



Models of Models: A Translational Route for Cancer Treatment and Drug Development

Lesley A. Ogilvie¹, Aleksandra Kovachev¹, Christoph Wierling¹, Bodo M. H. Lange¹ and Hans Lehrach^{1,2*}

¹Alacris Theranostics GmbH, Berlin, Germany, ²Max Planck Institute for Molecular Genetics, Berlin, Germany

OPEN ACCESS

Edited by:

Michael Breitenbach,
University of Salzburg, Austria

Reviewed by:

Michael Gerhard Löffler,
University of Salzburg, Austria
Hans-Werner Mewes,
Technische Universität München,
Germany

*Correspondence:

Hans Lehrach
lehrach@molgen.mpg.de

Specialty section:

This article was submitted to
Molecular and Cellular Oncology,
a section of the journal
Frontiers in Oncology

Received: 26 June 2017

Accepted: 01 September 2017

Published: 19 September 2017

Citation:

Ogilvie LA, Kovachev A, Wierling C,
Lange BMH and Lehrach H (2017)
Models of Models: A Translational
Route for Cancer Treatment and Drug
Development.
Front. Oncol. 7:219.
doi: 10.3389/fonc.2017.00219

Every patient and every disease is different. Each patient therefore requires a personalized treatment approach. For technical reasons, a personalized approach is feasible for treatment strategies such as surgery, but not for drug-based therapy or drug development. The development of individual mechanistic models of the disease process in every patient offers the possibility of attaining truly personalized drug-based therapy and prevention. The concept of virtual clinical trials and the integrated use of *in silico*, *in vitro*, and *in vivo* models in preclinical development could lead to significant gains in efficiency and order of magnitude increases in the cost effectiveness of drug development and approval. We have developed mechanistic computational models of large-scale cellular signal transduction networks for prediction of drug effects and functional responses, based on patient-specific multi-level omics profiles. However, a major barrier to the use of such models in a clinical and developmental context is the reliability of predictions. Here we detail how the approach of using “models of models” has the potential to impact cancer treatment and drug development. We describe the iterative refinement process that leverages the flexibility of experimental systems to generate highly dimensional data, which can be used to train and validate computational model parameters and improve model predictions. In this way, highly optimized computational models with robust predictive capacity can be generated. Such models open up a number of opportunities for cancer drug treatment and development, from enhancing the design of experimental studies, reducing costs, and improving animal welfare, to increasing the translational value of results generated.

Keywords: preclinical models, computational model, mechanistic modeling, genetically engineered mouse models, transgenic mice, model optimization

MECHANISTIC MODELS IN ONCOLOGY

Despite major breakthroughs in cancer research and therapy, the disease still remains one of the world’s major healthcare challenges. A challenge that is exacerbated in Europe due to an aging population (1) and associated increase in cancer incidence (2). Ultimately, identification of successful therapies is hampered by the high level of complexity and genetic heterogeneity existing even within single tumor types, causing a large fraction of patients to remain refractory to treatment even with the most effective drugs we have available today (3–7).

Whenever we deal with complex problems mistakes are unavoidable, often with catastrophic consequences. Computer models can be used to simulate complex situations prior to testing in reality, allowing us to make these inevitable mistakes *in silico* and helping us to successfully avoid their deleterious impacts.

Although used in many areas from the aviation industry to climate prediction, a computational modeling strategy is still not being applied within two areas that have a fundamental impact on our health, wellbeing, and even our survival: drug-based treatment and drug development. Both are areas in which we still proceed statistically, prescribing drugs that ultimately only help a very small fraction of the patients receiving them and that could have negative consequences. Adverse drug effects lead to nearly 200,000 deaths per year in Europe (8), with enormous associated economic costs (9).

To overcome this problem, we have to be able to generate sufficiently accurate models of an individual patient—a virtual self—that allow us to observe the effects of different therapies. The basis is a generic computational model with the ability to predict effects and side effects of different drugs and their combinations, which is individualized based on a detailed characterization of a patient by molecular, sensor, and other techniques. In oncology, for example, we should ideally reflect the heterogeneity of the tumor by modeling individual tumor cells, including the stroma, to determine response. In addition, aspects of the liver should also be considered to determine the pharmacogenetics of the drug. Side effects can be evaluated by incorporating a selection of normal cell types and if we want to predict the effects and side effects of immunotherapies, we should also include elements of the immune system.

Currently, we are building large-scale mechanistic computational models of the signal transduction networks in cells (or cell collectives), based on the ever-expanding biological knowledge base, e.g., on signaling pathways in human cells. For this, we are using PyBioS, currently in its third iteration (PyBioS3), an integrated software platform for the design, modeling, and simulation of cellular systems (10, 11). The current model integrates about 50 cancer-related signaling pathways and makes use of a large and growing information base on functional consequences of genetic variants and mechanistic drug action. See Ref. (12–16), for further details of the modeling system and its applications.

To provide personalized predictions, the models are typically individualized with next generation sequencing-derived omics data (e.g., genome and transcriptome) from a patient and in the case of cancer also from individual tumors. For drug response predictions, the drugs to be “screened” are regarded as molecular entities that typically affect molecular networks. This information is translated into systems of ordinary differential equations, which can be solved numerically to make predictions regarding the functional response of the system in response to perturbations, such as specific genetic variants and/or drugs and their combinations. Adaptation of the model to biological observations and experimental data calls for optimization approaches. As the models become more complex, it is becoming increasingly important to use advanced parameter estimation strategies (17–20) to fit the model to the data. This can be done based on data generated on the types of preclinical models discussed in this issue.

Cancer is suited particularly well to this type of approach, as it is fundamentally a cellular disease. In tandem, high levels of funding for cancer research in the last decades has generated much of the knowledge required to establish generic computational models, such as information on the basic mechanisms of cancer and drug action, including molecular targets [e.g., (21–25)]. Diseased tissues can also be obtained as surgical or biopsy material. This means we can actually observe the changes—often dramatic—occurring in the tumor genome and transcriptome, making it easier to understand the likely functional consequences. Last but not least, computing power is now at a level (26) that makes predictive modeling on a large scale a realistic prospect.

PREDICTING UNCERTAINTY

The use of such computational models for personalizing medicine in the clinic does, however, still face a number of challenges. One of the main barriers to routine implementation of computational models in clinical scenarios is the accuracy of the prediction. Just how reliable can predictive computational models be?

The generic mechanistic model we have created integrates major molecular species, i.e., representations of genes, proteins, protein complexes, metabolites, etc., and biochemical/cellular processes, together making up an *in silico* representation of the cellular signaling network. Furthermore, it integrates modifications of the molecular species that are associated with cancer onset and development, such as mutated genes and proteins, reflecting gain-of-function or loss-of-function of oncogenes or tumor suppressor genes. Understanding how such a complex system functions as a whole is inferred from examination of its individual parts and their interactions [e.g., see Ref. (27, 28)].

To ensure that such mechanistic models are predictive we need a detailed assessment of the most important underlying biological reactions. However, within the large-scale networks generated, much of this information, such as binding affinities, (de)phosphorylation rates and synthesis, and degradation rates, is not easily obtained experimentally. To overcome the lack of information on parameters needed for this, we originally used a Monte Carlo strategy, selecting multiple random parameter vectors for multiple solutions (12). The unknown kinetic parameters are repeatedly sampled from probability distributions of values and used in multiple parallel simulations.

As the models grow in size, representing more signaling pathways and cellular components, so does the inherent complexity of the model and the associated number of unknown parameters involved in each process (e.g., kinetic constants, component concentrations). The signal transduction model we are currently working with comprises hundreds of genes, their modifications, and associated interactions, equating to tens of thousands of parameters. To infer the unknown parameters within this growing large-scale network, parameter optimization and reverse engineering strategies are used to increase the accuracy of predictions. This essentially means using data generated in experimental systems, e.g., mouse models, organotypic cultures, and cell culture, as well as data from patients (if available), for evaluating drug effects and other functional responses on the phenotypic and molecular level. We are generating this type of data within

the scope of a number of national and international projects, such as the Horizon2020 project CanPathPro (www.canpathpro.eu); focused on the development of a combined experimental and systems biology platform for predictive modeling of cancer signaling, Treat20Plus, a German Federal Ministry of Education and Research funded project that uses a computational modeling approach to predict treatment outcome for metastatic skin cancer patients, and the recently concluded OncoTrack project [www.oncotrack.eu (15)], an IMI EU funded collaborative effort that aimed to develop and validate biomarkers for colon cancer, leveraging “virtual patient” models, and multi-level omics data to provide a personalized approach to the treatment of colon cancer. Using such data, we can train the model’s parameters and structure, and validate the predictions made in an iterative fashion.

When using models of such scale, we are faced with the problem that the number of unknown parameters significantly outnumber the datasets that can be accumulated, leading to limited identifiability of parameters. To identify model parameters, statistical methods such as Bayesian and frequentist estimation as well as global and local optimization techniques can be applied (29, 30). Partitioning of the datasets into training, validation, and test sets, i.e., cross validation, facilitates identification of optimized parameter vectors and provides an unbiased estimate of how these vectors actually perform with an independent dataset (31). However, due to the limited number of training datasets available, overfitting is likely to occur, e.g., predictions become overly influenced by “noise” in the dataset rather than the inherent trends, resulting in poor predictive performance. A number of approaches can be used to overcome overfitting, such as increasing the amount of data (real or *in silico*) and/or using regularization parameters, but with complex models these strategies may bring limited improvements. An alternative strategy is to reduce the number of parameters by simplifying the model using model reduction methods (32, 33). These approaches face the challenge of identifying simplified models that exhibit dynamics comparable with the original. Ensembles of parameters that fit the data can further reduce overfitting effects. Moreover, techniques from artificial intelligence, such as deep learning, can help to learn directly from the data in an unsupervised fashion, without even knowing the underlying model (34).

MODELS OF MODELS: THE USE OF PRECLINICAL MODELS FOR OPTIMIZING *IN SILICO* PREDICTIONS

We see preclinical experimental models, including PDXs and transgenic mice [genetically engineered mouse models (GEMMs)] as well as cell and organotypic cultures, as being an integral part of the development and optimization of mechanistic computational models with more robust predictive capacity.

In particular, the contribution of mouse models to the understanding of fundamental biological processes, cancer research, and drug development has been significant, albeit with inevitable pitfalls (35, 36). PDX models have been shown to recapitulate the major molecular features of the tumor of origin, and therefore have immense utility in translational cancer research and personalized

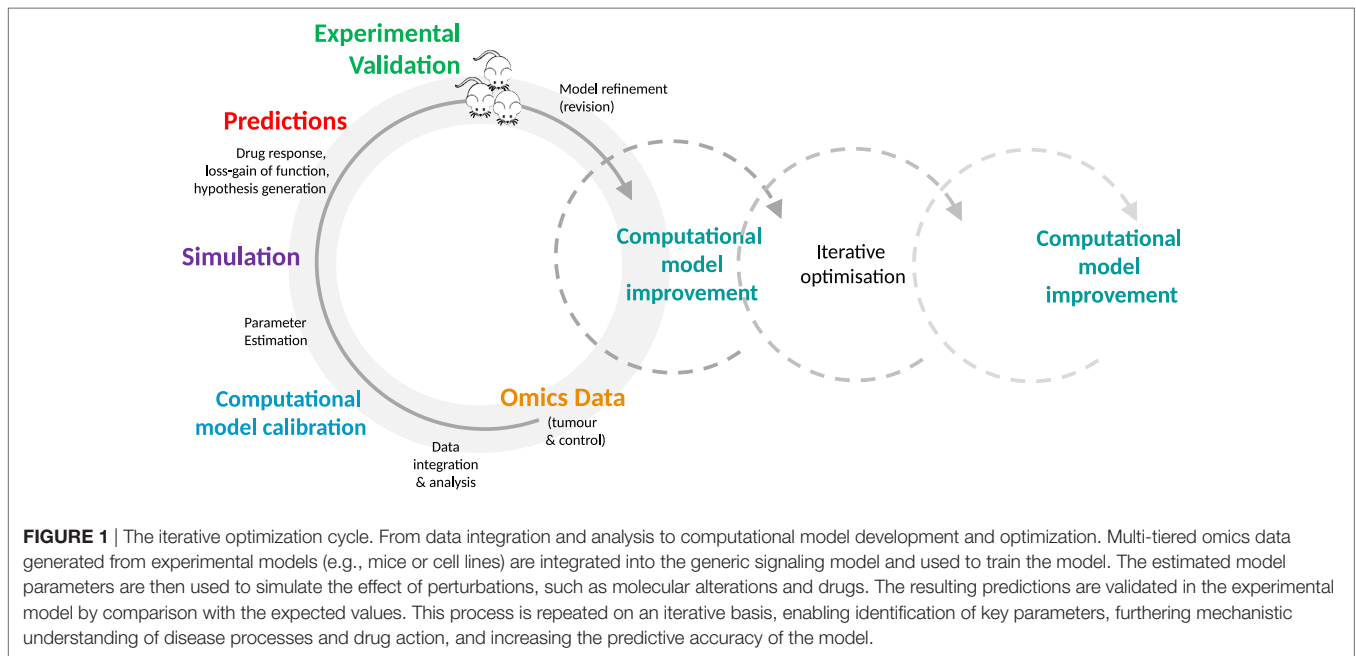
medicine applications (37, 38). Similarly, transgenic cancer mouse models, in which gene deletion or expression can be targeted in a spatial or temporal manner, are becoming an increasingly useful tool for understanding biological processes and disease development. These genetically engineered mice develop tumors *de novo*, which closely mimic both the histopathological and molecular features of human tumors, and provide an experimentally tractable *in vivo* platform for investigating disease mechanisms and determining response to therapies (36, 39).

While each preclinical system has its particular merits and pitfalls, it is clear they can provide a flexible and accurate experimental test bed for training and validating computational models. As part of a number of research projects (e.g., CanPathPro, Treat20Plus, and OncoTrack), the flexibility of preclinical systems is being leveraged to iteratively improve the accuracy of *in silico* predictions, regarding the functional consequences of molecular alterations on the signal transduction network, and the corresponding response to *in silico* drug treatment.

Even at the level of cell culture, an opportunity is provided to engage in a depth and breadth of experimentation that may not be feasible, cost and time-wise, using more complex preclinical models such as PDXs and GEMMs. Due to the simplicity and low cost of cellular systems, in-depth experimentation is made possible. A systematic comparison of predicted and observed responses of different cell models—in the case of oncology, tumor cells—to a variety of drugs and their combinations can be conducted in a quantitative and time-resolved manner. The observed phenotypic responses, e.g., does the cell respond to a drug or not, can be evaluated in the context of their specific molecular profiles. In addition, cell lines or organoids can be used for time-resolved analysis of the molecular changes (e.g., transcriptome, proteome, metabolome, and other omics-strategies) triggered by adding a drug or drugs under investigation. In-depth data are generated that is likely to be required for model optimization in high-dimensional parameter spaces. This detailed comparison of a cellular model at the phenotypic and molecular level with the predicted behavior of its computational avatar provides an essential data foundation that enables fine-scale optimization and increased predictive accuracy of *in silico* models. Computational models can be optimized in an iterative fashion through *in silico* perturbation experiments and subsequent validation of parameter information and functional response within experimental systems. See **Figure 1**, for a typical iterative workflow.

A stepwise process is taken in which the generic cellular signaling model is first adapted to the cancer being studied, with additional relevant pathways and mutations being added as modular elements [see (13, 16)]. Next, multi-tiered omics data from model samples (e.g., tumor and control samples from PDXs and transgenic mice or organotypic/tumor cell culture samples) is acquired and analyzed for alterations such as single nucleotide polymorphisms, gene fusions, and mutations. These data are integrated into the signaling network in modular format to individualize and calibrate the model.

Given that many parameters will be unknown, estimation strategies are applied [as detailed above and in Ref. (13, 16)] and the effects of perturbations (e.g., genetic variants and/or drugs), in combination with omics data (e.g., transcriptome of



the respective samples), on network components are simulated. The generated simulation data are analyzed to make predictions regarding functional effects within the network, using indicators such as Myc levels or caspase activity as a proxy for phenotypic effects (e.g., cell proliferation, apoptosis). These predictions, including drug response, can subsequently be validated in an experimental system such as a PDX or cell culture. This detailed comparison of experimental results and predictions enables us to further refine our definition of the computational parameter space used for modeling, identifying parameters that are key to generating accurate predictions and that influence the phenotype of interest.

We are taking such an iterative approach within CanPathPro (www.canpathpro.eu), an EU Horizon2020 funded project, which focuses on the development of a combined experimental and systems biology platform for predictive modeling of cancer signaling. GEMMs, GEMM-derived cell lines, and organotypic cultures are used to provide an accurate route toward mapping the functional changes (e.g., deregulated pathways, pathway modules, and expression signatures) associated with a variety of mutated or over-expressed oncogenes and tumor suppressor genes that lead to different lung or breast cancer phenotypes.

For *in silico* experiments, a generic mouse-specific computational model has been generated based on the human-specific ModCell™ model (12, 13), leveraging the conservation and homologies existing between human and mouse genes, proteins, network structure, individual signaling relevant protein–protein interactions, and post-translational events (40–45). The mouse-specific computational model is then “personalized” with multi-layered omics data (e.g., exome, transcriptome, quantitative proteome, and phosphoproteome data), from individual tumor and control mouse tissues at different disease stages. This provides the model with essential information on parameters, such as presence or absence of mutations (including their frequencies),

protein synthesis rates (derived from RNA-Seq data), and protein decay rates (e.g., derived from pulse chase experiments). Local and global optimization methods are employed to infer unknown parameters and ensembles of models with different parameter sets are used to simulate multiple hypothetical loss-/gain-of-function and under-/over-expression experiments or alternative interactions among the model components. These simulations lead to the formulation of testable hypotheses, such as determination of a specific cancer phenotype, pathway activity, and/or cross-talk. The perturbations that have the strongest effect on measurable read-out components are most likely to be used to reject or accept a hypothesis. Selected hypotheses can then be tested experimentally, first in cell lines and organotypic cultures and then in mouse models. Based on the outcome of these validation experiments, the quality and precision of the computational model predictions can be improved (see also **Figure 1** for a depiction of the iterative process). These systematic and detailed investigations will enhance the design of experiments and facilitate identification of new mechanistic interactions as well as synergistic and cross-talk effects between the cancer pathways.

Overall, the abundance of data that can be generated within preclinical systems provides a platform for validating computational model predictions, identifying the areas of the parameter space that correspond most closely to reality. More accurate identification of this space can reduce the gap between prediction and reality, improving the ability of the model to make accurate predictions, and potentially improving the translational capacity of results for the human system.

MODELS OF MODELS: ACCELERATING DRUG DEVELOPMENT

The drug development pipeline is notoriously fraught with difficulties. Preclinical models are a pivotal part of this pipeline,

bridging the translational gap from bench to bedside, however, approval rates of new drugs, especially in oncology, remain critically low—only ~5% of cancer drugs currently in Phase I trials will make it to market (46). A process with associated costs that amount to billions of dollars (47, 48). The reasons for such high attrition rates are complex and include the typically low response rate of patients to drugs, leading to the failure of late stage non-stratified clinical trials, usually after hundreds of millions to billions of dollars expenditure. Another key factor is the poor predictivity of preclinical models. Approximately 85% of drugs that have been successful in preclinical tests fail in early clinical trials, with cancer drugs making up the largest proportion of failures [reviewed in Ref. (49)]. Given that preclinical testing remains an integral part of the regulatory roadmap for drug development and approval, better approaches for improving the translational capacity of results generated are required.

In addition to virtual humans (either individual patients, or patients in large clinical trials), there is also the potential to model the large number of experimental models which are used during the preclinical phase of drug development. A highly optimized computational model opens up a range of opportunities for enhancing the design of experimental studies, thereby minimizing the number of experimental animals required, significantly reducing costs and improving animal welfare, and importantly, increasing the translational value of results generated.

The current strategy of directly extrapolating results of models to humans tends to ignore the enormous differences in the biology of models and human. It is really not very surprising that this will lead to imprecise predictions. After all, even the comparatively minor differences between different patients can cause enormous differences in response to already approved drugs (50). In a sense, we are trying to do the equivalent of directly transferring the results of a model plane in a wind tunnel to a large passenger plane, without taking into account key information on scaling effects and different aerodynamics.

To increase the translational success from experimental models to humans, we should first compare the results obtained

from experimental models to the predictions obtained by computational modeling, e.g., the effects of a drug on a computer model of the animal or cell model, adapting this computer model first. This adaptation can then be transferred to the human model of each individual patient; an equivalent strategy to that used in airplane design.

Deployment of *in silico* models at multiple stages throughout the drug development process (Figure 2) provides an opportunity to streamline the pipeline. During the early stages of drug development, *in silico* models can be deployed for selecting the most relevant drugs (and indeed models) for further development. Computational model-based knowledge gains in our understanding of the functional effects of disease-related molecular alterations could provide an effective pre-screening framework for selection of top priority candidates. By simulating the perturbation(s) occurring within the cellular transduction network, such as genetic alterations and drug treatment, *in silico* modeling has the capacity to improve understanding of disease progression and drug action. Model-based predictions of drug response/resistance and/or mode of action, based on the molecular profile of a specific experimental model, can in turn be independently validated in a preclinical experimental system using a variety of genetic manipulation techniques (e.g., RNA interference, over-expression analyses, etc.). In tandem, scope is provided to undertake experimental studies *in silico* that would not be possible due to cost and animal welfare constraints; for instance, more in-depth temporal investigations and extended screening of drug combinations.

A more robust translational route can be taken that combines the flexibility of optimized computational model predictions with experimental data, ultimately delivering benefits for patients.

OUTLOOK

Ultimately, the goal of any modeling approach is to generate results that will have a positive impact on patient outcomes and wellbeing, making it theoretically possible to identify the optimal

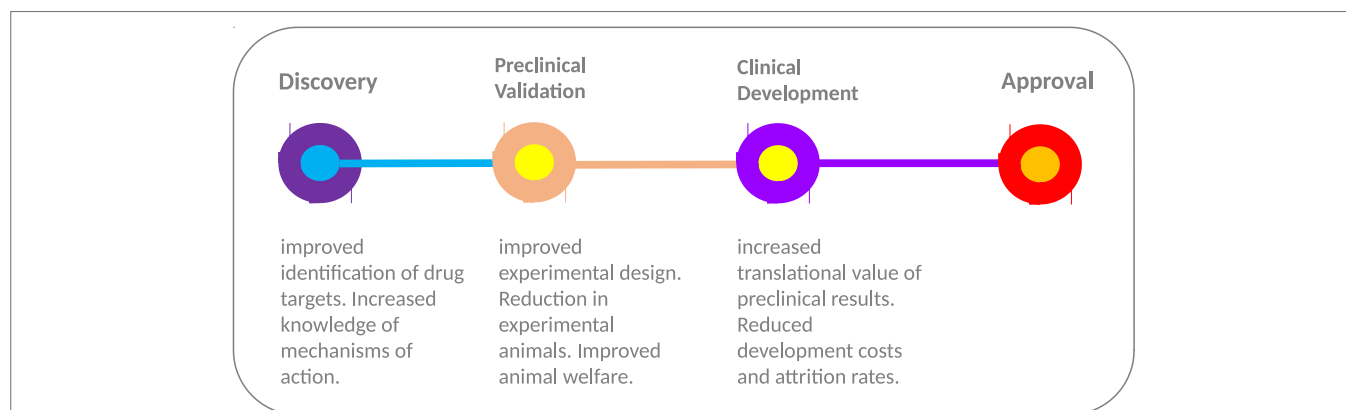


FIGURE 2 | From discovery to approval. The use of computational models throughout the drug development process provides the scope to improve experimental design and increase the translational value of early and preclinical stage results. The “models of models” approach provides a flexible test bed, enabling extended testing not feasible in animal models due to welfare and economic concerns. In combination with highly optimized computational models with more robust predictive capacity, the approach has the potential to increase the translational value of preclinical results and improve the high level of drug attrition rates, especially within the cancer arena.

therapy or preventive measure for every individual patient. Cancer models, whether experimental or computational, allow us to conduct investigations that are not possible on patients due to welfare, ethical, and economic considerations, generating information that will, in the short or long term, benefit patients. However, most model systems fail to fully recapitulate the human situation. In tandem with issues of experimental design, these inherent shortcomings will impact the translational value of models, exemplified by the high rates of attrition when preclinical findings are translated to clinical scenarios. It is also clear that existing computer models do not reflect the full complexity of the biological systems being simulated. Our own mechanistic models incorporate only a fraction of all human protein coding genes (ca. 2.5%) and do not reflect all aspects that play a role in a patient's response to drugs, such as the immune system, metabolism, and the microbial milieu associated with tumors. As our knowledge accumulates on these aspects, including the molecular mechanisms underlying interactions between tumor cells, their surrounding soma cells, and infiltrating cells of the immune system, more complex computational models can be developed that reflect the true complexity and heterogeneity of tumors. We are continually working on these components, but for now take a pragmatic approach that has enabled the generation of a generic mechanistic model of a large-scale cellular signaling network capable of predicting patient-specific responses to miRNA-based treatments (14). Within the framework of a number of preclinical and clinical studies, issues of accuracy, sensitivity, and uncertainty are being addressed, as the model is expanded, optimized, and validated.

REFERENCES

- European Commission. *The 2015 Ageing Report. Economic and Budgetary Projections for the EU 28 Member States (2013–2060)*. Report No.: 3 Directorate-General for Economic and Financial Affairs, European Economy (2015).
- Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. *GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11*. Lyon, France: International Agency for Research on Cancer (2013).
- Gerlinger M, Swanton C. How Darwinian models inform therapeutic failure initiated by clonal heterogeneity in cancer medicine. *Br J Cancer* (2010) 103(8):1139–43. doi:10.1038/sj.bjc.6605912
- Meacham CE, Morrison SJ. Tumour heterogeneity and cancer cell plasticity. *Nature* (2013) 501:328–37. doi:10.1038/nature12624
- Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* (2013) 501:338–45. doi:10.1038/nature12625
- Fisher R, Pusztai L, Swanton C. Cancer heterogeneity: implications for targeted therapeutics. *Br J Cancer* (2013) 108(3):479–85. doi:10.1038/bjc.2012.581
- Gerlinger M, Rowan AJ, Horswell A, Math M, Larkin J, Endesfelder D, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* (2016) 366(10):883–92. doi:10.1056/NEJMoa1113205
- European Commission. *Proposal for a Regulation Amending, as Regards Pharmacovigilance of Medicinal Products for Human Use. Regulation (EC) No 726/2004. Impact Assessment*. (2008). Available from: http://ec.europa.eu/health/files/pharmacovigilance/ia_vol1_en.pdf
- Sultana J, Cutroneo P, Trifiro G. Clinical and economic burden of adverse drug reactions. *J Pharmacol Pharmacother* (2013) 4:573–7. doi:10.4103/0976-500X.120957
- Wierling C, Herwig R, Lehrach H. Resources, standards and tools for systems biology. *Brief Funct Genomic Proteomic* (2007) 6(3):240–51. doi:10.1093/bfpg/elm027
- Klipp E, Liebermeister W, Wierling C, Lehrach H, Herwig R. *Systems Biology: A Textbook*. Weinheim: Wiley-VCH GmbH & Co. KGaA (2009).
- Wierling C, Kühn A, Hache H, Daskalaki A, Maschke-Dutz E, Psycheva S, et al. Prediction in the face of uncertainty: a Monte Carlo-based approach for systems biology of cancer treatment. *Mutat Res* (2012) 746(2):163–70. doi:10.1016/j.mrgentox.2012.01.005
- Wierling C, Kessler T, Ogilvie LA, Lange BM, Yaspo ML, Lehrach H. Network and systems biology: essential steps in virtualising drug discovery and development. *Drug Discov Today Technol* (2015) 15:33–40. doi:10.1016/j.ddtec.2015.07.002
- Röhr C, Kerick M, Fischer A, Kühn A, Kashofer K, Timmermann B, et al. High-throughput miRNA and mRNA sequencing of paired colorectal normal, tumor and metastasis tissues and bioinformatic modeling of miRNA-1 therapeutic applications. *PLoS One* (2013) 8(7):e67461. doi:10.1371/journal.pone.0067461
- Henderson D, Ogilvie LA, Hoyle N, Keilholz U, Lange B, Lehrach H. Personalized medicine approaches for colon cancer driven by genomics and systems biology: OncoTrack. *Biotechnol J* (2014) 9(9):1104–14. doi:10.1002/biot.201400109
- Ogilvie LA, Wierling C, Kessler T, Lehrach H, Lange BM. Predictive modeling of drug treatment in the area of personalized medicine. *Cancer Inform* (2015) 14:95–103. doi:10.4137/CIN.S1933
- Banga JR, Balsa-Canto E. Parameter estimation and optimal experimental design. *Essays Biochem* (2008) 45:195–210. doi:10.1042/BSE0450195
- Ashyraliyev M, Fomekong-Nanfack Y, Kaandorp JA, Blom JG. Systems biology: parameter estimation for biochemical models. *FEBS J* (2009) 276(4):886–902. doi:10.1111/j.1742-4658.2008.06844.x
- Cedersund G, Samuelsson O, Ball G, Tegnér J, Gomez-Cabrero D. Optimization in biology parameter estimation and the associated optimization problem. In: Geris L, Gomez-Cabrero D, editors. *Uncertainty in Biology: A Computational Modeling Approach*. Cham: Springer (2016). p. 177–97.
- Penas DR, González P, Egea JA, Doallo R, Banga JR. Parameter estimation in large-scale systems biology models: a parallel and self-adaptive cooperative strategy. *BMC Bioinformatics* (2017) 18(1):52. doi:10.1186/s12859-016-1452-4

AUTHOR CONTRIBUTIONS

LO wrote the first draft of the manuscript. LO, AK, CW, BL, and HL contributed to the writing of the manuscript. LO, AK, CW, BL, and HL jointly developed the structure and arguments for the paper. All authors reviewed and approved the final manuscript.

ACKNOWLEDGMENTS

The authors would like to thank their colleagues at Alacris Theranostics GmbH and the Dahlem Centre for Genome Research and Medical Systems Biology (DCGMS), for fruitful discussions and constructive criticism. This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 686282 (CanPathPro).

21. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* (2000) 100(1):57–70. doi:10.1016/S0092-8674(00)81683-9
22. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* (2011) 144(5):646–74. doi:10.1016/j.cell.2011.02.013
23. Santos R, Ursu O, Gaulton A, Bento AP, Donadi RS, Bologa CG, et al. A comprehensive map of molecular drug targets. *Nat Rev Drug Discov* (2017) 16(1):19–34. doi:10.1038/nrd.2016.230
24. Csermely P, Korcsmáros T, Kiss HJ, London G, Nussinov R. Structure and dynamics of molecular networks: a novel paradigm of drug discovery: a comprehensive review. *Pharmacol Ther* (2013) 138(3):333–408. doi:10.1016/j.pharmthera.2013.01.016
25. Karaman MW, Herrgard S, Treiber DK, Gallant P, Atteridge CE, Campbell B, et al. A quantitative analysis of kinase inhibitor selectivity. *Nat Biotechnol* (2008) 26(1):127–32. doi:10.1038/nbt1358
26. Processing Power Compared. Visualizing a 1 trillion-fold increase in computing performance. *Experts Exchange*. (2015). Available from: <http://pages.experts-exchange.com/processing-power-compared/>
27. Kitano H. Systems biology: a brief overview. *Science* (2002) 295(5560):1662–4. doi:10.1126/science.1069492
28. Klipp E, Liebermeister L, Wierling C, Kowald A. *Systems Biology. A Textbook*. Weinheim: Wiley-Blackwell (2016). 504 p.
29. Villaverde AF, Banga JR. Reverse engineering and identification in systems biology: strategies, perspectives and challenges. *J R Soc Interface* (2014) 11(91):20130505. doi:10.1098/rsif.2013.0505
30. Fröhlich F, Kaltenbacher B, Theis FJ, Hasenauer J. Scalable parameter estimation for genome-scale biochemical reaction networks. *PLoS Comput Biol* (2017) 13(1):e1005331. doi:10.1371/journal.pcbi.1005331
31. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. 2nd ed. New York: Springer Series in Statistics (2008). p. 241–9.
32. Henriques D, Villaverde AF, Rocha M, Saez-Rodriguez J, Banga JR. Data-driven reverse engineering of signaling pathways using ensembles of dynamic models. *PLoS Comput Biol* (2017) 13(2):e1005379. doi:10.1371/journal.pcbi.1005379
33. Rao S, Van der Schaft A, Van Eunen K, Bakker BM, Jayawardhana B. A model reduction method for biochemical reaction networks. *BMC Syst Biol* (2014) 8(1):52. doi:10.1186/1752-0509-8-52
34. Angermueller C, Pärnamaa T, Parts L, Stegle O. Deep learning for computational biology. *Mol Syst Biol* (2016) 12(7):878. doi:10.15252/msb.20156651
35. Day CP, Merlino G, Van Dyke T. Preclinical mouse cancer models: a maze of opportunities and challenges. *Cell* (2000) 163(1):39–53. doi:10.1016/j.cell.2015.08.068
36. Kersten K, de Visser KE, van Miltenburg MH, Jonkers J. Genetically engineered mouse models in oncology research and cancer medicine. *EMBO Mol Med* (2017) 9(2):137–53. doi:10.15252/emmm.201606857
37. Schütte M, Risch T, Abdavi-Azar N, Boehnke K, Schumacher D, Keil M, et al. Molecular dissection of colorectal cancer in pre-clinical models identifies biomarkers predicting sensitivity to EGFR inhibitors. *Nat Commun* (2017) 8:14262. doi:10.1038/ncomms14262
38. Morgan KM, Riedlinger GM, Rosenfeld J, Ganesan S, Pine SR. Patient-derived xenograft models of non-small cell lung cancer and their potential utility in personalized medicine. *Front Oncol* (2017) 7:2. doi:10.3389/fonc
39. van Miltenburg MH, Jonkers J. Using genetically engineered mouse models to validate candidate cancer genes and test new therapeutic approaches. *Curr Opin Genet Dev* (2012) 22(1):21–7. doi:10.1016/j.gde.2012.01.004
40. Kitano H. Biological robustness. *Nat Rev Genet* (2004) 5(11):826–37. doi:10.1038/nrg1471
41. Pires-daSilva A, Sommer RJ. The evolution of signalling pathways in animal development. *Nat Rev Genet* (2003) 4(1):39–49. doi:10.1038/nrg977
42. Cheng Y, Ma Z, Kim BH, Wu W, Cayting P, Boyle AP, et al. Principles of regulatory information conservation between mouse and human. *Nature* (2014) 515(7527):371–5. doi:10.1038/nature13985
43. Stergachis AB, Neph S, Sandstrom R, Haugen E, Reynolds AP, Zhang M, et al. Conservation of trans-acting circuitry during mammalian regulatory evolution. *Nature* (2014) 515(7527):365–70. doi:10.1038/nature13972
44. Tan CS, Bodenmiller B, Pasculescu A, Jovanovic M, Hengartner MO, Jørgensen C, et al. Comparative analysis reveals conserved protein phosphorylation networks implicated in multiple diseases. *Sci Signal* (2009) 2(81):ra39. doi:10.1126/scisignal.2000316
45. Kiel C, Aydin D, Serrano L. Association rate constants of ras-effector interactions are evolutionarily conserved. *PLoS Comput Biol* (2008) 4(12):e1000245. doi:10.1371/journal.pcbi.1000245
46. Thomas DW, Burns J, Audette J, Carroll A, Dow-Hygelund C, Hay M. *Clinical Development Success Rates 2006–2015*. Biotechnology Innovation Organization (BIO) (2016). Available from: https://www.bio.org/sites/default/files/Clinical_Development_Success_Rates_2006-2015_-_BIO,_Biomedtracker,_Amplion_2016.pdf [Accessed 30 May 2017].
47. Herper M. *The Cost of Creating a New Drug Now \$5 Billion, Pushing Big Pharma to Change*. New Jersey: Forbes, Pharma & Healthcare (2013). Available from: <https://www.forbes.com/sites/matthewherper/2013/08/11/how-the-staggering-cost-of-inventing-new-drugs-is-shaping-the-future-of-medicine/#3df7a07e13c3>
48. DiMasi JA, Grabowski HG, Hansen RW. *Innovation in the Pharmaceutical Industry: New Estimates of R&D Costs*. Boston: Tufts Center for the Study of Drug Development (2014). Available from: http://csdd.tufts.edu/news/complete_story/cost_study_press_event_webcast
49. Mak IW, Evaniew N, Ghert M. Lost in translation: animal models and clinical trials in cancer treatment. *Am J Transl Res* (2014) 6(2):114–8.
50. Schork NJ. Personalized medicine: time for one-person trials. *Nature* (2015) 520(7549):609–11. doi:10.1038/520609a

Conflict of Interest Statement: All authors are affiliated with Alacris Theranostics GmbH, which aims to develop “virtual patient” models for use in therapy choice and drug development.

The reviewer, ML, and handling editor declared their shared affiliation.

Copyright © 2017 Ogilvie, Kovachev, Wierling, Lange and Lehrach. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.