Novel drug candidates for the treatment of metastatic colorectal cancer through global inverse gene expression profiling

Vera van Noort¹,²,*, Sebastian Schölch³,*, Murat Iskar¹,*, Georg Zeller¹, Kristina Ostertag⁴, Christine Schweitzer³, Kristin Werner³, Jürgen Weitz³, Moritz Koch³,†, Peer Bork¹,⁵,†

* these authors contributed equally to the work
† to whom correspondence should be addressed

¹ Structural and Computational Biology Unit
European Molecular Biology Laboratory
Meyerhofstrasse 1
D-69117 Heidelberg, Germany

² Centre of Microbial and Plant Genetics
KU Leuven
Kasteelpark Arenberg 22
B-3001 Leuven, Belgium

³ Department of Gastrointestinal, Thoracic and Vascular Surgery
Medizinische Fakultät Carl Gustav Carus
Technische Universität Dresden
Fetscherstr. 74
D-01307 Dresden, Germany

⁴ Department of General, Gastrointestinal and Transplant Surgery
University Hospital Heidelberg
University of Heidelberg
Im Neuenheimer Feld 110
D-69120 Heidelberg, Germany

⁵ Max-Delbrück-Centre (MDC) for Molecular Medicine
Robert-Rössle-Str. 10
D-13092 Berlin, Germany

The authors of this manuscript have no conflict of interest.

Running title:
New drug candidates for metastatic CRC
ABSTRACT

Drug-induced gene expression profiles that invert disease profiles have recently been illustrated to be a starting point for drug-repositioning. In this study, we validate this approach and focus on prediction of novel drugs for colorectal cancer (CRC), for which there is a pressing need to find novel anti-metastatic compounds. We computationally predicted three novel and still unknown compounds against CRC; citalopram (an anti-depressant), troglitazone (an anti-diabetic) and enilconazole (a fungicide). We verified the compounds by in vitro assays of clonogenic survival, proliferation and migration and in a subcutaneous mouse model. We found evidence that the mode-of-action of these compounds may be through inhibition of TGF-β signaling. Furthermore, one compound, citalopram, reduced tumor size as well as the number of circulating tumor cells and metastases in an orthotopic mouse model of CRC. This study proposes citalopram as a potential therapeutic option for CRC patients illustrating the potential of systems pharmacology.

INTRODUCTION

Rational drug discovery has gained momentum by the recently increased availability of large-scale datasets on biological activities of small molecules. Integration and utilization of these heterogeneous resources requires computational tools, as well as expert knowledge to guide the drug discovery or repositioning process. To enable the latter, a pioneering systems pharmacology study built the Connectivity Map (CMap), a collection of genome-wide gene expression readouts of cell-lines treated with more than 1,000 drug-like chemicals [1]. This approach is fundamentally based on the observation that, in many cases, drugs with similar mechanism of action in vivo also elicit similar expression responses in this in vitro model system, thus providing a starting point for drug repositioning based on expression profile similarity [1, 2]. Moreover, it was proposed that a drug-induced gene expression profile that (partially) inverts a disease-associated expression profile could hint at a potential treatment for the respective disease and indeed, by applying this concept, novel compounds have been
recently proposed for inflammatory bowel disease, non-small cell lung cancer and muscle atrophy [3-6]. Thus, the repositioning of drugs based on specifically tailored gene signatures has been demonstrated for these diseases. However, although an enrichment of drug-disease relations has been found among anti-correlated profiles when all diseases are considered together, it remains to be investigated how well this approach works for each individual disease where an expression profile is available.

A systematic analysis requires addressing various biases that are inherent to the CMap data, such as batch effects which can be partially removed with a normalization procedure [7]. Furthermore, to predict new treatment candidates for a disease, the disease itself needs to be clearly defined and characterized at the molecular level. Finally, a robust gene expression profile, associated with the disease that is also relevant for the disease etiology is necessary to make accurate predictions.

Here we present a global, unbiased approach for using drug-induced gene-expression profiles to retrieve novel candidate drugs with many cancer profiles being inversely matched. Our analysis shows a particularly strong signal for colorectal cancer (CRC) which is the third most common cancer in western countries [8]. CRC patients with non-metastatic disease have excellent long-term survival as in most cases localized primary tumors can be surgically removed. However, survival rates drop significantly with the occurrence of distant metastases. Although hepatic metastases can be surgically removed in a curative intention, thereby providing long-term cure in a significant number of patients, metastatic disease often recurs, and patients succumb to recurrent metastatic disease or primarily unresectable multifocal metastases [9, 10]. The fate of CRC patients is therefore closely linked to the occurrence of distant metastases. Although the clinical significance of metastasis in CRC is evident, the process of metastasis is still poorly understood [11]. It is widely accepted that circulating tumor cells play a pivotal role in distant tumor dissemination and their detection is closely linked to the prognosis of CRC patients [12-16]. To actively leave the tumor, enter the blood stream, attach to the vessel endothelium in the
target organ, invade, proliferate and form new lesions, abilities far beyond those of a bulk tumor cells (e.g., migration and invasion capacities) are essential for metastatic tumor cells. In that respect, to reduce the risk of (further) tumor cell dissemination in CRC patients, especially in the neoadjuvant situation (e.g., in locally advanced rectal cancer requiring neoadjuvant treatment or the "liver first" approach in synchronously metastasized CRC [17]), there is a huge clinical need for (novel) anti-metastatic drugs.

We illustrate the power of our approach by experimentally validating three predicted novel drugs that could be repositioned for CRC. We specifically test the three predicted drugs for their potential to inhibit metastasis. We confirm in vitro that these drugs inhibit cell migration. Secondly, we validate that they inhibit tumor growth in vivo in a subcutaneous mouse model. Finally, using an orthotopic mouse model of CRC, we confirm the anti-metastatic properties of citalopram in vivo. Taken together, from a broad and comprehensive bioinformatics approach, via various computational and experimental filtering and validation procedures, we arrived at a very specific, clinical relevant indication in the context of CRC and propose that the antidepressant drug citalopram can serve at least as a lead for treatment of metastatic CRC.

MATERIALS AND METHODS

Connectivity Map Data

The Connectivity Map (CMap, build 02) is a large-scale microarray resource cataloguing transcriptomic responses of four human cell lines (promyelocytic leukemia HL60, breast adenocarcinoma MCF7, prostate cancer PC3 and skin melanoma SKMEL5 cell lines) to various small-molecule treatments (referred to as drugs for simplicity). It contains 6,100 treatment instances for 1,309 unique drugs, of which ca 650 are "Food and Drug Administration" (FDA) approved[1, 18]. To eliminate various biases in CMap (e.g., batch effect, microarray platforms), we filtered and normalized the raw CMap data set (downloaded from http://www.broadinstitute.org/cmap/) as described in detail previously [7]. For this study, we
retrieved 4,849 drug-induced gene expression profiles from three main cell lines (HL60, MCF7 and PC3) representing 1,144 distinct drugs (for a detailed list, refer to Table S3).

**Disease associated microarrays**

Microarray experiments on specific diseases were identified in the public microarray repository of NCBI GEO (http://www.ncbi.nlm.nih.gov/geo/) [19] using disease medical subject heading (MeSH) terms that were attributed to each study [20-22]. In this study, we restricted our analysis to disease MeSH terms that are associated with at least one CMap drug. Moreover, we only analysed raw datasets conducted on HG-U133A and HG-U133-Plus2 microarray platforms to maximize comparability across resources. Next, all individual samples were manually annotated as "disease" or "healthy controls". In order to obtain reliable disease signatures, microarray samples (disease vs healthy) were retained only if derived from disease-relevant primary tissues. The RMA procedure was employed to normalize each dataset separately [23]. In cases, where multiple studies were found for the same disease, we selected one representative per disease, such that distinction between disease and healthy samples is maximized based on inter-array Pearson's correlation (highest area under curve in receiver operating characteristic analysis), and discarded the others. Finally, in each disease-related gene expression profile, only probe sets present in HT_HG-U133A (HT Human Genome U133 Array Plate Set, 22277 probe sets) were kept and ranked according to their fold change. In total, we obtained here a total of 40 disease-associated gene expression profiles from 1,748 individual microarray samples (for details refer to Table S1).

**Similarity score of drugs and diseases**

To calculate profile similarity of a drug with a disease, we adapted a profile comparison method initially developed for drug-drug comparisons [7]. First, a disease-related signature was defined from the top and bottom 250 ranked probe sets of drug-associated gene expression profiles. We then computed disease-drug similarity by querying the up and down regulated signatures of the disease within drug-induced gene expression profiles (HT_HG-U133A, 22277 probe sets without
Present/Absent call filtering) using gene set enrichment analysis (GSEA, weighted) [24]. At last, two enrichments scores for up/down disease signatures were averaged to derive a final score for disease-drug similarity. We assessed whether known drug-disease associations (indications, excluding ‘antidotes’ from http://www.drugs.com/medical_conditions.html update Oct 7th, 2013) are significantly anti-correlated using a nonparametric Wilcoxon signed-rank test. To this end, similarity scores from each cell line were ranked in the range of [0, 1] and pooled together. For each disease, we compared the similarity scores of known drug associations to all other drugs present in CMap (Figure S1 and Table S2). Significantly anti-correlated diseases were defined as P-value < 0.05 after false discovery rate (FDR) correction for multiple testing.

Metastatic colon cancer gene signature

To identify novel therapeutics against metastatic CRC, we utilized a consensus gene signature defined by Jorissen et al.[25]. From two independent sets of colorectal cancers, a metastatic signature was derived that showed consistent expression changes between early stage (A) and metastatic stage (D) tumors. In total there were 86 stage A tumors, 91 stage D primary tumors and 30 stage D metastases. The signature was derived from consistent changes between three comparisons; i) early stage (A) (n=44) and metastatic stage (D) (n=61) primary tumors from the first independent set ii) early stage (A) (n=42) and metastatic stage (D) (n=32) primary tumors from the second set iii) early stage (A) (n=42) primary tumors and metastatic stage (D) (n=30) metastases from the second set. The signature consists of one hundred sixty-three probe sets representing in total one hundred twenty-seven unique genes (73 up and 54 down-regulated)(Table S3). Probe sets not present in CMap dataset (HT_HG-U133A platform) were removed from the signature. After filtering, the metastasis signature was compared against the CMap reference collection that contains 4,849 gene expression profiles from three cell lines treated with 1,144 distinct drugs. We used the scoring approach described above to quantify profile similarity between drugs and the metastasis signature. For each drug, a final score was obtained by averaging the enrichment scores across all replicates from three cell lines (Table
S4). We further examined the top 20 drugs with lowest scores (anti-correlated) and manually searched in the literature whether these drugs were previously associated with colon cancer or metastasis.

**Cell lines and reagents**

The colorectal cancer cell lines HCT 116 and HT-29 were obtained from ATCC (Manassas, VA, USA) and maintained in DMEM (PAA, Piscataway, NJ, USA) + 10% FCS (PAA), 100 U/mL penicillin (PAA) and 100 μg/mL streptomycin (PAA) in a humidified atmosphere of 5% CO₂ at 37°C. The cell lines were tested for authenticity by DSMZ (Braunschweig, Germany) on a regular basis. Enilconazole (Imazalil) and troglitazone were obtained from Sigma-Aldrich (St. Louis, MO, USA) and Neochema (Bodenheim, Germany), citalopram was obtained from Lundbeck (Hamburg, Germany).

**MTT Proliferation Assay**

The WST-1 reagent (Roche, Mannheim, Germany) is reduced by metabolically active cells to formazan, an insoluble purple dye which can be measured in a spectrophotometer, and was used according to the manufacturer’s instructions. Cells were seeded in quadruplicates in increasing cell numbers in 96-well plates (125/well, 250/well, 500/well, 1000/well). After 12h adhesion time the cells were treated with the compounds as indicated. 120 h later 10 µl of WST-1 reagent were added to each well and the cells were again incubated for 4h at 37°C and 5% CO₂. After incubation and 1 min shaking, the absorbance was measured on a Genios Microplate Reader (Tecan, Männedorf, Switzerland).

**Clonogenic Survival Assay**

Clonogenic survival assay was done as previously described [26, 27]. In brief, tumor cells were plated in triplicates to yield 50 to 100 colonies per culture flask, treated with the compounds in
various concentrations and incubated for 14 to 21 days. Colonies of more than 50 cells, as assessed by microscopic inspection, were scored as survivors and counted.

**Migration Assay**

The migration of tumor cells under treatment was measured via transwell assays as described previously [26, 27]. Briefly, cells were incubated with the compounds in various concentrations for 2 hours. Then 200 µl of cell suspension (3 × 105 cells/mL) were added in triplicate transwells (Matrigel-coated transwell inserts (8 µmol/L pore size; Becton Dickinson)). After 18 hours of incubation, migratory tumor cells had invaded the underside of the membrane, were fixed, stained in thiazine and eosine solution and sealed on slides. Quantification was done by microscopic counting.

**TGF-β signaling assay**

To demonstrate the effects of the three compounds on the TGF-β signalling pathway, we utilized a cell line stably expressing a secreted alkaline phosphatase (SEAP) reporter inducible by SMAD3/4-inducible elements (SBE), called HEK-Blue TGF-β sensor cell line (Invivogen, San Diego, CA, USA). This way, downstream activation of the TGF-β signalling pathway can be measured spectrophotometrically in the cell culture supernatants. The effects of the three compounds on TGF-β signaling pathway were analyzed using the HEK-Blue TGF-β kit (Invivogen, San Diego, CA, USA) according to the manufacturer’s instructions.

**Animal experiments**

All animal experiments strictly adhered to local and federal regulations as well as FELASA guidelines, and were approved by the local authorities prior to initiation.

For the subcutaneous (s.c.) experiments, 10⁷ tumor cells in 100 µl of PBS were injected bilaterally into the flanks of NOD.Cg-Prkdc<sup>cscid</sup> Il2rg<sup>tm1Wjl</sup>/SzJ (NSG) mice. When tumors reached a size of 10-12 mm in largest diameter, the mice were euthanized, the tumors excised, and cut
into tumor fragments of 1x1 mm size. The tumor fragments were implanted subcutaneously into recipient NSG mice and monitored thrice weekly for tumor growth. Once tumors were established, growing and had a size of 7-8 mm in largest diameter, the treatment was initiated.

For orthotopic tumor cell injection, the mice were anesthetized by isoflurane inhalation. A midline laparotomy was performed and the cecum was exteriorized and 10⁵ HCT 116 tumor cells in 20 µl Matrigel (BD, Franklin Lakes, NJ, USA) were injected subserosally by help of a microinjection pump (WPI, Sarasota, Fl, USA) under microscopic visual control. After injection, the abdomen was closed with PDS II 6-0 (J&J Ethicon, Somerville, NJ, USA) and wound clips.

The treatment regimens were as follows:

Citalopram: 60 mg/kg BW i.p. daily, dissolved in 100 µl NaCl 0.9% in the s.c. experiment. Due to low tolerability (significant weight loss due to loss of appetite), the dose was reduced to 20 mg/kg i.p. daily for the orthotopic experiment.

Enilconazole: 20 mg/kg BW i.p. daily, dissolved in 20 µl DMSO

Troglitazone: 30 mg/kg BW i.p. every third day, dissolved in 100 µl 9% Solutol HS 15

Quantification of circulating tumor cells

35 days after orthotopic tumor cell injection, the mice in the control group were moribund and the experiment was terminated. The mice were anesthetized by isoflurane inhalation, blood samples were taken via intracardiac puncture and the mice were euthanized. The blood was layered over LSM 1077 lymphocyte gradient medium (PAA, Piscataway, NJ, USA) and the PBMC fraction was enriched according to the manufacturer’s instructions. The resulting PBMC were stained with Alexa-Fluor488-anti-human EpCAM antibody (Biolegend), EpCAM-positive cells were identified and counted under a fluorescence microscope (Leica). The number of CTC per mL of blood was calculated according to the total blood volume obtained by cardiac puncture.

Histological work-up
Upon euthanasia, the livers and lungs of the animals were removed and immediately fixed in 4% PFA. After paraffin embedding, the FFPE tissue was cut in 10 µm (liver) or 2 µm (lung) sections using a Leica microtome (Wetzlar, Germany). Lungs were cut in a representative layer in the coronal plane; livers were cut in three representative layers (cranial third, middle third, caudal third) in the transversal plane to be able to detect the majority of metastatic lesions. The tissue sections were stained with a standard H/E protocol and metastases were counted.

Statistical analysis

Datasets were screened for outliers with the ROUT method (Q = 10%) [28], a total of 4 statistically significant outliers were identified (all in the orthotopic citalopram experiment: 1 among the tumor weight values, 2 among the tumor volume values, 1 among the CTC number values). The findings were checked for biologic validity and the values were excluded when the cause of the deviation was considered artificial (e.g., faulty measurements). 3 out of 4 outliers were found to be biologically invalid values and excluded from further analysis; the outlier among the CTC values was considered biologically possible and was therefore not excluded. Student’s t test was used to compare means. In case of time-series, only the last time-point was tested. All analyses were two-tailed; p < 0.05 was considered statistically significant. Key: * p < 0.05; ** p < 0.01; *** p < 0.001

PI/Annexin-V apoptosis assay

5x 10^4 (HT-29) or 3x 10^4 (HCT116) CRC cells per well were seeded in 24 well plates, allowed to adhere for 12 hours and then treated with the compounds, negative or positive controls as indicated. 72 hours after incubation, the cells were harvested with trypsin and washed twice with FACS binding buffer (10mM Hepes pH 7.4, 2.5mM CaCl2, 140 mM NaCl). Then the cells were stained with Propidium iodide und Annexin V-FITC (Becton Dickinson) and 10,000 cells were analyzed and quantified by flow cytometry on a FACS Calibur (Becton Dickinson). Annexin-positive cells were considered apoptotic and the ratio apoptotic cells / total cells was calculated.
RESULTS

Reverse profile searches for retrieval of drug-disease associations

To evaluate the general validity of the inverse signature approach, we assessed how well the known disease-drug indications were recapitulated by the expression-based inverse correlation of disease-drug relations for individual diseases. To this end, we first normalized the CMap dataset to obtain 4,849 drug-induced gene expression profiles from three main cell lines (HL60, MCF7 and PC3) treated with 1,144 distinct drugs [7]. In addition, we extracted disease-associated microarray experiments from publicly-available NCBI GEO microarray repository using Medical Subject Heading (MeSH) descriptors [19]. For this analysis, we next generated disease-associated gene expression profiles for 40 distinct diseases (details in Materials and Methods).

Indeed, in the normalized CMap data significant anti-correlations were found between drug and disease gene expression profiles for known drug-disease associations for seven out of 40 diseases (Figure S1). Anti-correlations were most successful in the retrieval of disease-drug relations involving anti-cancer agents that were additionally linked to multiple cancers (six out of seven significant diseases), including colorectal cancer (Figure S1). Based on this result, we focused on colorectal cancer to predict novel drug compounds aiming specifically at an inhibition of metastasis as the latter is of considerable clinical relevance.

Prediction of novel drug compounds for metastatic CRC

As CRC scored very prominently among the general disease-drug relations and because there are insufficient treatment options, particularly for inhibiting metastasis, we chose to illustrate the power of the inverse-signature method by predicting and validating drugs used in other indications as candidates for CRC treatment. Several gene expression signatures for CRC are available, but the most pressing need is for drugs that inhibit metastasis which could also be used as an adjuvant therapy. As opposed to previous studies that have used healthy versus
diseased states [29], we rather chose to compare non-metastatic tumors versus metastatic primary tumors and metastases, to define a gene signature of metastatic potential rather than a general cancer signature. Several groups have studied the gene expression differences between primary and metastatic colon cancers to explain the molecular basis of the metastatic process and predict clinical outcomes [25, 30-32]. In this study, we decided to employ the reproducible gene expression signature from [25] that emerged as a consensus from two independent comparisons of Dukes’ stage D versus stage A colorectal cancers (Table S3, for details see Materials and Methods). This signature is based on primary and metastatic tumors from patients in different stages of colorectal cancer and includes genes that have been implicated in the development of metastasis before. We predict novel anti-metastatic drugs by finding anti-correlated drug-induced gene expression profiles, i.e. we are looking for genes that are upregulated in the metastasis profile, but downregulated by adding a drug to one of the three CMap cell lines (Figure 1).

The validity of the procedure is illustrated by the fact that among the top chemicals with inverse drug-induced gene expression profiles, compared to the metastatic colon cancer gene signature (Figure 1, Table S4), were verteporfin [33], resveratrol [34] and novobiocin [35]; all of which have already been shown to inhibit migration and/or the metastatic development of cancer cells (Table 1, [33-35]). Upon drug treatment by any of these three drugs (Figure 2), fibronectin 1, collagen type V, notch homolog 3, platelet derived growth factor C, versican, integrin beta 5 and secreted phosphoprotein 1 are down-regulated. Four of the proteins encoded by these genes are involved in interactions between the extracellular matrix (ECM) and ECM-receptors (Figure S2). The proteins fibronectin 1 and its receptor integrin beta 5 together are vital for cell migration [36]. This confirms the metastatic CRC gene signature to be appropriate for our purpose. The candidate drugs found by inverse profile search do not fully overlap in the set of genes whose expression is inverted relative to the disease profile. Thus, to find the most relevant drug candidates, we refined our list of top 20 candidate compounds by hierarchical clustering of the drug-induced gene expression data from individual experiments in
different cell lines and batches (Figure 1) and continue with those that cluster together (Figure 2). The clustering shows that some marketed drugs such as enilconazole (antifungal), citalopram (antidepressant) and troglitazone (antidiabetic) induce regulatory responses similar to verteporfin and resveratrol (Figure 2).

The best scoring candidate drugs, namely enilconazole, citalopram and troglitazone, were selected for further experimental validations. Among these, only troglitazone (an anti-diabetic and anti-inflammatory drug) has been previously associated with decreased migratory behaviour of various cell types [37-39] but not with colorectal cancer. For example, troglitazone has been screened in breast cancer metastasis. It has been reported that troglitazone reduces the migration, adhesion, and spreading of human breast cancer cells on fibronectin (FN)-coated plates independent of its PPARγ action [37, 39] and to inhibit pulmonary metastasis of osteosarcoma [38]. Second, citalopram, an antidepressant drug of the selective serotonin reuptake inhibitor (SSRI) class is approved to be used in humans but not an obvious candidate for screening against metastatic CRC. The third candidate enilconazole is a fungicide that has only been approved for topical use in veterinary medicine. Hence, the three candidates to be validated experimentally for their efficacy against metastatic CRC are not only chemically diverse, but come with different levels of expectation.

Mode of action

To better understand the underlying basis of our predictions, we set out to find the main mode of action of the three selected compounds. There are various signalling pathways that could play a role. One of them, transforming growth factor-beta (TGF-β) exerts its effects on cell proliferation, differentiation and migration in part through its modulation of extracellular matrix components, such as fibronectin and plasminogen activator inhibitor-1 (PAI-1) [40] which we find downregulated by the candidate drugs. It has recently been shown that resveratrol inhibits the TGF-β1-induced increase in cell adhesion, migration and invasion of A549 lung cancer cells [41]. Troglitazone has also been shown to have anti-migratory properties
in glioma cells, associated with transcriptional repression of TGF-β(1-3) and their receptors I and II and with reduced TGF-β release [42] and to inhibit the TGF-β-induced epithelial-mesenchymal transition of primary alveolar epithelial cells [43]. Based on similarity in down-regulated genes (Figure 2), in particular the strong inhibition of the fibronectin 1 gene and integrin beta 5, our expectation is that our other candidate drugs would also exert their anti-migratory effect through inhibition of TGF-β signalling. To demonstrate the effects of the three compounds on the TGF-β signalling pathway, we utilized a cell line stably expressing a reporter of TGF-β signalling (see Materials and Methods). The three compounds demonstrated significant inhibition of TGF-β signalling and are likely to inhibit cell migration and invasion with the same mode of action as resveratrol (Figure 3) thus strengthening our predictions based on a global computational screen.

**Experimental validation of citalopram, troglitazone and enilconazole against metastatic CRC**

We verified the predictions by in vitro tests showing that the three compounds troglitazone, citalopram and enilconazole significantly inhibited cell migration and clonogenic survival of HCT 116 and HT-29 human CRC cells, both important hallmarks of metastasis (Figure 4). While we were specifically searching for anti-metastatic compounds, we additionally sought to avoid growth stimulatory effects on primary tumors. For this purpose, the three compounds were tested for effects on proliferation. They did in fact inhibit this process albeit to a different extent (Figure 4).

Based on the promising results of the in vitro assays, next we tested the three compounds in a subcutaneous tumor model in mice (Figure 5A)[44]. We anticipated inhibition of tumor growth as all three compounds had demonstrated anti-proliferative effects. Immunodeficient mice (NOD scid gamma) were subcutaneously implanted with tumor fragments and monitored for tumor growth. All three compounds significantly inhibited the growth of subcutaneous tumors, but enilconazole and citalopram showed the most profound effects (Figure 5A). Of these two,
citalopram was tolerated best by the mice and was investigated further. PI/Annexin-based FACS assays evaluating the fraction of apoptotic cells confirmed that the compounds were indeed inhibiting metastasis and not generally cytotoxic on tumors (Figure S3).

As subcutaneous tumor models primarily simulate growth of the primary tumor and can reproduce the metastatic cascade only to a limited extent, we next investigated citalopram in an orthotopic mouse model of colorectal cancer, which is able to reproduce the metastatic cascade and mimic the clinical course of colorectal cancer [44]. We established an orthotopic model of CRC by surgical implantation of a tumor in the cecal wall (see Materials and Methods). Mice were treated with citalopram or vehicle. After 35 days, all mice in the control group were moribund due to extensive tumor burden, while the mice in the treatment group were in significantly better condition. The mice were sacrificed and tumor volume and tumor weight were measured, which were both significantly reduced in the citalopram group as compared to the controls (Figure 5B). As a direct correlate of the metastatic activity of a tumor, we also quantified the circulating tumor cells (CTC) in the blood of the animals (Figure S4). CTC numbers were significantly reduced in the citalopram-treated mice, indicating a reduced metastatic activity of the tumors after citalopram treatment. Finally, the number of metastases in the livers and lungs were quantified and revealed a significantly reduced number of metastatic deposits in the livers of the animals (p < 0.01). The number of metastatic nodules in the lungs was also reduced by citalopram treatment, however failed to reach statistical significance (p=0.27). These results show that citalopram can inhibit both tumor growth as well as metastasis of CRC.

**DISCUSSION**

Using a systems pharmacology approach based on the integration of publicly available drug-induced gene expression data, we inferred that the anti-diabetic drug troglitazone, the antifungal drug enilconazole and the antidepressant citalopram could potentially lead to novel therapeutics for the prevention of metastasis in colorectal cancer, thereby reducing the leading
cause of death in CRC patients. Based on the known mode of action of the top correlated drugs and functional analysis of inhibited genes, we propose that the main mode of action of these compounds is via the inhibition of TGF-β signaling. We performed experimental validations for the anti-cancer activities of all three compounds using \textit{in vitro} and subcutaneous tumor models in mice and we showed the efficacy of citalopram in an orthotopic mouse model of CRC which is the clinically most relevant model to examine the effects of a compound on the process of metastasis.

In an orthotopic model, we demonstrated that citalopram had effects on both CTC numbers, the gold-standard surrogate marker for metastatic activity [45] as well as on the number of hepatic metastases, which is the clinically apparent outcome of a tumor’s metastatic activity.

Due to low tolerability in the subcutaneous model, we could not test enilconazole and troglitazone in the orthotopic CRC model. Still, the \textit{in vitro} experiments show the desired effect so that they represent promising leads to arrive at similar compounds, which are tolerated better and can be tested against metastatic CRC in the future.

 Whereas the number of hepatic metastases was significantly reduced in the citalopram treated orthotopic CRC mice, the number of pulmonary foci was reduced as well, although this effect failed to reach statistical significance. This can most likely be explained by the filter effect of the hepatic capillary bed. As previously shown, the number of circulating tumor cells is significantly reduced in the blood after having passed the liver (i.e., the hepatic veins) as compared to before passing the liver (i.e., in the portal vein, [46]). This filter effect may have weakened the impact of the treatment on pulmonary metastasis.

 The heterogeneity of cancers and especially CRC makes us hypothesize that even better novel therapies could be developed through systems pharmacology. There are many different subtypes, both genetically, such as microsatellite-stable (MSS) versus unstable (MSI) or BRAF-driven versus KRAS-driven and histologically, such as serrated versus classical CRC [47]. In the future, the sub-classification and molecular characterization of these CRC subtypes into more
specific diseases as well as the development of corresponding murine disease models that mimic the specific disease subtypes will allow for the discovery of better targeted, stratified therapies.

We conclude that our implementation of the reverse gene expression profiling concept shows very promising results for the retrieval of novel candidate drugs; in particular for the treatment of cancer-like disease as we have shown in our general analysis. Indeed, CMap employs cancer cell lines to establish drug-induced gene expression profiles where many genes involved in cancer-relevant pathways are expressed. In the LINCS consortium, more data are being generated for various types of human primary cells and cell lines that will even improve the power of this approach (http://www.lincsproject.org). More relevant cell line or tissue models will make this approach more broadly applicable to other types of diseases in addition to cancer.

**Grant Support**

This work has received funding through the CancerBiome project (European Research Council project reference 268985).

**References**


### TABLES

<table>
<thead>
<tr>
<th>Small Molecule</th>
<th>PUBCHEM</th>
<th>Cmap Cell</th>
<th>CMA</th>
<th>CMAP Experiment ID</th>
<th>GSEA score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verteporfin</td>
<td>5362420</td>
<td>MCF7</td>
<td>678</td>
<td>5500024028849050407172.F09</td>
<td>-0.58</td>
</tr>
<tr>
<td>Enilconazole</td>
<td>37175</td>
<td>HL60</td>
<td>661</td>
<td>550002403070072107987.B09</td>
<td>-0.57</td>
</tr>
<tr>
<td>N-acetyl-L-leucine</td>
<td>1995</td>
<td>HL60</td>
<td>660</td>
<td>550002403070072107987.E06</td>
<td>-0.53</td>
</tr>
<tr>
<td>Troglitazone</td>
<td>5591</td>
<td>PC3</td>
<td>727</td>
<td>550002403172100807773.C04</td>
<td>-0.47</td>
</tr>
<tr>
<td>Resveratrol</td>
<td>5056</td>
<td>HL60</td>
<td>640</td>
<td>640641112706.G01</td>
<td>-0.46</td>
</tr>
<tr>
<td>Felodipine</td>
<td>3333</td>
<td>HL60</td>
<td>634</td>
<td>550002402414212006604.F03</td>
<td>-0.45</td>
</tr>
<tr>
<td>Mycophenolic acid</td>
<td>4272</td>
<td>MCF7</td>
<td>690</td>
<td>550002403071907257.B06</td>
<td>-0.40</td>
</tr>
<tr>
<td>Citalopram</td>
<td>2771</td>
<td>MCF7</td>
<td>720</td>
<td>550002403172310087770.B03</td>
<td>-0.39</td>
</tr>
<tr>
<td>Novobiocin</td>
<td>4545</td>
<td>HL60</td>
<td>658</td>
<td>55000240307602207028.G04</td>
<td>-0.38</td>
</tr>
<tr>
<td>Hydrocortison</td>
<td>3640</td>
<td>PC3</td>
<td>714</td>
<td>550002403172310087776.H03</td>
<td>-0.37</td>
</tr>
</tbody>
</table>

**Table 1** Per cell-line best inverse signature GSEA scores [-1 to 1] for candidate and proposed drugs (in bold).
FIGURE LEGENDS

Figure 1 | Using anti-correlated profiles to find candidate therapeutic chemicals.

First a metastatic disease signature was created by comparing metastatic colorectal cancers with non-metastatic cancers. The disease signature was then used to search for reverse signatures in the Connectivity Map data by Gene Set Enrichment Analysis. The resulting candidate drugs do not necessarily have similar gene expression profiles; e.g. different subsets of up-regulated genes from the disease signature can be down-regulated in the drug-induced expression profile. Therefore, an extra refinement step was added by clustering the expression profiles of candidate drugs and selecting from the candidates, those that have similar gene profiles for further validation.

Figure 2 | Evaluation of gene expression profile clustering of candidate chemicals.

Clustering of Connectivity Map data of individual experiments of 10 refined drug candidates (most anti-correlated profiles per cell line) and top 30 anti-correlated genes. Purple up-regulated in metastatic CRC signature, blue down regulated in metastatic CRC signature. Normalized fold changes dark blue, down-regulated in drug-induced profile, yellow up-regulated in drug-induced profile.

Figure 3 | Mode of action of candidate chemicals. Normalized TGF-beta signalling activity with different concentrations of citalopram, troglitazone and enilconazole. Mean and SEMs for five replicates are indicated by symbols and error bars.

Figure 4 | Inhibition of proliferation, clonogenic survival and migration of Human HCT 116 and HT-29 Colorectal Cancer Cell lines.

Graphs represent the means and SEMs of relative proliferation, clonogenic survival rates and relative number of migrating cells at different concentrations of citalopram, troglitazone and enilconazole (see Methods). Proliferation assay: Mean and SEMs for quadruplicate wells in
increasing cell numbers/well (250, 500, 1000, 2000 cells/well) are indicated by symbols and error bars. Migration & Clonogenic survival assays: Mean and SEMs for triplicate wells are indicated by symbols and error bars.

**Figure 5 | Inhibition of In vivo CRC tumor growth and CRC metastasis.**

**A** Immunodeficient (NOD scid gamma) mice were injected subcutaneously with HCT 116 CRC and treated with citalopram, troglitazone or enilconazole or their respective vehicles (see methods). Tumor volumes were measured in both treatments and control mice as indicated. **B** NOD scid gamma mice were injected with HCT 116 CRC cells in the cecal wall and treated with either citalopram or vehicle (NaCl 0.9%). After euthanasia, the total tumor size and volumes were measured. Blood was taken from control and treated mice and circulating tumor cells counted. **C** Histology of livers and lungs of orthotopic CRC mice as in B. Exemplary pictures of H/E stainings of livers and lungs of control animals (left) and citalopram-treated animals (middle). Histograms (right) show mean numbers of liver and lung metastases in control animals and citalopram treated animals.

In s.c. experiments, the symbols and error bars represent mean and SEM of 12-15 animals / group. In orthotopic experiments, data (mean ± SEM) of 11 animals per group are depicted.
Figure 1

Metastatic disease signature

Up-regulated 71 genes
Down-regulated 57 genes

Reverse profile search

Connectivity map
Drug-induced expression profiles

Gene set enrichment analysis

Candidate drugs with similar profiles
Drug-induced expression profile clustering

Rank | Chemical name | Drug indication
--- | -------------- | -----------------
1 | Verteporfin | Photosensitizer for photodynamic therapy
2 | Resveratrol | Natural phenol
3 | Novobiocin | Antibiotic
4 | Mycophenolic acid | Immunosuppressive
5 | N-acetyl-L-Leucine | Vertigo treatment
6 | Enilconazole | Antifungal
7 | Hydrocortisone | Immunosuppressive
8 | Felodipine | Anti-hypertensive
9 | Citalopram | Antidepressant
10 | Troglitazone | Anti-diabetic
Figure 2
Figure 3

TGF-β signalling activity

- Citalopram
- Troglitazone
- Enilconazole

Normalized extinction vs. nM concentration
Figure 4
Figure 5

**A** subcutaneous HCT 116 on NOD scid gamma mice

**B** orthotopic HCT 116 on NOD scid gamma mice

**C**

- Liver
- Lung

**Control**

- Liver metastases
- Lung metastases

**Citalopram**

- Liver metastases
- Lung metastases

**Troglitazone**

- Liver metastases
- Lung metastases

**Enilconazole**

- Liver metastases
- Lung metastases

Fold tumor volume (mean ± SEM) vs. day of therapy for Citalopram, Troglitazone, and Enilconazole.

Tumor weight (g), tumor volume (mm³), and circulating tumor cells per ml blood for Control and Citalopram.

Liver and Lung metastases for Control and Citalopram.
Novel drug candidates for the treatment of metastatic colorectal cancer through global inverse gene expression profiling

Vera v van Noort, Sebastian Schölch, Murat Iskar, et al.

Cancer Res  Published OnlineFirst July 18, 2014.

Updated version  Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-13-3540

Supplementary Material  Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2014/07/21/0008-5472.CAN-13-3540.DC1

Author Manuscript  Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.