulation of a Top-Down Multiscale Simulation Model of Tumor Response to Treatment. The Oncosimulator Concept. 2010;(May 2014).
328. Li XL, Oduola WO, Qian L, Dougherty ER. Integrating Multiscale Modeling with Drug Effects for Cancer Treatment. *Cancer Inform.* 2015;14(Suppl 5):21–31.
329. Stamatakos GS, Georgiadi EC, Graf N, Kolokotroni EA, Dionysiou DD. Exploiting Clinical Trial Data Drastically Narrows the Window of Possible Solutions to the Problem of Clinical Adaptation of a Multiscale Cancer Model. *PLoS One* [Internet]. 2011 Mar 3;6(3):e17594. Available from: <https://doi.org/10.1371/journal.pone.0017594>
330. Wan S, Coveney P V. Rapid and accurate ranking of binding affinities of epidermal growth factor receptor sequences with selected lung cancer drugs. *J R Soc Interface.* 2011;8(61):1114–27.
331. Wang Z, Deisboeck TS. Application of ANOVA-Based Global Sensitivity Analysis to a Multiscale Cancer Model. *This Vol is Dedic to cancer patients.* :27.
332. Roniotis A, Panourgias K, Ekaterinaris J, Marias K, Sakkalis V. Approximating the diffusion–reaction equation for developing glioma models for the ContraCancrum Project: a showcase. *This Vol is Dedic to cancer patients.* :80.
333. Folarin AA, Stamatakos GS. Molecular Personalization of Cancer Treatment via a Multiscale Simulation Model of Tumor Response to Therapy. The Paradigm of Glioblastoma Treated with Temozolomide. *This Vol is Dedic to cancer patients.* 2010;23.
334. Schera F, Weiler G, Neri E, Kiefer S, Graf N. The p-medicine portal—a collaboration platform for research in personalised medicine. *Ecanermedicalscience.* 2014;8.
335. Georgiadi EC, Dionysiou DD, Graf N, Stamatakos GS. Towards in silico oncology: adapting a four dimensional nephroblastoma treatment model to a clinical trial case based on multi-method sensitivity analysis. *Comput Biol Med.* 2012;42(11):1064–78.
336. Blazewicz M, Georgiadi EC, Pukacki J, Stamatakos GS. Development of the p-medicine oncosimulator as a parallel treatment support system. In: *Proceedings of the 2014 6th International Advanced Research Workshop on In Silico Oncology and Cancer Investigation - The CHIC Project Workshop (IARWISOCI).* 2014. p. 1–5.
337. Argyri KD, Dionysiou DD, Misichroni FD, Stamatakos GS. Numerical simulation of vascular tumour growth under antiangiogenic treatment: addressing the paradigm of single-agent bevacizumab therapy with the use of experimental data. *Biol Direct.* 2016;11(1):1–31.
338. Stamatakos G, Dionysiou D, Lunzer A, Belleman R, Kolokotroni E, Georgiadi E, et al. The technologically integrated oncosimulator: Combining multiscale cancer modeling with information technology in the in silico oncology context. *IEEE J Biomed Heal Informatics.* 2014;18(3):840–54.
339. Ouzounoglou EN, Dionysiou DD, Stanulla M, Stamatakos GS. Towards patient

- personalization of an Acute Lymphoblastic Leukemia Model during the oral administration of prednisone in children: Initiating the ALL Oncosimulator. In: Proceedings of the 2012 5th International Advanced Research Workshop on In Silico Oncology and Cancer Investigation. 2012. p. 1–4.
340. Kolokotroni E, Ouzounoglou E, Systems C, Stanulla M, Dionysiou DD, Systems C. In Silico Oncology : Developing and Clinically Adapting the Acute Lymphoblastic Leukemia (ALL) Oncosimulator by Exploiting Pathway Based Gene Expression Analysis in the Context of t ... In Silico Oncology : Developing and Clinically Adapting the Acute L. 2015;(January 2014).
 341. Ouzounoglou E, Kolokotroni E, Stanulla M, Stamatakis GS. A study on the predictability of acute lymphoblastic leukaemia response to treatment using a hybrid oncosimulator. Interface Focus. 2018;8(1):20160163.
 342. Bucur A, van Leeuwen J, Christodoulou N, Sigdel K, Argyri K, Koumakis L, et al. Workflow-driven clinical decision support for personalized oncology. BMC Med Inform Decis Mak. 2016;16(2):87.
 343. Deisboeck TS, Zhang L, Martin S. Advancing cancer systems biology: Introducing the Center for the Development of a Virtual Tumor, CViT. Cancer Inform. 2007;5:1–8.
 344. Tomita M, Shimizu T, Saito K, Venter JC, Hashimoto K, Matsuzaki Y, et al. E-CELL: Software environment for whole cell simulation. Genome Informatics. 1997;8:147–55.
 345. Hedley WJ, Nelson MR, Bellivant DP, Nielsen PF. A short introduction to CellML. Philos Trans R Soc London Ser A Math Phys Eng Sci. 2001;359(1783):1073–89.
 346. Christie GR, Nielsen PMF, Blackett SA, Bradley CP, Hunter PJ. FieldML: concepts and implementation. Philos Trans R Soc A Math Phys Eng Sci. 2009;367(1895):1869–84.
 347. Johnson D, Cooper J, McKeever S. TumorML: Concept and requirements of an in silico cancer modelling markup language. In: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE; 2011. p. 441–4.
 348. Wang Z, Zhang L, Sagotsky J, Deisboeck TS. Simulating non-small cell lung cancer with a multiscale agent-based model. Theor Biol Med Model. 2007;4(1):1–14.
 349. Johnson D, McKeever S, Deisboeck TS, Wang Z. Connecting digital cancer model repositories with markup: introducing TumorML version 1.0. ACM SIGBioinformatics Rec. 2013;3(3):5–11.
 350. Sakalis V, Sfakianakis S, Tzamali E, Marias K, Stamatakis G, Misichroni F, et al. Web-Based Workflow Planning Platform Supporting the Design and Execution of Complex Multiscale Cancer Models. IEEE J Biomed Heal Informatics. 2014;18(3):824–31.
 351. Athale C, Mansury Y, Deisboeck TS. Simulating the impact of a molecular ‘decision-process’ on cellular phenotype and multicellular patterns in brain tumors. J Theor Biol. 2005;233(4):469–81.

352. Shlomi T, Benyamini T, Gottlieb E, Sharan R, Ruppin E. Genome-Scale Metabolic Modeling Elucidates the Role of Proliferative Adaptation in Causing the Warburg Effect. *PLOS Comput Biol*. 2011 Mar;7(3):e1002018.
353. Stamatakis GS, Kolokotroni E, Dionysiou D, Veith C, Kim Y-J, Franz A, et al. In silico oncology: exploiting clinical studies to clinically adapt and validate multiscale oncosimulators. In: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE; 2013. p. 5545–9.
354. Kolokotroni E, Dionysiou D, Veith C, Kim Y-J, Sabczynski J, Franz A, et al. In Silico Oncology: Quantification of the In Vivo Antitumor Efficacy of Cisplatin-Based Doublet Therapy in Non-Small Cell Lung Cancer (NSCLC) through a Multiscale Mechanistic Model. *PLOS Comput Biol* [Internet]. 2016 Sep 22;12(9):e1005093. Available from: <https://doi.org/10.1371/journal.pcbi.1005093>
355. Antonopoulos M, Stamatakis G. In Silico Neuro-Oncology: Brownian Motion-Based Mathematical Treatment as a Potential Platform for Modeling the Infiltration of Glioma Cells into Normal Brain Tissue. *Cancer Inform*. 2015;14(Suppl 4):33–40.
356. Ouzounoglou E, Dionysiou D, Stamatakis GS. Differentiation resistance through altered retinoblastoma protein function in acute lymphoblastic leukemia: in silico modeling of the deregulations in the G1/S restriction point pathway. *BMC Syst Biol*. 2016;10(1):23.
357. Stamatakis GS, Giatili SG. A numerical handling of the boundary conditions imposed by the skull on an inhomogeneous diffusion-reaction model of glioblastoma invasion into the brain: Clinical validation aspects. *Cancer Inform*. 2017;16.
358. Stamatakis G, Graf N. Computational horizons in cancer (chic). *Clin Ther*. 2017;39(8):e107–8.
359. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue Expression (GTEx) project. Vol. 45, *Nature Genetics*. Nature Publishing Group; 2013. p. 580–5.
360. Lee HJ, Palm J, Grimes SM, Ji HP. The Cancer Genome Atlas Clinical Explorer: A web and mobile interface for identifying clinical-genomic driver associations. *Genome Med* [Internet]. 2015;7(1):1–14. Available from: <http://dx.doi.org/10.1186/s13073-015-0226-3>
361. Wang Z, Jensen MA, Zenklusen JC. A practical guide to The Cancer Genome Atlas (TCGA). In: *Methods in Molecular Biology*. Humana Press Inc.; 2016. p. 111–41.
362. Shaul YD, Yuan B, Thiru P, Nutter-Upham A, McCallum S, Lanzkron C, et al. MERAV: A tool for comparing gene expression across human tissues and cell types. *Nucleic Acids Res*. 2016;44(D1):D560–6.
363. Jensen MA, Ferretti V, Grossman RL, Staudt LM. The NCI Genomic Data Commons as an engine for precision medicine. *Blood*. 2017;130(4):453–9.
364. Li J, Lu Y, Akbani R, Ju Z, Roebuck PL, Liu W, et al. TCPA: A resource for cancer functional

- proteomics data. Vol. 10, Nature Methods. Nature Publishing Group; 2013. p. 1046–7.
365. Pontén F, Jirström K, Uhlen M. The Human Protein Atlas—a tool for pathology. J Pathol. 2008 Dec;216(4):387–93.

A Complex Multi-scale Biomolecular Regulation in Cancer

E Environment

Databases (ST 4)

- ExoCarta
- Evpedia
- MatrixDB
- MatrisomeDB

- Vesiclepedia
- CancerPDF
- HCSD
- SPD

B Biomolecules Involved in Cancer

Databases (ST 1)

Genome

- GenBank

Transcriptome

- GEO
- TCGA
- ICGC
- cBioPortal

Proteome

- HPP
- HPA

Metabolome

- HMDB
- BMRB

C Pathways

Databases (ST 2)

- PathDB
- KEGG
- REACTOME
- PC
- InnateDB
- PANTHER

- GTDB
- SMPDB
- PAGED
- Pathguide
- NetPath
- PATRIC

- PDS
- MetaCyc
- CST
- WikiPathways

D Networks

Databases (ST 3)

- BioGRID
- STRING
- DIP
- BIND
- IntAct
- GeneNet

- HumanNet
- BioModels
- HPRD
- CPDB
- EHMN
- MINT

F Cell Lines for Culture

Databases (ST 5)

- CCLE
- CLDB
- Cellosaurus
- CLO
- COSMIC
- ICLAC
- hPSCreg

- IGRhCellID
- CLIFF
- GEMiCCL
- CLASTR

G Histopathological Imaging

Databases (ST 6)

- TCIA
- GTEx
- TMAD
- CDSA
- Histomics ML
- TCGA-

- READ
- BreakHis
- HICL
- BreCaHAD
- Pathology PhotoStore
- MRF-ANN

- Camelyon17

H Mutation & Drugs

Databases (ST 7)

- DrugBank
- KEGG Drug
- OncoKB
- TTD
- OMIM
- PDTT
- PharmGKB

- STITCH
- MATADOR
- GDSC
- DGIdb
- PharmacOD
- PanDrugs

Scales		Network				Environment			Cell Line			Multi-scale							
Software		CellDesigner	COPIASI	CellNetAnalyzer	BoolNet	GIMsim	CellWare	Morpheus	BioFVM	E-Cell	Virtual Cell	ThinkCell	Compucell	CHASTE	ELECANS	PhysiCell	PhysiBoSS	TISON	
Year		2003	2006	2008	2010	2012	2003	2014	2015	1999	2008	2009	2004	2013	2013	2018	2019	2021	
Features List	Boolean/Qualitative Networks			●	●	●											●	●	
	Boolean/Quantitative Networks				●													●	
	Quantitative Networks	●	●				●	●		●	●								
	Stochastic Semantics		●		●					●	●		●				●	●	
	ODE		●				●	●			●	●	●	●				●	
	Agent-based Modeling												●	●	●	●	●		
	2D Tissue							●		●	●		●	●	●	●	●	●	
	3D Tissue								●	●	●	●	●	●	●	●	●	●	
	Biomolecular Diffusion Continuous									●	●		●	●				●	
	Biomolecular Diffusion Discrete							●	●	●	●			●	●	●	●	●	
	Biomolecular Diffusion Hybrid						●				●							●	
	Parallelized		●	●	●			●	●		●	●		●				●	●
	GUI	●	●	●		●	●	●	●	●	●	●	●	●		●		●	●
	Web-based Platform	●	●			●		●		●	●	●							●
	Language	Java	C++	MATLAB	R	Java	Java	C++	C++	C++	C++	C++/Java	Python/C	C++	C#	C++	C++	C++	C#
Reference PMID's	doi:10.1016/S1478-5382(03)02370-9	17032683	17408509	20378558	22144167	14871872	24443380	26656933	10068694	19045830	19874625	14764549	23516352	23390137	29474446	30169736	doi:10.1101/2021.05.04.442539		

Platform Adaptability

User's Citations

- CHASTE
- CompuCell3D
- Repast
- VCe
- CellSys
- TinkerCell
- SemSim
- RuleMonkey
- OpenCell
- NetLogo
- ELECANs

Supplementary Table 11