An Integrative Bioinformatics Approach Rapidly Identifies *in vivo*-validated Drug Candidates with Novel Mechanisms of Action in Rheumatoid Arthritis

Aaron C. Daugherty[1], Carl Farrington[1], Isaac Hakim[1], Andrew M. Radin[1], Elizabeth S. Noblin[1]
Mei-Sze Chua[2], Marina Sirota[3], Mark Rabe[4], Guy Fernald[4], Daniel Ford[4], Laura Schaevitz[4,5],
Andrew A. Radin[1,6]

[1]twoXAR, Inc. Palo Alto, CA. All authors contributed equally to this work. [2]Stanford [3]UCSF [4]Vium, Inc. [5]Email: laura@vium.com [6]Email: aradin@twoXAR.com

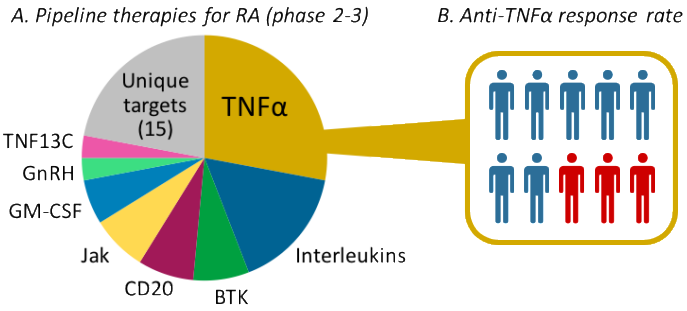
Abstract

Rheumatoid arthritis (RA) is an area of active drug development, with over 100 candidates in clinical trials. However, most act on a small number of immunomodulatory targets. Drug candidates that act through new targets or mechanisms could expand treatment options for RA. We applied a data-driven bioinformatics approach and *in vivo* screen to identify and test new drug candidates and targets that could form the basis of future drug development in RA.

A computational model of RA was constructed by integrating patient gene expression data, molecular interactions, and clinical drug-disease associations. Candidates were scored based on their predicted efficacy across these data types. FDA-approved treatments for RA were significantly enriched among the top-ranked candidates. Ten high scoring novel candidates were then screened in the collagen-induced arthritis model of RA in rats. Therapeutic treatment with three candidates (exenatide, olopatadine, and TXR-112 [data not shown]) significantly reduced ankle size, alleviated limb inflammation, improved joint histopathology, and reduced mobility impairments tracked by a novel digital motion endpoint. These candidates do not act on common RA therapeutic targets. However, links between known candidate pharmacology and pathological processes in RA suggest hypothetical mechanisms that could contribute to efficacy.

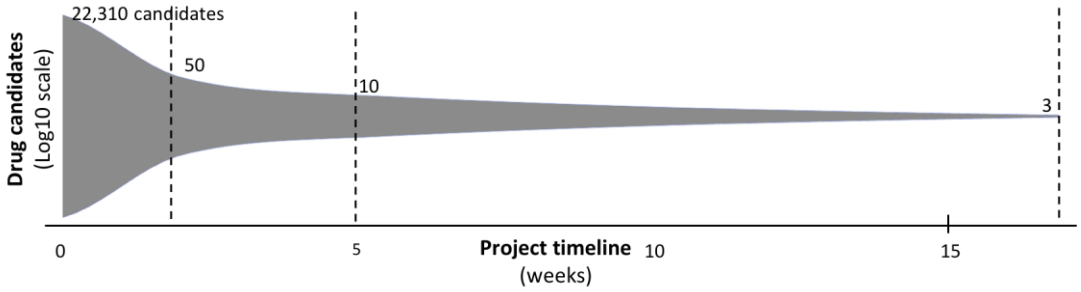
Future studies will inform the druggable targets, pathways, and mechanisms that could contribute to each candidate's efficacy in RA. The candidates could themselves be modified and optimized to increase efficacy in RA. Novel targets identified in these studies could also be the basis of new drug discovery initiatives.

Background: More diverse therapies for RA are needed



The RA drug development pipeline is heavily enriched for known immunomodulatory targets. For example, over 25% of all candidates in Phase 2 or Phase 3 clinical trials target TNFα (based on data from Thompson Reuters Integrity, accessed in October 2016). Although therapies targeting TNFα represent over 25% of the RA drug development pipeline, these therapies are ineffective at reducing RA symptoms or slowing disease progression in 30% of patients[1]. Abbreviations: TNF, tumor necrosis factor; BTK, Bruton's tyrosine kinase; GM-CSF, granulocyte macrophage colony stimulating factor; GnRH, gonadotropin releasing hormone.

Overview: Rapid discovery and *in vivo*-validation of candidates



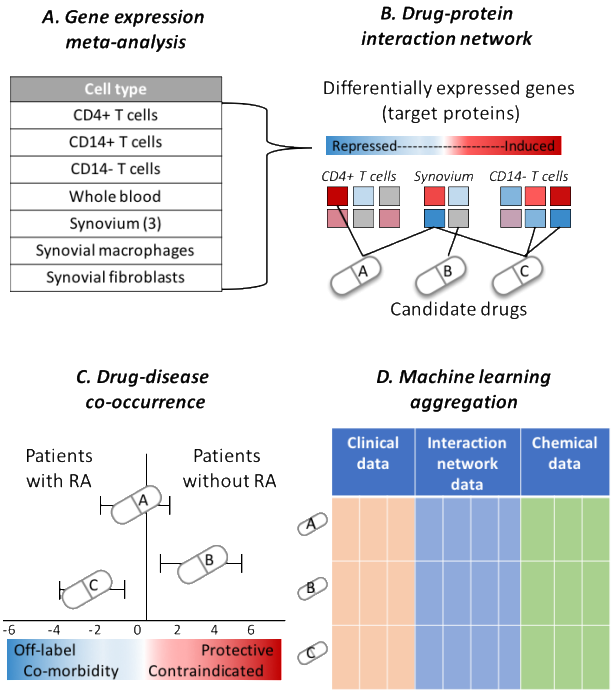
twoXAR's DUMA drug discovery platform rapidly identifies high-potential treatments from among thousands of possible drug candidates. In Phase 1, 22,310 large and small molecules present in DrugBank[2] and the Therapeutic Targets Database[3] were scored based on predicted therapeutic potential to yield 50 high-probability candidates. In Phase 2, algorithm evaluation and candidate due diligence identified 10 optimal candidates. In Phase 3, these 10 candidates were tested in an animal model of RA and 3 lead candidates were identified. The entirety of the project, from conception to completion, was finished in 4 months.

Phase 1: Integrative bioinformatic drug predictions

Gene expression:	Drug candidates:	Protein-protein interactions:	Clinical data:
994 samples	22,310	210,057	418,797

Highlight: twoXAR's approach identifies high-potential candidates by integrating many data points from multiple biomedical data sources. In the case of RA, data used to support drug candidate prediction included: gene expression data from RA patients and healthy controls, drug-protein interactions, and records of drug use among patients with and without RA.

Supporting data:



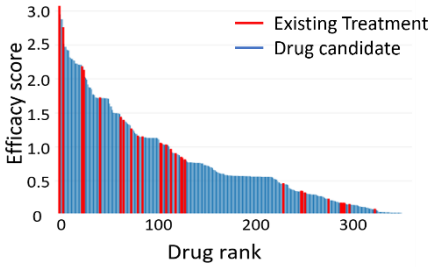
A. Gene expression meta-analysis. A meta-analysis of gene expression changes in RA patients versus healthy controls was performed using data spanning multiple tissues and cell types implicated in the pathology of RA. All data were downloaded from NCBI Gene Expression Omnibus[4] and processed using a standard microarray analysis pipeline.

B. Drug-protein interaction network. Information on drug-protein interaction was collected for 22,310 drug candidates from DrugBank and the Therapeutic Target Database along with protein-protein interaction data from DrPIAS[5]. Interaction information was integrated with protein expression changes (approximated from corresponding gene expression changes), and candidates were scored according to the number and confidence of interactions. In this method, scoring favors drugs that interact with proteins misexpressed in multiple gene expression datasets.

C. Drug-disease co-occurrence. Drug-disease co-occurrence scores were calculated by accessing disease diagnoses and medication use from participants in the FDA Adverse Event Reporting System. Odds-ratios of drug candidate use among patients with and without a diagnosis of RA were calculated. Candidates that tend to co-occur with an RA diagnoses (positive scores) may represent efficacious off-label treatments or medications used to treat co-morbidities of RA. Candidates that do not tend to co-occur with an RA diagnosis (negative scores) may represent treatments that are protective against RA or contraindicated in patients with RA.

D. Composite scoring. Candidate drug efficacy scores and features from methods B and C were aggregated into a final score using a proprietary machine learning-based approach.

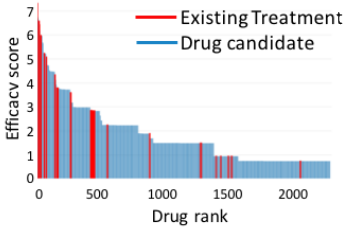
Phase 2: Evaluation & candidate selection



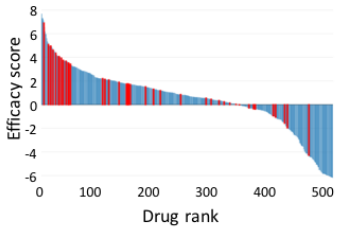
Highlight: To evaluate the predictive capacity of scoring methods, we examined scores for FDA-approved and clinically used RA treatments (collectively, "Existing Treatments"). Enrichment of these treatments near the top of a ranked list of candidates correlates with the capacity of each method to predict novel treatments. RA treatments were most highly enriched in the aggregate scoring method (shown here, $p = 9x10^{-20}$). After excluding RA treatments from the ranked candidate list, the 50 top remaining candidates were evaluated based on novelty and safety in RA. Candidates that had been suggested as RA treatments (by research or patent), deemed unsafe for RA, or unavailable for purchase were excluded.

Supporting data:

Interaction network



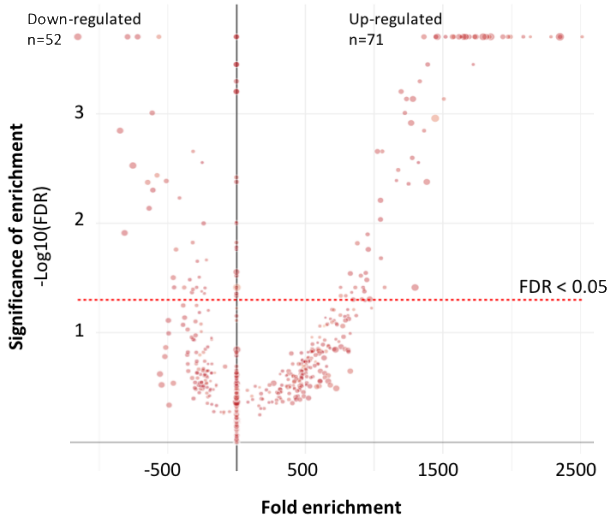
Drug-disease co-occurrence



<u>A. Existing treatment enrichment in efficacy scoring methods.</u> As for the aggregate scoring method, existing treatments for RA were enriched among top-scoring candidates in the drug-protein interaction (left, $p = 1.7x10^{-15}$) and drug-disease co-occurrence (right, $p = 4x10^{-3}$) scorning methods.

<u>B. Reactome pathways enriched in RA gene expression meta-analysis.</u> The biologic relevance of differentially expressed genes in the RA gene expression meta-analysis was assessed by quantifying enrichment of Reactome pathways using a GSEA[6]-like method. The graph (top) denotes all pathways, represented by dots, that were enriched for down-regulated genes in RA (left) or enriched for up-regulated genes (right). The size of the dot corresponds to the number of all proteins within that pathway that were detected in the RA meta-analysis. The most significant pathways (FDR < 0.05) were aggregated into general themes and reported in the accompanying tables.

B. Reactome pathways enriched in RA gene expression meta-analysis

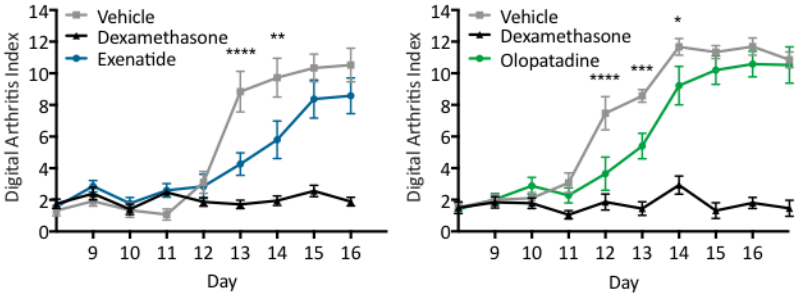


Significantly down-regulated themes
Translation processes
Transcription processes
Signal transduction pathways: FGFR, PI3K
Phosphatidylinositol metabolism

Significantly up-regulated themes
Immune processes (e.g. T-cell receptor and cytokine signaling
ECM organization (e.g. collagen formation)
Hemostasis, including PDGF signaling
Glycosaminoglycan metabolism
Insulin signaling

Phase 3: Preclinical validation of candidates

Highlight: Ten optimal and highly ranked candidates were tested in the collagen-induced arthritis (CIA) model of RA in female Lewis rats. Of these, three candidates alleviated multiple symptoms of CIA. Shown here is the effect of two candidates (exenatide, 10 µg/kg administered subcutaneously, and olopatadine, 2 mg/kg delivered via oral gavage) on Digital Arthritis Index, which is calculated using continuous digital monitoring of animal movement. Higher Digital Arthritis Index scores correspond to more severely impaired mobility. As a positive control for all assays, animals were treated with 75 µg/kg dexamethasone.

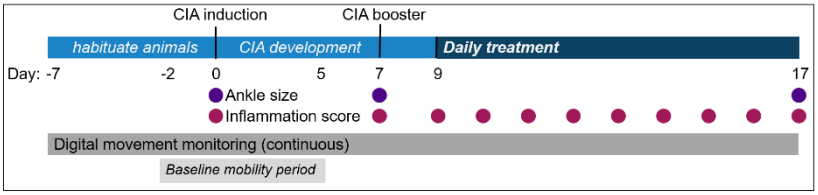


Supporting data:

A. Study design and endpoints. Female Lewis rats, aged 7-10 weeks, were injected with type II collagen in incomplete Freud's adjuvant and treated with candidate drugs according to the schedule shown. Movement was continually monitored, and mobility during study days -2 to 5 was used to randomly assign animals to treatment groups (8 animals per group). At the conclusion of the study, animals were sacrificed and ankle joints were processed for histopathology.

B. Candidate dosing. Dosing information for exenatide and olopatadine, two candidates that alleviated multiple CIA endpoints. Three doses of each candidate were tested; data for the maximally efficacious dose are shown.

A. Study design and endpoints

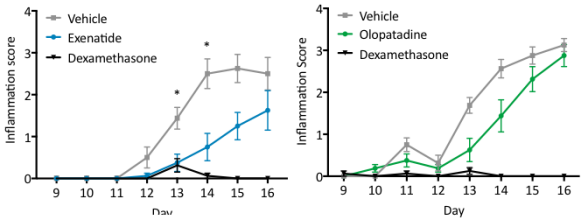


B. Candidate dosing

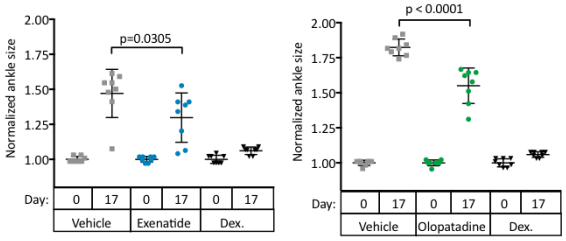
Drug	Original indication	Vehicle	Route	Dose
Exenatide	Type 2 diabetes	Saline	SC	10*, 50, 100 µg/kg
Olopatadine	Allergic conjunctivitis	1% DMSO (saline)	PO	1, 2* & 4 mg/kg

*most effective dose

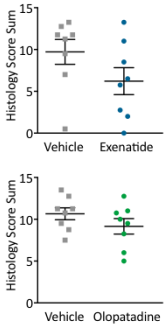
C: Hind-limb Inflammation Score



D: Ankle Size

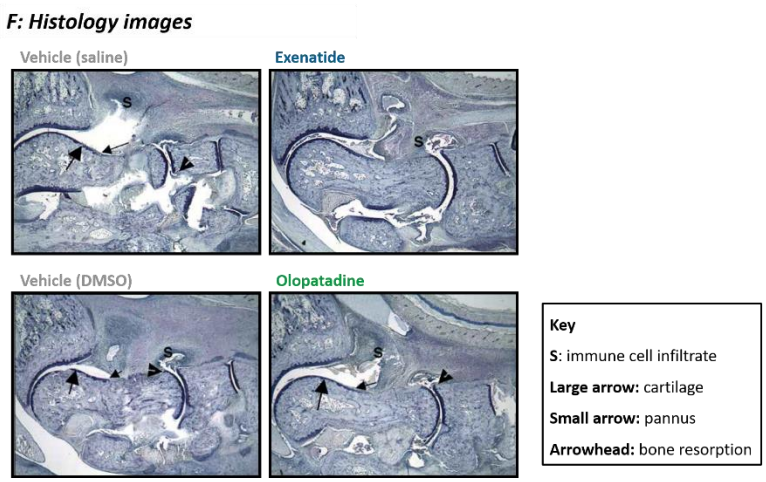


E: Histology Scores



C. Hind-limb inflammation. Hind-limbs were scored daily based on commonly used criteria, with 0 denoting no inflammation, and 4 denoting severe inflammation and swelling. Scores for left and right limbs for each animal were averaged.

D. Ankle size. Ankles were measured by caliper in triplicate on the indicated days. Sizes of left and right ankles for each animal were averaged, then normalized to ankle size at day 0.

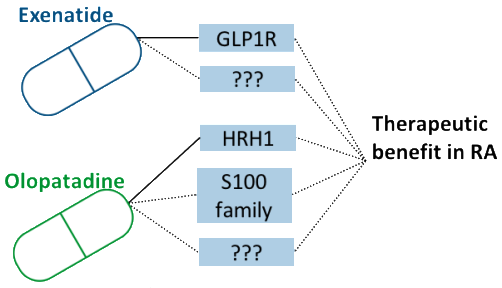


E. Histology scores. Scores for several histology parameters assayed for left ankle joints (inflammation, pannus, cartilage degradation, bone resorption, and periosteal bone formation) were summed for each animal in the indicated treatment group.

F. Histology images. Representative histological sections from the left ankle joints of animals sacrificed at the end of the study (day 17)

In all graphs, gray data points represent vehicle controls (saline for exenatide and 1% DMSO in saline for olopatadine). Error bars represent standard error of the mean. P-values were calculated using 2-way ANOVA with Tukey's multiple comparison test (ankle size, digital arthritis index) or Kruskall-Wallis with Dunn's correction (inflammation scores). * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$, and **** denotes $p < 0.0001$.

Conclusions & next steps



Exenatide and olopatadine do not act on common therapeutic targets for RA. In type 2 diabetes, the therapeutic target of exenatide is the glucagon-like peptide 1 receptor[7]. In allergic conjunctivitis, olopatadine is a histamine H1 receptor antagonist and mast cell stabilizer[8]. It also may interact with the S100 family of proteins, which have been linked to calcium signaling, immune function, and inflammation[8]. Results in the CIA model suggest that exenatide and olopatadine improve endpoints by acting on either these known targets in their approved indication, or yet-to-be-discovered off-targets, that are being engaged in a way that is therapeutic for RA. Future work will focus on identifying the target(s) of these candidates that lead to improved outcomes in animal models of RA.

References

1. Rosenthal, K. S. *et al.* (2015). Rheumatoid arthritis vaccine therapies: perspectives and lessons from therapeutic ligand epitope antigen presentation system vaccines for models of rheumatoid arthritis. *Expert Rev Vaccines* 14, 891-908.

2. Law, V., Knox, C., Djoumbou, Y., Jewison, T., Guo, A. C., Liu, Y., et al. (2014). DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res.*, 42(Database issue), D1091–7.

3. Zhu F, Shi Z, Qin C, Tao L, Liu X, Xu F, Zhang L, Song Y, Liu XH, Zhang JX, Han BC, Zhang P, Chen YZ. (2012) Therapeutic target database update 2012: a resource for facilitating target-oriented drug discovery. *Nucleic Acids Res*. 40(D1): D1128-1136.

4. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A. (2013). NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res.* 41(Database issue):D991-5.

5. Sugaya, N., & Furuya, T. (2011). Dr. PIAS: an integrative system for assessing the druggability of protein-protein interactions. *BMC Bioinformatics*, 12(1), 50.

6. Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S. & Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* 102, 15545-15550.

7. McCormack, P. L. (2014). Exenatide Twice Daily: A Review of Its Use in the Management of Patients with Type 2 Diabetes Mellitus. *Drugs*, 74(3), 325–351.

8. Kishimoto, K., Kaneko, S., Ohmori, K., Tamura, T., & Hasegawa, K. (2006). Olopatadine Suppresses the Migration of THP-1 Monocytes Induced by S100A12 Protein. *Mediators of Inflammation*, (4), 1–5.