Adaptation and influential nodes of biological networks

Final report

Executive summary:

As the major result of this project, we described a novel network adaptation mechanism showing that the network core encodes already learned responses, while the network periphery is needed to acquire novel responses. In addition, we extended neuronal Hebbian learning to the molecular level showing that signaling networks of individual, non-neuronal cells learn similarly to our brain. We started the experimental validation of both theories. We developed Translocatome: a database predicting translocation of 10366 proteins, and the EntOpt network visualization Cytoscape plugin, the only network visualization method which is able to show network modules (protein complexes) visually. Investigating the *C. elegans* neuronal network (connectome), we developed a program able to determine path-dependent graphlets (3-node structural network segments along a neuronal pathway). We also developed a toolkit to predict the sign (activation or inhibition) of *C. elegans* synapses. In agreement with previous data on the sign-balance of other networks we showed a 4:1 ratio of activating connections to inhibiting ones. We developed 6 fruitful collaborations and wrote 3 key reviews papers. The work resulted in 23 papers (cumulative impact of journals: 95), 2 PhDs, 8 MSc theses and 18 research awards. From the network dynamics studies the Turbine middle-size company was established which won contests of Bayer, Johnson & Johnson and Roche and completed its 2nd investment round with 3 million EUR.

The LINK group (http://linkgroup.hu) developed several network-related databases and analysis methods between 2004 and 2015 with the support of the predecessors of this research grant. In the current project spanning between 2015 and 2020 we accomplished the following project elements, as planned in the original research proposal.

I. Databases

1. In the start of this work we have **extended our earlier SignaLink** (http://signalink.org) **database**. This signaling database was specialized to zebrafish containing 389 signaling proteins, 178 microRNAs and their 6756 interactions, and was published as SignaFish (http://signafish.org) in the journal Zebrafish (Csályi et al, 2016).

2. We have developed a program enabling the **automatic updates of our formerly developed ComPPI database** (http://comppi.linkgroup.hu), i.e. compartmentalized protein-protein interaction database and made its update.

3. **We developed the Translocatome database** (http://translocatome.linkgroup.hu, Mendik et al, 2019), which contains human proteins translocating between different cellular compartments. The database contains a highly curated "gold standard" dataset of 213 proteins and a true-negative dataset of 139 proteins using an in-house developed Manual Curation Framework. Using the XGBoost machine learning method we predicted 1133 high confidence translocating proteins. The database was published in the 2019 Nucleic Acid Research database issue.

4. We are preparing a perspectives paper on **RNA translocation**, since it turned out that this is a very ill-defined but at the same time very promising concept.

5. With the work of a 4-membered sub-group led by Nina Kunsic, our PhD student in cooperation with Reka Albert (external member of the Hungarian Academy), we **extended the dynamic ephitelial-mesenchymal transition (EMT) network** of Reka Albert (Cancer Res. 74,5963) by several micro-RNA-s and proteins, discovering partial EMT attractors and MET transition. The manuscript describing our results is ready for submission. We used this extended network to initiate a number of new projects, including learning of signaling networks (via intrinsically disordered proteins of the network, chromatin memory and microRNA-memory), compartmentalized signaling network analysis (based on our database http://comppi.linkgroup.hu/ developed formerly in frame of this grant), as well as assessment of the effects of protein translocation on EMT signaling. Extension of this network including stem cell development is in progress.

6. We started a program to **develop inter-cellular signaling networks** of heterogeneous cancer, stromal and immune cells.

7. As a key application of our research work, in cooperation with the startup company of our research group (http://turbine.ai), **we have assembled a cancer-specific, dynamic signaling network** containing all known cancer drivers and other key cancer-related signaling proteins from the COSMIC Census, IntOGen, KEGG Cancer, Netpath, Network of Cancer Genes, CancerDR and GDSC databases, as well as from former, much smaller dynamic cancer models. The Simulated Cell model of Turbine currently contains more than 1600 proteins and microRNAs and their over 5000 interactions. Turbine have specified this model to more than hundred cancer cell lines so far including their transcriptional (proteomic) and mutational profiles. The network includes all major cancer hallmarks including cancer stem cells.

II. Network analysis methods

1. **The upgrade** we developed at the start of this project **to the Cytoscape plugin of our ModuLand network community analysis program** developed in the preceding project (http://apps.cytoscape.org/apps/moduland20) was downloaded 5231 users by the end of 2019 and received only maximal (5 star) evaluations.

2. In the first year of this study we developed and published **a novel network representation theory** which includes network nodes not as dimensionless vertices but as probability distributions, and using the minimal information discrimination measure and the Kulback-Leibler divergence results in a new family of network visualization, coarse graining and ordering methods (Kovács et al, 2015). Later we developed **the EntOpt network visualization Cytoscape plugin based on this theory** (http://apps.cytoscape.org/apps/entoptlayout, Ágg et al, 2019). This program is able to visualize network modules (e.g. protein complexes), which is not shown by any other currently available network visualization tools. The plugin was downloaded by 5890 users by the end of 2019 and received only maximal (5 star) evaluations.

3. Bence Szabó (http://linkgroup.hu/benceszabo.php) and his sub-group has developed a **Cytoscape plugin of our former C++ network social dilemma game program, NetworGame** (http://linkgroup.hu/NetworGame.php). Bence identified minimal node-sets (3

to 4 nodes) whose cooperation turns large complex, real-world networks to total cooperation and organized a sub-group to utilize these results.

III. New scientific results A. Novel theories – and their applications

1. Extending and streamlining our previous concept on the plastic-rigid transitions of networks (https://arxiv.org/abs/1511.01239) we have developed a theory describing the **decision making mechanisms of complex systems** to known stimuli, which already have a previously encoded response, as opposed to novel stimuli, which require a brand new response (Csermely, 2018a; 2018b). Encoded responses often mobilize a small subset of central nodes, which are closely linked and give a "consensus" response. These responses are usually fast. On the contrary, newly developed responses often mobilize a wide range of network nodes also containing peripheral nodes. These responses usually develop slowly and represent the "wisdom of crowds" or "slow democracy". This paper received 4 (all together) 8-star recommendations by Faculty of 1000 (F1000 Prime). We used this theory to summarize the adaptation of cancer cell networks (Csermely, 2020a). We established a 5-membered sub-group led by our PhD student, Borbála Kovács examining the signaling network changes in colon adenoma and carcinoma development, as well as in the development of treatment resistance in prostate cancer. We are particularly interested in finding those network edges which are "newly appearing" in form of the increasing abundance of their signaling nodes during cancer development, especially at the network periphery as predicted by our theoretical work (Csermely, 2018a). Recently, Botond Mészáros (http://linkgroup.hu/botondmeszaros.php), a multi-awarded member of our team designed an algorithm which embeds networks to multidimensional space according to the neighboring nodes, and uses this network representation to find shortcuts between distant network regions. We currently implement and test this promising approach.

2. We developed a new theory on the **learning mechanisms of molecular networks** including the persistent ordering of intrinsically disordered proteins in signal-induced complex formation, resilience of activated signaling cascades and the development of chromatin memory – among others. The paper (which will be the cover story of the 2020 April issue of TiBS) demonstrates for the first time that **the well-known Hebbian learning rule of neuronal networks** (those neurons which are involved together in a learning process will have an increased connections strength as a result of learning) **is also valid at the molecular level in the learning processes of a large variety of signaling networks of non-neuronal single cells** (Csermely et al, 2020). We established a 4-membered sub-group led by our core-member, Donát Buszlai (http://linkgroup.hu/donatbuszlai.php) developing a computational tool to apply Hebbian learning rules to the extended epithelial mesenchymal signaling network made by Nina Kunsic in our group. We use intrinsically disordered proteins and experimentally identified cellular memory-related elements as special "learning nodes" of this network. We also started to examine the learning processes of interactomes.

3. In the first year of this project we described the **involvement of the RasGAP family in the learning processes of *C. elegans*** (Gyurkó et al, 2015). Later we developed a 4-membered sub-group led by our PhD student, Bánk Fenyves (and also involving our core-member, Zsolt Vassy) to analyze the structure of the ***C. elegans* neuronal network (worm-connectome) determining path-dependent 3-graphlets** (3-node structural network segments), especially those, which are used in forward or backward movement of the worm. In this inquiry we defined approximately 6 billion pathways for each sensory-motor neuron sets and identified – for the first time in the literature – pathway-related motifs used by specific pathways in the

worm connectome. We updated the work using the new *C. elegans* connectome published in 2019 and plan to submit it soon. In connection of this work **we developed a tool to predict the sign (activation or inhibition) of *C. elegans* synapses**. The tool shows a 4:1 ratio of activating connections to inhibiting ones which is in agreement with previously published data and the "sign-balance" observed in other networks. We have finished the manuscript of this study which will be submitted soon. We also examined the sign-balance of signaling networks and found the increase of activating contacts during cancer development.

III. New scientific results B. Collaborative works

We used our knowledge on interactomes and signaling networks in cooperation with other groups.

4. We published a paper together with colleagues from the Harvard University, Sweden and Norway, where we described that slow growing tumors may quite some times develop faster and more serious metastases being finally lethal than fast growing tumors based on both molecular dynamics and patient data (Adami et al, 2017).

5. We published two papers on the interactomes and signaling networks of cancer cells showing the importance of the network neighbors of proteins involved in cancer (i.e. mutated or differentially expressed; Korcsmáros et al, 2017), as well as the cross-talks between signaling pathways in cancer development (Módos et al, 2017).

6. We have assessed the changes of microRNA signaling networks in hypercholesterinemic myocardium together with Peter Ferdinandy and his group (Ágg et al, 2018).

7. We identified novel disease biomarkers using tightly connected groups of disease-related interactomes working together with a research group in Beijing (Sun et al, 2019).

8. In collaboration with Tibor Szarvas from the Clinics of Urology of the Semmelweis University, we started the analysis of prostate cancer cell line signaling networks resistant and permissive for therapy using their proteomic data listed already in Section III/1.

9. As already mentioned in Section I/5 we have developed a successful collaboration with Reka Albert, an international top expert of network dynamics and an external member of the Hungarian Academy. Reka spent her sabbathical in our lab.

IV. Review papers

We have published two review papers in the special issue of the Seminars in Cell and Developmental Biology on the changes of signaling networks during cancer initiation and development. In these papers we described a new model on the interaction of KRAS and YAP1/beta-catenin signaling pathways (Nussinov et al, 2016), as well as the importance of the intercellular signaling networks of heterogeneous cancer cells and surrounding stromal cells (Csermely et al, 2016). Continuing our previous work on heat shock proteins we published a summary on their molecular mechanism (Sőti and Csermely, 2017) and a review on the complex function of heat shock factor 1 (Barna et al, 2018).

V. Utilization and dissemination of the results

1. To utilize the Turbine methodology we described above, 4 members of our research group (including the PI of the current project) out of the 5 founders total established the Turbine start up company (http://turbine.ai) in 2015. Turbine developed the largest Simulated Cell model of the world based on the human signaling, metabolic networks and interactome and uses a number of novel artificial intelligence methods to simulate the complex signaling interactions in it. With these unique tools Turbine successfully predicted the effects of

millions of drug combinations for various large pharmaceutical companies shortening and guiding their pipelines. The technology is also used for patient cohort stratification and design the break of drug resistance. The startup was selected as one of the 4 winners of the 405 applicants (representing 66 countries) of Bayer's 2016 accelerator program, was the number 1 health startup of Pioneers'17 having more than 4000 participants and became the best startup of Central and Eastern Europe in 2017. The startup became finalist in contests of Johnson & Johnson and Roche, its founders became the world-wide bests of MIT and Forbes young innovators. In the last two years Turbine grew from the 5 founders to a team of more than 50, established a GmbH and a UK Ltd., is already serving three of the largest 10 pharmaceutical companies of the world in paid contracts and completed its second investment round with the value of 3 million EUR. The company started to develop its own drug candidates, which are predicted by signaling network simulations and validated in simulation-driven, targeted experiments.

2. The cumulative impact of the journals, where the 23 papers of this project were published, is 95. Based on the results of our research, the PI was elected as a regular member of the Hungarian Academy of Sciences and received the Széchenyi-Award (one of the highest Hungarian research awards) in 2019. Dániel Gyurkó and Dániel Veres received their PhD, Andrea Császár, Levente Dobronyi, Bánk Fenyves, Iván Fekete, Borbála Kovács and Péter Mendik completed their MD-theses, Sebestyén Kamp and Tibor Kőszegi defended their MSc thesis (10 theses total), Bánk Fenyves, Nina Kunsic, Borbála Kovács and Péter Mendik continued their PhD at the Semmelweis University. There are 4 MSc/MD theses (Brede N. Christensen MD, Balázs Kovács, Zsolt Nagy and Nóra Ordasi MSc-s) running in the lab currently. Gábor Szilágyi received a 1st prize of the President of the Semmelweis University for his research thesis. Peter Mendik, Klára Schulz and Dániel Veres received the Hungarian Scholarship of Excellence. Students of the group received two first and one second prizes at the Hungarian national student research conference ("OTDK") and 4 first, 3 second and 1 third prizes at the Semmelweis University student research ("TDK") conference. Eszter Arany high school student received a grand prize of the Hungarian national research conference, TUDOK. Our results were summarized in 14 and 15 lectures on international and Hungarian scientific conferences, respectively, in 13 posters, as well as in 14 Hungarian and 6 English seminars. In addition, we had 10 Hungarian public lectures on networks and published a scientific paper in Gifted Child Quarterly (Csermely 2017d), a chapter in a Prufrock Press US book (Csermely, 2017a), a paper in the journal of the Hungarian Academy, "Magyar Tudomány" (Csermely, 2017c) and two chapters in Hungarian books (Csermely, 2017b; Csermely, 2020b) on the use of social networks to foster creative, deep thinking of talented people, the scientific advances of women and the use of networks in social sciences, respectively. The current grant was extended by the grant K131458 continuing the uninterrupted support of "OTKA"/NKFI from 1992 until 2024. During these studies we also received a Scientific Excellence grant of the Semmelweis University and a targeted grant for AI development. In 2019 together with the Semmelweis University, we became part of the multi-national Network Medicine Consortium organized by Brigham & Womens Hospital/Harvard University Medical School. Our PhD student, Péter Mendik MD received one of the 2019 the innovation grants of the Semmelweis University helping this cooperation.

VI. References

1. Adami, H.-O., Csermely, P., Veres, D.V., Emilsson, L., Loberg, M., Bretthauer, M. and Kalager, M. (2017) Are rapidly growing cancers more lethal? Eur. J. Cancer. 72, 210-214, IF: 6.2

2. Ágg, B., Baranyai, T., Makkos, A., Vető, B., Faragó, N., Zvara, Á., Giricz, Z., Veres, D.V., Csermely, P., Aranyi, T., Puskás, L.G., Varga, Z.V., and Ferdinandy, P. (2018) MicroRNA interactome analysis predicts post-transcriptional regulation of ADRB2 and PPP3R1 in the hypercholesterolemic myocardium. Sci. Rep. 8, 10134, IF: 4.3
3. Ágg, B., Császár, A., Szalay-Bekő, M., Veres, D., Mizsei, R., Ferdinandy, P., Csermely, P. and Kovács, I.A. (2019) The EntOptLayout Cytoscape plug-in for the efficient visualization of major protein complexes in protein-protein interaction and signalling networks. Bioinformatics 35, 4490-4492. IF: 5.4
4. Barna, J., Csermely, P. and Vellai, T. (2018) Roles of Heat Shock Factor 1 beyond the heat shock response. Cell. Mol. Life Sci. 75, 2897-2916, IF: 5.8
5. Csályi, K., Fazekas, D., Kadlecsik, T., Türei, D., Gul, L., Horváth, B., Dezső Módos, D., Demeter, A., Pápai, N., Lenti, K., Csermely, P., Vellai, T., Korcsmáros, T. and Varga, M. (2016) SignaFish: a zebrafish-specific signaling pathway resource. Zebrafish, 13, 541-544, IF: 2.0
6. Csermely, P. (2017a) Creative deep thinking: Talent support networks fostering the discovery of unexpected simplicity and the development of autonomy, integrity and wisdom. In: Jonathan Plucker, Anne Rinn & Matt Makel (Eds.) Giftedness: Reflecting Theory in Practice, Waco TX: Prufrock Press, pp. 283-299.
7. Csermely, P. (2017b) General features of the structure and dynamics of real world networks (in Hungarian). In: Crossing borders. (Ed.: G. Tolcsvay Nagy) Gondolat press, Budapest, Hungary, pp. 209-220.
8. Csermely, P. (2017c) Scientific progress of women -- as a network phenomenon restricted from the top of hierarchy (in Hungarian). Magyar Tudomány (Hungarian Science) 178, 624-626.
9. Csermely, P. (2017d) The network concept of creativity and deep thinking. Applications to social opinion formation and talent support. Gifted Child Quarterly 61, 194-201, IF: 1.3
10. Csermely, P. (2018a) The wisdom of networks: A general adaptation and learning mechanism of complex systems. The network core triggers fast responses to known stimuli; innovations require the slow network periphery and are encoded by core-remodeling. BioEssays, 40, 1700150, IF: 4.7
11. Csermely, P (2018b) Democracy as a successful evolutionary survival strategy (in Hungarian). Magyar Tudomány (Hungarian Science) 179, 615-623 (2018)
12. Csermely, P. (2020a) Adaptation of cancer cell networks. In: Bernhard Strauss, Ingemar Ernberg, Mina Bissell (Eds.) Cancer Theory. MIT Press, in press
13. Csermely, P. (2020b) Networks -- How can we use them (and how not) and how can they learn? (In Hungarian) New aspects in the description of the Hungarian language 7. Network description of language. (Ed.: G. Tolcsvay Nagy) ELTE BTK, Department of Hungarian Language, in press
14. Csermely, P., Korcsmaros, T. and Nussinov, R. (2016) Intracellular and intercellular signaling networks in cancer initiation, development and precision anti-cancer therapy. RAS acts as contextual signaling hub. Sem. Cell Dev. Biol. 58, 55-59, IF: 6.3
15. Csermely, P., Kunsic, N., Mendik, P., Kerestély, M., Faragó, T. Veres, D.V. and Tompa, P. (2020) Learning of signaling networks: molecular mechanisms. Trends in Biochemical Sciences, in press, IF: 16.9 – cover story of the 2020 April issue of TiBS
16. Gyurkó, M. D., Csermely, P., Sőti, C. and Steták, A. (2015) Distinct roles of the RasGAP family proteins in C. elegans associative learning and memory. Scientific Reports 5, 15084, IF: 5.5
17. Korcsmaros, T., Modos, D., Bulusu, K., Fazekas, D., Kubisch, J., Brooks, J., Marczell, I., Szabo, P., Vellai, T., Csermely, P., Lenti, K. and Bender, A. (2017) Neighbours of cancer-

related proteins have key influence on pathogenesis and could increase the drug target space for anti-cancer therapies. Systems Biology and Applications 3, 2
18. Kovács, I.A. Mizsei, R. and Csermely P. (2015) A unified data representation theory for network visualization, ordering and coarse-graining. Scientific Reports, 5, 13786, IF:5.5
19. Mendik, P., Dobronyi, L., Hári, F., Kerepesi, C., Maia-Moço, L., Buszlai, D., Csermely, P., Veres, D. (2019) Translocatome: a novel resource for the analysis of protein translocation between cellular organelles. Nucleic Acid Res. 47, D495-D505, IF: 11.6
20. Módos, D., Brooks, J., Fazekas, D., Ari, E., Vellai, T., Csermely, P., Korcsmáros, T. and Lenti, K. (2016) Identification of critical paralog groups with indispensable roles in the regulation of signaling flow. Sci. Rep. 6, 38588, IF: 5.2
21. Nussinov, R. Tsai, C-J., Janga, H., Korcsmáros, T. and Csermely, P. (2016) Oncogenic KRAS signaling and YAP1/β-catenin: Similar cell cycle control in tumor initiation. Sem. Cell Dev. Biol. 58, 79-85, IF: 6.3
22. Sőti, Cs. and Csermely, P. (2017) The multiple role of the 90 kDa heat shock protein in cellular processes (in Hungarian). (Retrospection of the outstanding papers of the last 50 years series). Biochemistry (Hung.) 61, 27-38.
23. Sun, D., Ren, X., Ari, E., Korcsmaros, T., Csermely, P. and Wu, L-Y. (2019) Discovering cooperative biomarkers for heterogeneous complex disease diagnoses. Briefings in Bioinformatics 20, 89-101, IF: 8.4

Budapest, 25th January 2020

(Prof. Peter Csermely, PI)