JCI 2021 – Monday, April 19th

Online

Abstracts Talks

Organized by:

BSC : Biotechnologie et signalisation cellulaire
CAMB : Laboratoire de conception et application de molécules bioactives
LBP : Laboratoire de bioimagerie et pathologies
LIT : Laboratoire d'innovation thérapeutique
ICube : Laboratoire des sciences de l'ingénieur, de l'informatique et de l'imagerie
IGBMC : Institut de génétique et de biologie moléculaire et cellulaire
PCBIS : Plateforme de chimie biologique intégrative de Strasbourg

Identification and characterization of new anti-infective solutions from the medieval Arabic pharmacopoeia

Basma Abdallah[1,2,3], Hassan Ait Benhassou[3], Laurence Choulier[2], Pierre Fechter[1]

[1] Biotechnologie et signalisation cellulaire, CNRS - Université de Strasbourg, UMR7242
[2] Laboratoire d'Innovation Thérapeutique, CNRS - Université de Strasbourg, UMR7200
[3] Fondation MAScIR, Rabat, Maroc

Since the discovery of penicillin in 1928 by Alexander Fleming, a wide range of effective antibiotics against bacterial pathogens was synthesized. These therapeutics gave hope to fight all infectious diseases and save the lives of millions of patients worldwide. However, an excessive or inappropriate use has led to the emergence of different resistance mechanisms and consequently we are today facing a major challenge due to the growing rate of multidrug-resistant pathogens.

From ancient times until the discovery of antibiotics, Nature has provided all the essential needs for Humans for various uses and to release from sufferings. Many of the natural remedies mentioned in the ancient scientific manuscripts were described largely with reference to their composition, therapeutic activities and for which disease the remedy was prescribed. These natural products were used for treating and preventing human diseases based on plants, animals, and minerals. In this regard, the objective of our multidisciplinary project is to re-exploit the remedies used in the mineral pharmacopoeia of the Arab Middle Ages (aqrābādhīn) in order to identify new active compounds against bacterial infections and to understand the role of minerals in such remedies. Even if plants, as we could expect, were extensively used (almost 80% of all ingredients), more than 36 different minerals have been found in 4 aqrābādhīn. When it came to remedies against infections that could be applied externally, the use of metals grew to 70%. To show that these ancient remedies could still provide original concepts/solutions, we focused on a simple remedy, containing mainly metals. We have been able to attribute a role for each ingredient, which span from galenic functions, to bactericidal, and anti-inflammative properties. It appeared that with a very simple recipe, mainly composed of metals, these physicians designed a complete and synergistic remedy to combat abscesses, while restricting the toxic effect of metals to the site of infection. It is also a first example showing that different metal manufactures were evolved to improve the therapeutic potentials of metals. The knowledge acquired by these physicians could thus deserve more attention, and unexpected features, original organo-metallic compounds or therapeutic synergy could still be found from such research. Nevertheless, as these remedies may contain toxic compounds, or compounds with weak activity, we are further looking for strategies to combine these ancient anti-infective solutions with more modern vectorization tools (nucleic acid aptamers) to circumvent the adverse effects

Search for new anti-cancer epi-drugs by a multi-disciplinary approach

Stefano Ciaco[1, 2], Yves Mély[1], Mattia Mori[2]

[1] *Laboratoire de Bioimagerie et Pathologies, CNRS - Université de Strasbourg, UMR7021, 74 route du Rhin 67400 Illkirch-Graffenstaden*
[2] *University of Siena, Department of Biotechnology Chemistry and Pharmacy*

Epigenetics is the study of heritable phenotype changes which do not involve alterations in the DNA sequence. One of the most studied epigenetic alterations is the methylation of the cytosines. Abnormal gain of DNA methylation in promoter regions of tumor suppressor genes (TSGs) plays a key role in gene silencing and transcriptional repression which is tightly associated to cancer initiation and development.

For a faithful transmission of the methylation pattern, DNA methyltransferase 1 (DNMT1) is coordinated by a macromolecular protein complex in which the Ubiquitin-like, containing PHD and RING fingers domains 1 (UHRF1) protein is an essential epigenetic regulator. During cell replication, UHRF1, through its SRA sub-domain, binds specifically the hemi-methylated (HM) DNA and flips out the methylated cytosine (5mC) from the DNA helix and subsequently recruits DNMT1 to methylate the cytosine on the opposite daughter strand.

In a previous work, through a collaboration between the University of Strasbourg and the University of Siena, a compound (UM63) able to inhibit the base flipping activity of SRA was disclosed and characterized. In this context, we moved forward combining virtual screening and molecular modelling with biophysical assays in solution and cells with a multidisciplinary strategy in order to overcome the limitations of UM63, such as carcinogenicy and interaction with the duplex. Moreover, strengthening the collaboration thought a cotutelle agreement, a multidisciplinary strategy to further optimize the potency, specificity and physicochemical properties of UM63 was established.

A few compounds bearing different scaffolds such as anthracene, naphthalene or benzoic acid scaffold, able to bind the 5mC binding pocket and inhibit the base-flipping process in the low micromolar range were identified, also showing a limited interaction with the DNA duplex. We have also established the low toxicity of the best two compounds, thus becoming valuable starting points for structure relationship activity (SRA) and hit-to-led optimization.

Multi-omics comparisons of different forms of centronuclear myopathies and the effects of several therapeutic strategies

Sarah Djeddi[1], David Reiss[1], Alexia Menuet[1], Sébastien Freismuth[1], Juliana de Carvalho Neves[1], Sarah Djerroud[1], Xènia Massana-Muñoz[1], Anne-Sophie Sosson[1], Christine Kretz[1], Wolfgang Raffelsberger[1], Céline Keime[1], Olivier M. Dorchies[2], Julie Thompson[3], Jocelyn Laporte[1]

[1] *Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS - Université de Strasbourg, UMR7104, INSERM U1258, 67404 Illkirch, France*
[2] *Pharmaceutical Biochemistry, Institute of Pharmaceutical Sciences of Western Switzerland (ISPSO), University of Geneva, Geneva, Switzerland*
[3] *Complex Systems and Translational Bioinformatics (CSTB), ICube laboratory – CNRS - Université de Strasbourg, Fédération de Médecine Translationnelle de Strasbourg (FMTS), Strasbourg, France*

Myopathies are progressive muscle disorders affecting children and adults in all populations and represent a significant burden for the patients and families. Centronuclear myopathies (CNM) are characterized by general muscle weakness and abnormal nuclear centralization in muscle fibers. They are caused by mutations in *MTM1*, *DNM2* or *BIN1*, encoding proteins implicated in membrane trafficking and organelle positioning. MTM1 and BIN1 are negative regulators of DNM2, and the downregulation of DNM2 in mice was shown to be an efficient pre-clinical therapy for all three CNM forms. The implicated pathways are however only partially known.

In order to investigate the molecular mechanisms and to decipher a common disease signature we analyzed the muscle transcriptome of treated and untreated MTM1, DNM2 and BIN1 mouse models. The treatments encompass a pharmacological approach using tamoxifen, an antisense oligonucleotide treatment to reduce the level of DNM2, or the modulation of DNM2 or BIN1 through genetic crosses. Data analysis uncovered a total of 25494 genes including 155 dysregulated genes shared by the three murine CNM forms, and primarily associated with muscle contraction, regeneration and inflammation. Tamoxifen treatment had almost no effect on gene expression, while DNM2 downregulation normalized 40% and BIN1 overexpression 96% of the dysregulated genes. The common therapy signature of the DNM2 downregulation and BIN1 overexpression revealed 42 essential genes for the rescue of the muscle phenotype. Overall, this work highlighted disease and therapy signatures, disclosed potential novel therapeutic targets. It also enabled the identification of biomarkers detectable in muscle and/or plasma and validated through RNA quantification, western blotting and ELISA assays that might be useful to follow the efficacy of prospective therapeutic approaches.

Fluovial: A Project to Detect and Identify Fluorinated Pollutants

Laura Duciel[1, 2], Marc-André Delsuc [2]

[1] *CASC4DE*
[2] *Institut de Génétique et de Biologie Moléculaire et Cellulaire, IGBMC CNRS - Université de Strasbourg, UMR7104, Institut National de la Santé et de la Recherche Médicale, U1258*

Fluor is a common element not metabolized in living organisms, present in perfluorinated polymers, some phytosanitary products, numerous drugs and many manufactured objects. Fluorinated molecules are extremely resistant and routinely used in numerous products synthesis, they are part of environment pollutants and specifically known as POP1 (Persistent Organic Pollutants). An example of these molecule is PFOS (PerFluoroOctaneSulfonic acid), a fluorinated surfactant used in water repellents products and well-known as an endocrine disruptor2. Norms exist to restrict the quantity of fluorinated molecules present in environment but no technique allows for now to detect and identify theses fluorinated pollutants. The ANR granted project “Fluovial” aims at reaching this objective through 19F NMR.

The idea is to record 19F NMR3 data-sets from known fluorinated compounds and apply an algorithm specifically tuned for this application, with specific processing steps. A Machine Learning algorithm is used for the analysis, detection and identification of fluorinated compounds from NMR spectra. First, we prepare data-sets and extract the features used for training. Data preparation includes a dimensionality reduction and a feature selection from the data-set to allow a quick and robust training. Next, we train a Random Forest4 algorithm which was chosen here as it provides good preliminary results and allows to get an explanation of how the classification is made to check if it bio-physically makes sense. Then an optimization of the algorithm is needed to get out the best results and increase the robustness of the classifier. It can be of multi-faceted nature, from the choice of the data included in the training to a database augmentation (e.g. artificial addition of noise) but foremost the optimization of the hyper-parameters5, which are used to control the learning process. Thus, with this algorithm applied to unknown compounds we should be able to detect and identify the fluorinated molecules present.

We obtain good results on the database examples, with more than 90% of true predictions. This will be extensively trained, with more families of fluorinated compounds and with extended conditions to be more robust and efficient on “real” environment extract samples. Furthermore, the explainability of the decision trees could be checked to ensure the biophysical sense of the classification.

References

1 Wania, F., & Mackay, D. (1996). Peer reviewed: tracking the distribution of persistent organic pollutants. *Environmental science & technology*, *30*(9), 390A-396A.
2 Du, G., Hu, J., Huang, H., Qin, Y., Han, X., Wu, D., ... & Wang, X. (2013). Perfluorooctane sulfonate (PFOS) affects hormone receptor activity, steroidogenesis, and expression of endocrine-related genes in vitro and in vivo. *Environmental toxicology and chemistry*, *32*(2), 353-360.
3 Danielson, M. A., & Falke, J. J. (1996). Use of 19F NMR to probe protein structure and conformational changes. *Annual review of biophysics and biomolecular structure*, *25*, 163.
4 Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.
5 Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B. (2011, December). Algorithms for hyper-parameter optimization. In *25th annual conference on neural information processing systems (NIPS 2011)* (Vol. 24). Neural Information Processing Systems Foundation.

Unexpected similarity between TNF-α and HIV-1 reverse transcriptase binding sites revealed by a novel 3D computer vision-inspired method

Merveille Eguida, Didier Rognan

Laboratoire d'Innovation Thérapeutique, CNRS - Université de Strasbourg UMR 7200, Illkirch.

Identifying binding sites similarity between unrelated targets across the proteome is valuable to drug design in a range of applications spanning polypharmacology, off-target detection and ligand design [1], but still is a challenge to state of the art binding site comparison methods, notably regarding local similarities arising from pocket microenvironment.

We therefore developed ProCare [2], a new computational approach to compare protein pockets with a 3D point cloud registration algorithm. In computer vision, point cloud registration is a fundamental problem of finding the best transformation (rotation, translation, scaling) to align two clouds of points. A protein pocket is here represented as an ensemble of 3D atomic coordinates annotated with microenvironment specific pharmacophoric properties. Following the characterization of each point with a hybrid shape-chemical descriptor (c-FPFH) [2-3], two pockets are aligned by superimposing their corresponding points sharing the most similar patterns. The alignments are then evaluated by estimating the proportion of matched points sharing the same pharmacophoric properties.

Out of a large-scale comparison where ~33,000 subpockets from different proteins in the sc-PDB database were compared to the pocket of the trimeric Tumor Necrosis Factor alpha (TNF-α) [4], ProCare suggested a similarity with the non-nucleoside binding site of HIV-1 reverse transcriptase (HIVRT) [2]. Extensive binding site comparisons using different structures of TNF-α [4] and HIVRT reinforced that similarity hypothesis, which was later confirmed by microscale thermophoresis assay where two out of three tested HIVRT FDA-approved drugs were found to bind to TNF-α (*K*D in the 20-40 μM range). Remarkably, other binding site comparison methods as well as ligand 2D fingerprints and 3D shape methods were not able to detect that similarity.

ProCare allows local comparison of protein pockets of different sizes while yielding visually interpretable results. The method is ideally suited to identify local and unobvious similarities among totally unrelated targets and appears as a promising idea generator for fragment-based ligand design, able to pick starting fragments at a proteomic scale, not necessarily influenced by existing ligand or cavity neighborhoods. ProCare is an open source software freely available at https://github.com/kimeguida/ProCare.

References

[1] Ehrt, C.; Brinkjost, T.; Koch, O. *J. Med. Chem*. 2016, 59, 4121–4151.
[2] Eguida, M.; Rognan, D. *J. Med. Chem.* 2020, 63, 7127–7142.
[3] Rusu, R. B.; Blodow, N.; Beetz, M. 2009, *IEEE Int. Conf. Robot. Autom.* 2009, 3212–3217.
[4] O'Connell, J.; Porter, J.; Kroeplien, B.; Norman, T.; Rapecki, S.; Davis, R.; McMillan, D.; Arakaki, T.; Burgin, A.; Fox III, D.; Ceska, T.; Lecomte, F.; Maloney, A.; Vugler, A.; Carrington, B.; Cossins, B. P.; Bourne, T.; Lawson, A. *Nat. Commun*. 2019, 10, 5795.

Semi-Procedural Textures using Point Process Texture Basis Functions

Pascal Guehl[1], Rémi Allègre[1], Jean-Michel Dischler[1], Bedrich Benes[2], Eric Galin[3]

[1]Laboratoire des sciences de l'ingénieur, de l'informatique et de l'imagerie, CNRS-Université de Strasbourg, UMR7357
[2] : Purdue University [West Lafayette] USA
[3] : Laboratoire d'informatique en Image et Systèmes d'information
Université Claude Bernard Lyon 1, CNRS, Institut National des Sciences Appliquées de Lyon

---

Texturing 3D models is one of the key ingredients of computer-generated images of movies and video games. Unfortunately, tools used by digital artists require years of expertise. To automate these content creation tasks, two approaches are commonly used but still suffer drawbacks. On the one hand, data-driven texture synthesis methods try to generate similar arbitrary size images from small input exemplars. They require some stationarity, fail to preserve structures, and control and editing are difficult because of the lack of parameters. On the other hand, procedural modeling uses set of rules such as noise functions to synthesize patterns and add details. It requires expert knowledge, trial and error authoring and handle limited types of structures. Tools used in production are based on procedural node graphs. Currently, inverse procedural modeling, which tries to retrieve full node graph from image(s), is an active research field. It needs classification tasks and cannot guarantee visual resemblance.

We introduce a novel semi-procedural approach that avoids drawbacks of procedural textures and leverages advantages of data-driven texture synthesis. We split synthesis in two parts: 1) structure synthesis, based on a procedural parametric model and 2) color details synthesis, being data-driven. The procedural model consists of a generic Point Process Texture Basis Function (PPTBF), which extends sparse convolution noises by defining rich convolution kernels. They consist of a window function multiplied with a correlated statistical mixture of Gabor functions, both designed to encapsulate a large span of common spatial stochastic structures, including cells, cracks, grains, scratches, spots, stains, and waves. Parameters can be prescribed automatically by supplying binary structure exemplars. As for noise-based Gaussian textures, the PPTBF is used as stand-alone function, avoiding classification tasks that occur when handling multiple procedural assets. Because the PPTBF is based on a single set of parameters it allows for continuous transitions between different visual structures and an easy control over its visual characteristics. Color is consistently synthesized from the exemplar using a multiscale parallel optimization method, constrained by the PPTBF. The generated textures and materials are parametric, infinite and avoid repetition. The data-driven part is automatic and guarantees strong visual resemblance with inputs.

This work has been published in the Computer Grapics Forum journal in July 2020 and presented at the Eurographics Symposium on Rendering conference in July 2020 where it got an award: Honorable Mention from the Best Papers committee.

Reference

1. Guehl, P., Allegre, R., Dischler, J.M., Benes, B. and Galin, E., 2020, July. Semi-Procedural Textures Using Point Process Texture Basis Functions. In Computer Graphics Forum (Vol. 39, No. 4, pp. 159-171).

The absence of PARP3 promotes genome instability and causes tumour aggressiveness in human prostate cancer cells

Daisy Harwood, Françoise Dantzer

Biotechnologie et signalisation cellulaire, CNRS - Université de Strasbourg, UMR7242, 67412 Illkirch

Poly(ADP-ribosyl)ation is a post-translational modification of proteins that is catalyzed by enzymes called poly(ADP-ribose) polymerases or PARPs, a family of 17 members1. Within this family, PARPs 1 and 2 have been the centre of attention of the past decades, and the inhibition of PARP1 is now a major strategy used in anti-cancer therapy, based on the principle of synthetic lethality2.

During recent years, our team has been focused on determining the biochemical and biological specificities of PARP3, the third member of this family which has been less well described. Consequently, PARP3 has been identified as a key element in the repair pathway choice of double strand DNA breaks3, in mitotic segregation4, in epithelial to mesenchymal transition (EMT)5and in the activation of the oncogenic pathway mTORC2 in breast cancer6.

During my PhD, I aimed to explore the role of PARP3 in tumour aggressiveness in prostate cancer. The results acquired so far suggest that the absence of PARP3 (generation of PARP3-deficient clones using Crispr/Cas9) in PC-3 human prostate cancer cells causes the acquisition of aggressive (induction of a state of partial EMT) and migratory properties that are associated with genome instability that is defined by oxidative stress, DNA damage repaired by the BER pathway and replication stress.

My current experiments aim to define the replication stress in my model (Repli-seq, molecular combing) and to specify the impact of this genome instability on EMT properties (activation of stem cells, chemoresistance). The goal of my work is to understand how PARP3 maintains genome stability to limit tumour progression in prostate cancer.

References

1. Schreiber V, *et al*. Poly(ADP-ribose): novel functions for an old molecule. *Nat Rev Mol Cell Biol*. **2006**Jul;7(7):517-28.
2.Bryant HE, *et al*. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature*. **2005**Apr 14;434(7035):913-7.
3. Beck C, *et al*. PARP3 affects the relative contribution of homologous recombination and nonhomologous end-joining pathways. *Nucleic Acids Res*. **2014**May;42(9):5616-32.
4. Boehler C, *et al*. Poly(ADP-ribose) polymerase 3 (PARP3), a newcomer in cellular response to DNA damage and mitotic progression. *Proc Natl Acad Sci U S A*. **2011**Feb 15;108(7):2783-8.
5. Karicheva O, *et al*. PARP3 controls TGFβ and ROS driven epithelial-to-mesenchymal transition and stemness by stimulating a TG2-Snail-E-cadherin axis. *Oncotarget*. **2016**Sep 27;7(39):64109-64123.
6. Beck C, *et al*. PARP3, a new therapeutic target to alter Rictor/mTORC2 signaling and tumor progression in BRCA1-associated cancers. *Cell Death Differ*. **2019**Sep;26(9):1615-1630.

Caveolin-1/EREG/YAP axis in the treatment resistance of caveolin 1-expressing head and neck squamous cell carcinoma

Aude Jehl, Mickaël Burgy, Sophie Martin

Laboratoire de Bioimagerie et Pathologies, CNRS - Université de Strasbourg, UMR7021, 74 route du Rhin 67400 Illkirch-Graffenstaden

---

Head and neck squamous cell carcinoma (HNSCC) represent the 5th most common cancer worldwide with an annual incidence and mortality estimated to be around 600 000 and 375 000 cases, respectively[1]. Locally-advanced head and neck squamous cell carcinoma (LA-HNSCC) (stage III/IV) represents about 60% of patients at diagnosis.

They require a combination of therapies including a primary surgery followed by radiotherapy or chemoradiation in patients at high risk of relapse. The epithelial growth factor receptor (EGFR)-targeting antibody cetuximab combined with radiotherapy is the only targeted therapy that has been proven effective for the treatment of LA-HNSCC. Recurrences arise in 50% of patients with HNSCC in the years following treatment. In clinico-pathological practice, it is difficult to assign patients into classes of risk since no reliable biomarkers are available to predict the outcome of Human Papillomavirus-unrelated HNSCC.

A body of evidence suggests that caveolin-1 may be involved in the resistance of HNSCC to anti-EGFR therapies. In the present study, we showed that caveolin-1 overexpressing cells exert better surviving capacities and remain proliferative and motile when exposed to chemo-radiotherapy. Resistance to the standard chemo-radiotherapy regimen is mediated by the caveolin-1/EREG/YAP axis. In addition, patient whose tumor overexpressed Caveolin-1 recurred regionally a few years after treatment.

Altogether our observations suggest that high expression of Cav1 might be predictive of locoregional relapse of LA-HNSCC due to resistance to the standard chemo-radiotherapy regimen.

Reference

[1] Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. International journal of cancer 2015;136:E359-86

Involvement of the GPRASP1 protein in the development of tolerance to analgesic effects of delta opioid receptor agonists.

Juliette Kaeffer[1], Adam Medina[1], Glenn Marie Le Coz[1], Raphaëlle Quillet[1], Khadija Elhabazi[1], Dominique Massotte[2], Franck Vandermoere[3], Frédéric Simonin[1], Sandra Lecat[1]

[1] Biotechnologie et signalisation cellulaire,
CNRS - Université de Strasbourg, UMR7242, Illkirch
[2] Institut des Neurosciences Cellulaires et Intégratives,
CNRS - Université de Strasbourg, UPR3212
3 : Institut de Génomique Fonctionnelle
Université de Montpellier, CNRS UMR5203, INSERM U1191, F-34094, Montpellier

---

The delta opioid receptor (DOR) is a G protein-coupled receptor (GPCR) involved in pain homeostasis. Activation of DOR by agonists induces analgesia in chronic pain models. However, repeated administrations of DOR agonists lead to a rapid loss of their analgesic effect called analgesic tolerance. We are studying the molecular mechanisms involved in analgesic tolerance with two working hypothesis: (i) tolerance at the cellular level could be caused by a rapid loss of activated DOR due to their degradation into lysosomes (ii) Alternatively, DOR repeated stimulations (leading to analgesia),could induce counteractive signaling pathways leading to activation of pro-nociceptive signals in order to restore pain homeostasis.

My project is to characterize the role of the intracellular protein GPCR-associated sorting protein 1 (GPRASP1) in the analgesic tolerance to DOR agonists. Indeed, GPRASP1 has been identified by our laboratory as a protein that interacts with DOR *in vitro* (1). It has been proposed by others to target DOR receptor for degradation (2).

We have generated GPRASP1-deficient mice expressing a fluorescent DOR (DOR-eGFP) by crossing our in-house GPRASP1-KO mice with knock-in DOR-EGFP mice (3). This unique mice model allows me to perform behavioural analysis of pain and to track DOR-GFP with specific robust antibodies and by fluorescence. I have shown that analgesic tolerance to repeated DOR activation is absent in GPRASP1-KO mice using two chronic pain models. In the brain, DOR are degraded to the same extent in WT and KO-GPRASP1 mice after repeated stimulation as measured by two techniques. My aim is now to compare DOR interactome, trafficking and signaling between tissues of WT and KO-GPRASP1 mice treated by repeated agonist stimulation. Since GPRASP1 associates with other GPCRs (4), the elucidation of its involvement in the mechanism tolerance may provide insight into the molecular adaptions to chronic activation of other GPCRs.

References

1. Simonin, F. *et al.*Identification of a novel family of G protein-coupled receptor associated sorting proteins. J. Neurochem, **89**, 766-775(2004)
2. Whistler, J.L. *et al.*Modulation of postendocytic sorting of G protein-coupled receptors. Science,**297**, 615–620 (2002)
3. Scherrer, G. *et al*. Knockin mice expressing fluorescent delta-opioid receptors uncover G protein-coupled receptor dynamics in vivo. Proc Natl Acad Sci USA.; **103**, 9691–9696 (2006)
4. Kaeffer J, *et al*.GPRASP/ARMCX protein family: potential involvement in health and diseases revealed by their novel interacting partners. Current topics in medicinal chemistry, **21**:227-254 (2021)

End-to-end deep representation learning for time series clustering: a comparative study

Baptiste Lafabregue[1], Jonathan Weber [1], Pierre Gançarski[3], Germain Forestier[2]

1 : Institut de Recherche en Informatique Mathématiques Automatique Signal Université de Haute-Alsace (UHA) Mulhouse - Colmar : EA7499
2 : Institut de Recherche en Informatique Mathématiques Automatique Signal - IRIMAS - UR 7499, Université de Haute-Alsace (UHA) Mulhouse - Colmar : UR7499
3 : Laboratoire des sciences de l'ingénieur, de l'informatique et de l'imagerie CNRS-Université de Strasbourg : FR3627

Time series are ubiquitous in data mining applications. Similar to other types of data, annotations can be challenging to acquire, thus preventing from training Time Series Classification (TSC) models.

In this context, clustering can be an alternative to cluster time series into homogeneous groups allowing a better analysis of the structure of the data. Time series clustering has been investigated for many years and multiple approaches have already been proposed. Following the advent of deep learning in computer vision, researchers recently started to study the use of deep clustering to cluster time series data. The existing approaches mostly rely on representation learning (imported from computer vision), which consists in learning a representation of the data and performing the clustering task using this new representation.

The goal of this paper is to provide a careful study and an experimental comparison of the existing literature on time series representation learning for deep clustering. In this paper, we went beyond the sole comparison of existing approaches and proposed to decompose deep clustering methods into three main components: (1) network architecture, (2) pretext loss, and (3) clustering loss.

We evaluated all combinations of these components (totaling 300 different models) with the objective to study their relative influence on the clustering performance. We also experimentally compared the most efficient combinations we identified with existing non-deep clustering methods. Experiments were performed using the largest repository of time series datasets (the UCR/UEA archive) composed of 128 univariate and 30 multivariate datasets.

Finally, we proposed an extension of the Class Activation Maps (CAM) method to the unsupervised case which allows to identify patterns providing highlights on how the network clustered the time series.

Drug-sponge lipid nanocarrier for in situ cargo loading and release using dynamic covalent chemistry

Fei Liu[1,2], Yosuke Niko[1,3], Redouane Bouchaala[1], Luc Mercier[1], Olivier Lefebvre[4], Bohdan Andreiuk[1], Thierry Vandamme[2], Jacky Goetz[4], Nicolas Anton[2], Andrey Klymchenko[1]

[1] Laboratoire de Bioimagerie et Pathologies, UMR 7021 CNRS - Université de Strasbourg, 74 route du Rhin, 67401, Illkirch, France
[2] Laboratoire de Conception et Application de Molécules Bioactives, CNRS - University of Strasbourg, UMR 7199, 74 route du Rhin, 67401 Illkirch Cedex, France
[3] Research and Education Faculty, Multidisciplinary Science Cluster, Interdisciplinary Science Unit, Kochi University, 2-5-1, Akebono-cho,Kochi-shi, Kochi, 780-8520, Japan.
[4] MN3T, Inserm U1109, LabEx Medalis, Fédération de Médecine Translationnelle de Strasbourg (FMTS), University of Strasbourg, F-67200, France

---

Currently, drug-delivery strategies using nanocarriers (NCs) deal with encapsulation of cargo or its covalently modified prodrug. Lipid nanoemulsions are particularly attractive “green” nano-carriers, as they are composed of materials generally recognized as safe (GRAS).[1] However, liquid nature of their oil core limits makes controlled encapsulation of cargo rather challenging. Herein, we describe a concept of reversible pH-controlled capture and delivery of active cargo based on dynamic covalent chemistry inside lipid nano-droplets (nanoemulsions), coined as “drug sponge”.[2]

The idea is to design highly lipophilic hydrazide/ capable of reacting with a free cargo-ketone (fluorescent dyes and drugs) directly inside lipid NCs, yielding a lipophilic hydrazone prodrug efficiently captured in the oil core. By using FRET as a tool for the determination of cargo (as acceptor/donor) in nanocarriers, we are able to identify whether those functional NCs have the capability of spontaneously accumulating cargo-ketones, yielding formulations stable against cargo leakage at neutral, and further released their dye/drug cargo at acetic pH in solution and live cells. Doxorubicin-loaded drug-sponge NCs showed cytotoxicity in four cancer cell lines and capacity to inhibit tumor growth in subcutaneous xenografts of mice.

Finally, unprecedented extraction of dye/drug cargos directly from cells and tissues (i.e. detoxification) was realized by the drug-sponge NCs, opening the route to new detoxification strategies.

References

1. Klymchenko A. S., Liu F., Collot M., Anton N., Dye-Loaded Nanoemulsions: Biomimetic Fluorescent Nanocarriers for Bioimaging and Nanomedicine, Advanced Healthcare Materials, 2021, 10, 2001289.
2. Liu, F.; Niko, Y.; Bouchaala, R.; Mercier, L.; Lefebvre, O.; Andreiuk, B.; Vandamme, T.; Goetz, J. G.; Anton, N.; Klymchenko, A. S. Drug-sponge lipid nanocarrier for in situ cargo loading and release using dynamic covalent chemistry. Angew Chem. Int. Ed. 2021, 60, 2021, 6573-6580

N-heterocyclic carbene platinum complexes as anti-cancer stem cell therapeutics.

Conor Mccartin[1], Monique Dontenwill[2], Stéphane Bellemin-Lapponnaz[3], Sylvie Fournel[1] and Antoine Kichler[1]

[1] Conception et application de molécules bioactives – CNRS - Université de Strasbourg, UMR7199, Illkirch
[2] Laboratoire de Bioimagerie et Pathologies – CNRS - Université de Strasbourg, UMR7021, Illkirch
[3] Institut de Physique et Chimie des Matériaux de Strasbourg (IPCMS) CNRS – Université de Strasbourg, UMR7504, 23 rue du Loess - Strasbourg Cedex 2

The identification of a stem cell-like niche within tumours, referred to as cancer stem cells (CSCs) has changed the manner in which cancer biogenesis and resistance is regarded. The niche, which possesses the stem cell like characteristics of auto-regeneration and the ability to differentiate, is purported to be involved in the persistent clinical reoccurrence of tumours, through their resistance to classical therapeutic treatments, allowing the niche to repopulate the bulk tumour mass1. Thus, the continued development of metal-based anti-cancer therapeutics which may overcome this issue is of great interest.

Recently, it has been revealed that the CSCs possess high mitochondrial activity, suggesting the importance of this organelle in the metabolism of CSCs and its possible implication in resistance2. It is for this reason that the development of mitochondrial targeting anti-cancer agents has gathered attention in recent years3.

The basis for this work is one such NHC-Pt(II)-PEI complex with a coordinated polyethyleneimine ligand which has been developed by S. Bellemin-Laponnaz. This compound has shown promise due to its ability to kill p53 deficient cisplatin resistant cells as well as the ability to reduce tumour size in an *in-vivo* mouse model with no observable side-effects. The compound localises significantly in the mitochondria, where it induces mitochondrial superoxide production, and inhibits mitochondrial respiration4. This suggests a different mode of action to other platinum anti-cancer drugs. We are interested in the study of the effect of this compound against cancer stem cells, and its mechanism of action. Studies so far show a high level of toxicity towards cancer stem cells, inducing a cell death through a non-apoptotic mechanism which implicates an activation of autophagy, which is different to that of currently used anti-cancer platinum compounds.

Modulation of expression of therapeutic targets in glioblastoma; Example of the α5β1 integrin

Mélissa Messé, Chloé Bernhard, Marie-Cécile Mercier, Sophie Foppolo, Monique Dontenwill

Laboratoire de Bioimagerie et Pathologies – CNRS - Université de Strasbourg, UMR7021, Illkirch

Glioblastoma (GBM) is the most frequent and agressive tumor of the adult central nervous system. With an incidence of 3/100000 worldwide, it remains of bad prognosis and no new efficient therapies have been described since 2005 1. High inter- and intra-tumoral molecular heterogeneity explain in part the failure of numerous GBM clinical trials. In addition the existence of GBM stem cells in the tumors was linked to therapy resistance and recurrence. We demonstrated earlier that α5β1 integrin is a GBM potential therapeutic target as an high expression (at the mRNA and protein levels) is linked to a worse patient prognostic 2,3. We also
observed that α5 protein expression is highly heterogeneous between tumors and in different locations of a given tumor. The aim of our work is to understand how this integrin expression is modulated in GBM.

We first hypothesized that distinct populations of GBM stem cells may exist either expressing
or not the integrin. We showed that the patient-derived stem cell lines (about 10) examined did not express the integrin in stem culture conditions (neurospheres) but interestingly half of them did after differenciation (adherent cells). GBM stem cells reside in specific niches such as perivascular or hypoxic area. We next wondered if hypoxia may modulate α5β1 integrin expression. Tumor stem cell lines were subjected to different conditions of incubation either in normoxia (21% O2) or in hypoxia (1% O2). We also used chemically-induced hypoxia with cobalt chloride (COCL2) or desferoxiamine (DFO), We followed kinetically (24h, 48h and 72h) the impact of hypoxia on α5β1 integrin induction. Our results show that even in stem cell culture conditions, expression of the integrin was induced or increased in hypoxia but only for some of the stem cell lines. Taken together, these data confirm that different populations of stem cells exist in GBM and that α5β1 integrin expression may be induced by different signaling pathways in different cells.

Here we add new evidence that molecular switches may occur either when stem cells differentiate to tumor cells but also directly in stem cells in hypoxic niches. Work is in progress
to elucidate the different molecular mechanisms implicated in the trigger of α5 integrin expression which associates with GBM increased aggressivity. Characterization of regulators may also offer the possibility to propose new therapeutic targets.

References

1. Stupp, R. et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N. Engl. J. Med. 352, 987–996 (2005).
2. Renner, G. et al. Integrin α5β1 and p53 convergent pathways in the control of anti-apoptotic proteins PEA-15 and survivin in high-grade glioma. Cell Death Differ. 23, 640–653 (2016).
3. Janouskova, H. et al. Integrin α5β1 plays a critical role in resistance to temozolomide by interfering with the p53 pathway in high-grade glioma. Cancer Res. 72, 3463–3470 (2012)

New strategies for genome annotation using deep learning algorithm

Nicolas Scalzitti, Anne Jeannin-Girardon, Pierre Collet, Olivier Poch, Julie Thompson

Complex Systems and Translational Bioinformatics (CSTB), ICube laboratory, UMR7357, University of Strasbourg, 1 rue Eugène Boeckel, 67000, Strasbourg, France

---

High-throughput technologies are constantly generating huge amounts of genomic sequences that represent an essential source of information for studying and understanding living organisms. However, without a crucial "annotation" step to add information, these raw sequences are difficult to exploit and sometimes even useless. One of the main challenges of annotation is to identify genes and characterize their internal structure in the genome (Salzberg, 2019). In Eukaryotes, this step is very complex in particular for protein coding genes. Indeed, the architecture of these genes is organized in a mosaic of exons and introns (Mudge and Harrow, 2016) delimited by boundaries called splice sites. There are two types of splice sites, the 5' (donor) and 3' (acceptor) sites, which are respectively the junction between exon-introns and intron-exons (Matera and Wang, 2014). The splice sites are mainly characterized by the presence of GT (5') and AG (3') dinucleotides, embedded in a longer, more divergent pattern of about ten nucleotides. To help identify these sites, many prediction programs based on machine learning algorithms have been developed (Jaganathan *et al.*, 2019), (Wang *et al.*, 2019). Unfortunately, these programs are often dedicated to model organisms (e.g A. *Thaliana*, C. *Elegans*) or to human, and still generate too many annotation errors (Drăgan *et al.*, 2016) that affect the quality of the protein sequence data included in public databases.

In this context, based on the hypothesis of a universal structure of splice sites, we have developed of a new artificial intelligence algorithm based on reference dataset (Scalzitti *et al.*, 2020), to predict multi-species splice sites (from Human to Protists). We first developed a quality control protocol that allowed us to identify good and bad quality data. This was then integrated in a new deep learning approach based on a convolutional neural network called Spliceator. The high-quality dataset allows Spliceator to achieve comparable or even better performance compared to other best recent methods.

References

- Drăgan,M.-A. *et al.* (2016) GeneValidator: identify problems with protein-coding gene predictions. *Bioinformatics*, **32**, 1559–1561.
- Jaganathan,K. *et al.* (2019) Predicting Splicing from Primary Sequence with Deep Learning. *Cell*, **176**, 535-548.e24.
- Matera,A.G. and Wang,Z. (2014) A day in the life of the spliceosome. *Nat Rev Mol Cell Biol*, **15**, 108–121.
- Mudge,J.M. and Harrow,J. (2016) The state of play in higher eukaryote gene annotation. *Nat Rev Genet*, **17**, 758–772.
- Salzberg,S.L. (2019) Next-generation genome annotation: we still struggle to get it right. *Genome Biol*, **20**, 92, s13059-019-1715–2.
- Scalzitti,N. *et al.* (2020) A benchmark study of ab initio gene prediction methods in diverse eukaryotic organisms. *BMC Genomics*, **21**, 293.
- Wang,R. *et al.* (2019) SpliceFinder: ab initio prediction of splice sites using convolutional neural network. *BMC Bioinformatics*, **20**, 652.

Buchwald-Hartwig amination in alcohols

Steinsoultz Philippe[1], Bailly Aurélien[2], Schmitt Martine[1],
Grimaud Laurence[2] and BihelL Frédéric[1]

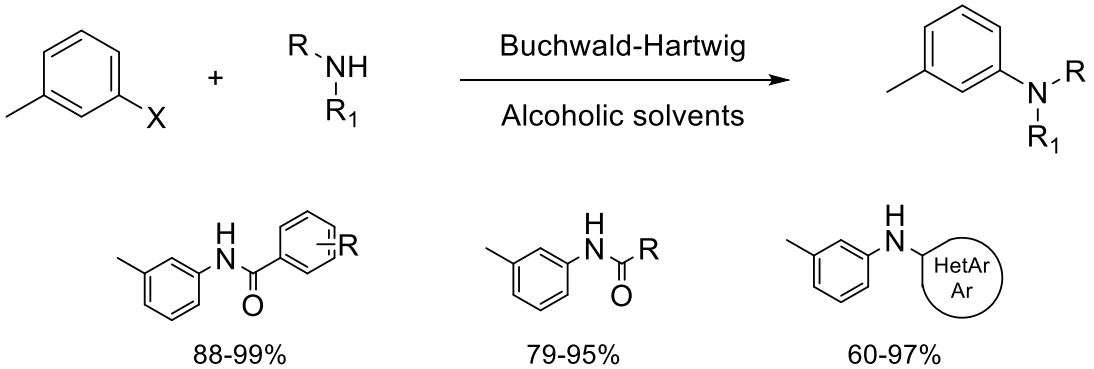
[1]Laboratoire d'Innovation Thérapeutique, UMR7200, CNRS - Université de Strasbourg
74 route du Rhin 67400 Illkirch-Graffenstaden
[2] Laboratoire de biomolécules (LBM), Département de chimie, UMR7203, CNRS - Sorbonne Université, École normale supérieure, 75005 Paris, France

Aromatic amines are essential molecules, and can be found in many natural products, as well as in synthetic building blocks used in agrochemical or pharmaceutical industries. Historically prepared by multistep process (nitration, reduction), nucleophilic aromatic substitution or by copper-mediated Ullmann-type couplings these strategies show drawbacks such as a limited scope and functional group tolerance due to harsh reaction conditions. In 1994, the palladium catalyzed amination developed by Buchwald and Hartwig emerged and is now considered as a fundamental transformation.

Micellar chemistry has been used to performed metal-catalysed reactions following green chemistry concepts, such as Suzuki-Miyaura[1] Buchwald-Hartwig[1] or Ullmann-type[3] reactions. However micellar medium still presents some limitations in terms of solubilisation, requiring the use of organic co-solvents.

To bypass this limitation, we explored the opportunity to use alcoholic solvents as suitable alternatives. These solvents are considered as "green" by chemical industries[4]: they can be obtained from renewable feedstock; they are safer to use and except for the methanol, their toxicity for human health and environment is lower compared to other classical organic solvent.

Using ethanol as solvent and aqueous KOH as base, we optimized Buchwald-Hartwig amination reactions at only 50°C. The cross-coupling reaction is achieved in few hours, giving excellent yields with a large scope of substrates, including amides, indoles, carbamates, and anilines. Aryl iodides, bromides and chlorides can be used efficiently. Mechanistic studies were performed using cyclic voltammetry, ^{31}P & ^{19}F NMR experiments. We performed these studies in order to understand the nature of the palladium species present in alcohol, as well as the kinetic of the reaction.



References

1 S. Handa *et al., Science,* **2015**, 349, page 1087-page 1091.
2 P. Wagner *et al., Green Chemistry,* **2014**, 16, page 4170-page 4178.
3 M. Bollenbach *et al., Chem. Eur. J.* **2017,** 23, page 13676-page 13683.
4 F. P. Byrne *et al., Sustainable Chemical Processes,* **2016**, 4.

MISTIC: A prediction tool to reveal disease-relevant deleterious missense variants

Thomas Weber[1], Kirsley Chennen[2], Xavière Lornage[3], Arnaud Kress[1], Johann Bohm[3], Julie Thompson[1], Jocelyn Laporte[2], Olivier Poch[1]

[1] Laboratoire des sciences de l'ingénieur, de l'informatique et de l'imagerie (ICube), CNRS-Université de Strasbourg, École Nationale du Génie de l'Eau et de l'Environnement de Strasbourg (ENGEES), Institut National des Sciences Appliquées (INSA), UMR7357 300 bd Sébastien Brant - 67412 Illkirch
[2] Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC) CNRS-Université de Strasbourg, UMR7104, Inserm, U964, Parc D'Innovation 1 Rue Laurent Fries - 67404 Illkirch

---

The advent of high-throughput sequencing methods has made it possible to study these Mendelian diseases at the highest resolution available, the nucleotide level. A patient's genome can now be fully sequenced in order to study its genetic variations1. However, the technical limitations today lie in the exploitation of the data produced, particularly due to its volume and complexity.

Approximately five million single nucleotide variations (SNV)2 are present in each human genome resulting from both evolution and inter-individual diversity. Among these, 150,000 are found in the protein-coding regions, called exome, with an impact on gene products close to zero. Nevertheless, in rare cases, SNVs can lead to functional or structural modifications that result in Mendelian disease.

To date, the most studied SNV is the class of missense variations. A missense corresponds to a SNV leading to a change of amino acid in the peptide sequence product during translation. Unlike most damaging classes, which can for example, result in a stop codon, missense SNVs are difficult to study due to their variable consequences. Out of the 150,000 SNVs present in each human exome, the number of missense variations is estimated to be around 1,500 and finding the causative variation for a rare disease in an exome study is like searching for a needle in a haystack. Prediction of the impact of a missense is currently based on multiple parameters: frequency in the general population (large genomic databases), conservation during evolution, physico-chemical properties of the reference and the new amino acid, location in the protein (domain).

With the increase of computational power and emergence of artificial intelligence methods, different algorithms3,4 have been developed to help both researchers and physicians to find disease-causing variations in clinical studies.

We present MISsense deleTeriousness predICtor (MISTIC), a new original prediction tool based on an original combination of two complementary machine-learning algorithms that integrates 115 features, ranging from multi-ethnic minor allele frequencies and evolutionary conservation, to physiochemical and biochemical properties of amino acids. Our approach also uses training sets with a wide spectrum of variant profiles, including both positive (deleterious) and negative (benign) variants. Compared to recent state-of-the-art ensemble prediction tools in various benchmark tests, MISTIC exhibits the best and most consistent performance, notably with the highest AUC value (0.95). Importantly, MISTIC maintains its high performance in the specific case of discriminating deleterious variants from rare benign variants (allele frequency <1%) or population-specific benign variants (no allele frequency). In a clinical usage context, MISTIC drastically reduces the list of candidate variants (<30%) and has a median ranking of the “causative” deleterious variants among the top 25 variants.