doi:10.1038/nature08506



ARTICLES

# Predicting new molecular targets for known drugs

Michael J. Keiser[1,2]*, Vincent Setola[3]*, John J. Irwin[1], Christian Laggner[1], Atheir I. Abbas[4], Sandra J. Hufeisen[5], Niels H. Jensen[5], Michael B. Kuijer[3], Roberto C. Matos[3], Thuy B. Tran[3], Ryan Whaley[3], Richard A. Glennon[6], Jérôme Hert[1], Kelan L. H. Thomas[1,7], Douglas D. Edwards[1], Brian K. Shoichet[1] & Bryan L. Roth[3,5]

**Although drugs are intended to be selective, at least some bind to several physiological targets, explaining side effects and efficacy. Because many drug–target combinations exist, it would be useful to explore possible interactions computationally. Here we compared 3,665 US Food and Drug Administration (FDA)-approved and investigational drugs against hundreds of targets, defining each target by its ligands. Chemical similarities between drugs and ligand sets predicted thousands of unanticipated associations. Thirty were tested experimentally, including the antagonism of the $\beta_1$ receptor by the transporter inhibitor Prozac, the inhibition of the 5-hydroxytryptamine (5-HT) transporter by the ion channel drug Vadilex, and antagonism of the histamine $H_4$ receptor by the enzyme inhibitor Rescriptor. Overall, 23 new drug–target associations were confirmed, five of which were potent (<100 nM). The physiological relevance of one, the drug *N*,*N*-dimethyltryptamine (DMT) on serotonergic receptors, was confirmed in a knockout mouse. The chemical similarity approach is systematic and comprehensive, and may suggest side-effects and new indications for many drugs.**

The creation of target-specific 'magic bullets' has been a therapeutic goal since Ehrlich[1], and a pragmatic criterion in drug design for 30 years. Still, several lines of evidence suggest that drugs may have many physiological targets[2–5]. Psychiatric medications, for instance, notoriously act through multiple molecular targets, and this 'polypharmacology' is probably therapeutically essential[6]. Recent kinase drugs, such as Gleevec and Sutent, although perhaps designed for specificity, modulate several targets, and these 'off-target' activities may also be essential for efficacy[7,8]. Conversely, anti-Parkinsonian drugs such as Permax and Dostinex activate not only dopamine receptors but also 5-$HT_{2B}$ serotonin receptors, thereby causing valvular heart disease and severely restricting their use[9].

## Predicting drug polypharmacology

Drug polypharmacology has inspired efforts to predict and characterize drug–target associations[10–15]. Several groups have used phenotypic and chemical similarities among molecules to identify those with multiple targets[16,17], and early drug candidates are screened against molecular target panels[18]. To predict new targets for established drugs, a previous group looked for side-effects shared between two molecules[19], whereas another group linked targets by drugs that bind to more than one of them[20]. Indeed, using easily accessible associations, one can map 332 targets by the 290 drugs that bind to at least two of them, resulting in a network with 972 connections (Fig. 1a). It seemed interesting to calculate a related map that predicts new off-target effects.

Accordingly, we used a statistics-based chemoinformatics approach to predict new off-targets for 878 purchasable FDA-approved small-molecule drugs and 2,787 pharmaceutical compounds. Unlike bioinformatics methods, which might use the sequence or structural similarity among targets, this similarity ensemble approach (SEA)[21] compares targets by the similarity of the ligands that bind to them, expressed as expectation values, adapting the BLAST algorithms[21–23] (other methods such as naive Bayesian classifiers[23,24] may also be used, see Supplementary Table 1). The approach thus captures ligand-based similarities among what would otherwise be considered disparate proteins. The 3,665 drugs were compared against 65,241 ligands organized into 246 targets drawn from the MDL Drug Data Report (MDDR) database[25], yielding 901,590 drug–target comparisons.

Most drugs had no significant similarities to most ligand sets. However, 6,928 pairs of drugs and ligand sets were similar, with expectation values (*E*-values) better than $1 \times 10^{-10}$. We analysed these predictions retrospectively against known associations and prospectively for unreported drug polypharmacology.

## Retrospective drug–target predictions

We first compared the predicted drug–target associations from the MDDR database against reported associations with affinities better than 1 μM in a second database, the World of Molecular Bioactivity (WOMBAT)[26]. For instance, the MDDR annotates Azopt (brinzolamide) only as an "antiglaucoma agent", but WOMBAT reports that it binds carbonic anhydrase II at 3 nM. Correspondingly, when screened internally against all MDDR molecular targets, SEA associated this drug with "carbonic anhydrase inhibitors" with an *E*-value of $8.3 \times 10^{-139}$. For 184 of the 746 drugs in WOMBAT, the predicted MDDR target agreed with the annotated WOMBAT target with *E*-values of $1 \times 10^{-10}$ or better, recapitulating 19% of the off-targets missing from the MDDR (Supplementary Table 2). Another 257 drug–target predictions were unannotated in either database, and may suggest new polypharmacology.

[1]Department of Pharmaceutical Chemistry, [2]Graduate Group in Bioinformatics, University of California San Francisco, 1700 4th Street, San Francisco, California 94143-2550, USA. [3]NIMH Psychoactive Drug Screening Program, Department of Pharmacology, [4]Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106, USA. [5]Department of Pharmacology and Division of Medicinal Chemistry and Natural Products, The University of North Carolina Chapel Hill School of Medicine, Chapel Hill, North Carolina 27759, USA. [6]Department of Medicinal Chemistry, School of Pharmacy, Medical College of Virginia Campus, Virginia Commonwealth University, 410 North 12th Street, PO Box 980540, Richmond, Virginia 23298-0540, USA. [7]University of Michigan Health System, 1500 East Medical Center Drive, Ann Arbor, Michigan 48109, USA.
*These authors contributed equally to this work.



©2009 Macmillan Publishers Limited. All rights reserved

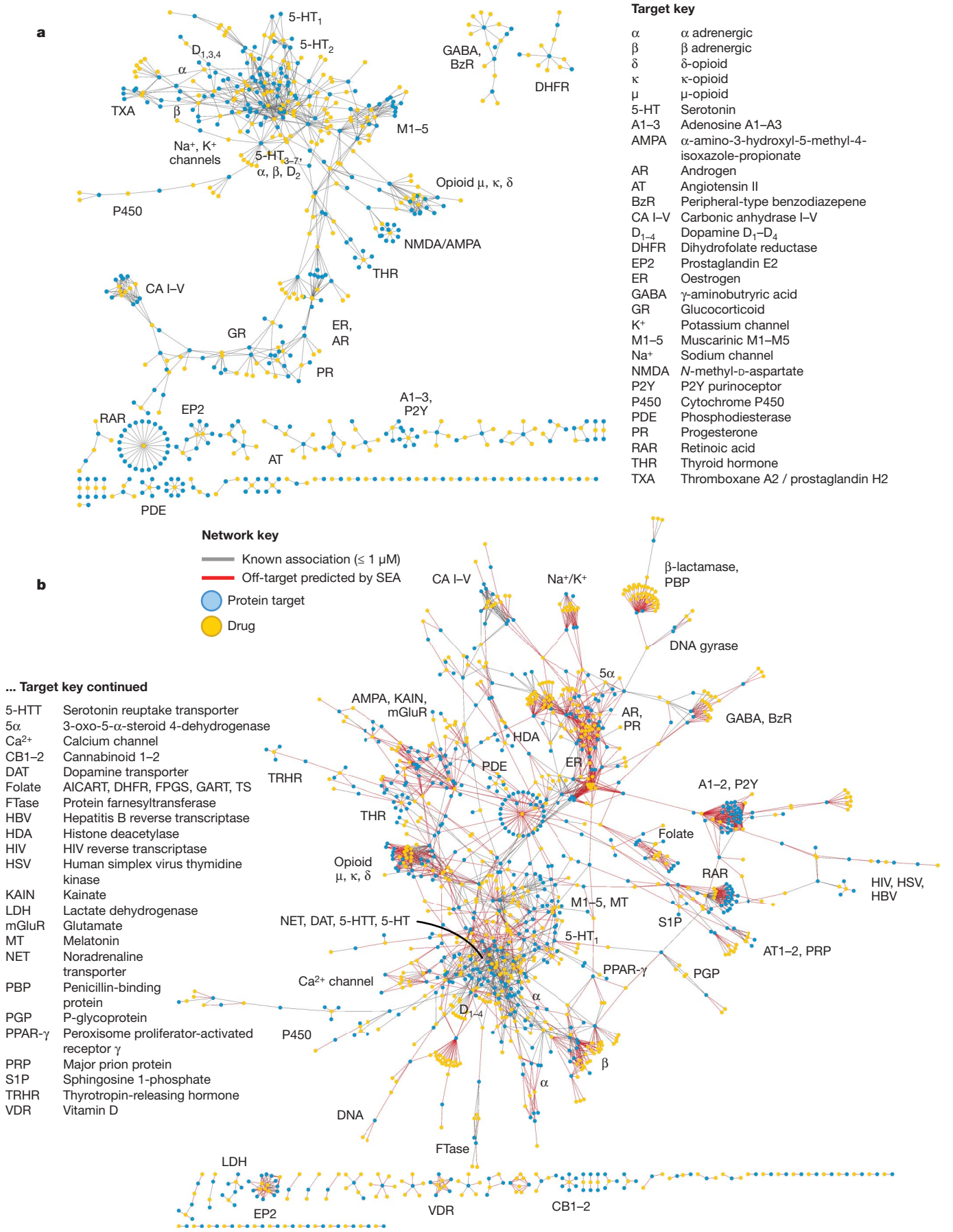



**Figure 1 | Drug–target networks, before and after predicting off-targets.** **a**, Known drug–target network. Each drug (gold) is linked to its known protein targets (cyan) by a grey edge. Each edge denotes a $K_i$ of 1 μM or better for that drug to its target. **b**, Predicted drug–target network. Drugs and proteins are linked as per the known drug–target network in **a**, but with the addition of red edges representing SEA off-target predictions with $E$-values $\leq 10^{-10}$.



©2009 Macmillan Publishers Limited. All rights reserved

A second retrospective test predicted targets for the 3,665 drugs uncharacterized in either database but known in the literature. Of the 6,928 drug off-targets predicted, we discarded 430 as highly similar by structure to known target ligands, and another 2,666 as trivial. This left 3,832 predictions, of which we inspected 184 by literature search and by interrogating other databases. Of these, 42 turned out to be known associations (Supplementary Table 3). For instance, when we screened the drug Revanil (lisuride) against the MDDR ligand–target sets, its best *E*-value was as an $\alpha_2$ adrenergic antagonist, and when we screened the drug Permax (pergolide) it had an *E*-value of $8.70 \times 10^{-29}$ as a 5-$HT_{1D}$ receptor agonist. Consistent with these predictions, Revanil has been reported to bind adrenergic $\alpha_2$ at 0.055 nM and Permax binds the 5-$HT_{1D}$ receptor at 13 nM (Supplementary Table 3), although neither activity was reported in the MDDR or WOMBAT databases.

## New drug–target predictions

For many of these 184 predictions we found no literature precedent. We therefore tested 30 predictions that were experimentally accessible to us. In radioligand competition binding assays, 23 of these (77%) yielded inhibition constants ($K_i$ values) less than 15 μM (lower $K_i$ values indicate higher affinity) (Tables 1 and 2 and Supplementary Fig. 1). Fifteen of these 23 were to aminergic G-protein-coupled receptors (GPCRs) (Table 1), and the remainder crossed major receptor classification boundaries (Table 2). For instance, the $\alpha_1$ antagonist Doralese was predicted and observed to bind to the dopamine $D_4$ receptor—both $\alpha_1$ and $D_4$ are aminergic GPCRs. Conversely, the HIV-1 reverse transcriptase (enzyme) inhibitor Rescriptor was predicted and observed to bind to histamine $H_4$ receptor (GPCR); this prediction crosses major target boundaries. For several predictions, we tested multiple receptor subtypes because the MDDR left these unspecified; for example, for a predicted '$\alpha_1$ adrenergic blocker', we tested the drug at $\alpha_{1A}$, $\alpha_{1B}$ and $\alpha_{1D}$ subtypes; we count these as a single target. In total, 14 drugs bound 23 previously unknown targets, with 13 having sub-micromolar and five having sub-100 nM affinities (Tables 1 and 2). In cases such as Doralese's, the affinity for the discovered off-target dopamine $D_4$, to which it binds with a $K_i$ of 18 nM (Fig. 2a), was better than that for its known therapeutic targets, $\alpha_{1A}$ and $\alpha_{1B}$ adrenergic receptors, for which its $K_i$ values are 611 and 226 nM, respectively[27].

The question arises as to how interesting and biologically relevant these new off-targets are. This can be evaluated by the following criteria: when the new targets contribute to the primary activity of the drug, when they may mediate drug side effects, or when they are unrelated by sequence, structure and function to the canonical targets. Although not all of the newly predicted off-targets fall into these three categories, several fall into each.

## New targets as primary sites of action

The new targets can improve our understanding of drug action. DMT is an endogenous metabolite and a notorious hallucinogen. Recently, the molecule was characterized as a $\sigma_1$-receptor regulator at micromolar concentrations, an association implicated in its hallucinogenic properties[28,29]. This surprised us because many drugs, including non-hallucinogens, bind promiscuously to the $\sigma_1$ receptor with higher affinity than DMT[30]. Also, the hallucinogenic characteristics of DMT are consistent with other hallucinogens thought to act through serotonergic receptors, some of which the molecule is known to bind[31–33]. We therefore screened DMT against the 1,133 WOMBAT targets. SEA predicted it to be similar against multiple serotonergic (5-HT) ligand sets, with *E*-values ranging from $9.2 \times 10^{-81}$ to $7.4 \times 10^{-6}$. Upon testing, we find DMT binds 5-$HT_{1A}$, 5-$HT_{1B}$, 5-$HT_{1D}$, 5-$HT_{2A}$, 5-$HT_{2B}$, 5-$HT_{2C}$, 5-$HT_{5A}$, 5-$HT_6$ and 5-$HT_7$ receptors with affinities ranging from 39 nM to 2.1 μM (Supplementary Table 4 and Supplementary Fig. 2). Of these, three were previously unknown (Table 1), and all had substantially greater affinities for DMT than that represented by its 14.75 μM dissociation constant ($K_d$) for $\sigma_1$ (ref. 28). To investigate further the role of serotonin receptors in DMT-induced hallucination, we turned to a cell-based assay and an animal model that are predictive of hallucinatory actions[34]. Consistent with SEA prediction, we find that DMT is not only a potent partial agonist at 5-$HT_{2A}$ (Fig. 2g) as has been reported[31], but it also induces head twitch response in wild-type but not in 5-$HT_{2A}$ knockout mice (Fig. 2h), which is new to this study. The half-maximum effective concentration ($EC_{50}$) of DMT at 5-$HT_{2A}$ is 100-fold lower (better) than that observed for $\sigma_1$ (ref. 28). These observations support 5-$HT_{2A}$ as the primary target for DMT's hallucinogenic effects.

Similarly, the new off-targets for Sedalande—a neuroleptic and anxiolytic derived from haloperidol—may illuminate this drug's therapeutic effects. Although used in psychiatric clinical trials as far back as the early 1960s[35], neither its mechanism of action in the central nervous system (CNS), nor that of the related Dimetholizine, is well understood. In addition to new activities against $\alpha_1$ adrenergic receptors (1.2–240 nM; Fig. 2b and Table 1), Dimetholizine was found to bind the $D_2$ and 5-$HT_{1A}$ receptors, and Sedalande to bind the 5-$HT_{1D}$ receptor (Table 1 and Supplementary Fig. 1). This probably contributes to the CNS activity of both drugs, given the association of the former with anxiety and aggression modulation, and the activity of many antipsychotics against the $D_2$ receptor. We also found analogues of Sedalande that were active against 5-$HT_{1D}$, often at affinities comparable to or greater than those of Sedalande itself (Supplementary Table 5 and Supplementary Fig. 3). This supports the possibility of optimizing these drugs for new indications.

An example of a drug now being investigated for a new indication is Fabahistin. Used since the 1950s as a symptomatic antihistamine, Fabahistin is now being investigated for Alzheimer's disease. When screened against 1,133 WOMBAT targets, SEA found an extraordinary similarity to 5-$HT_{5A}$ ligands, with an *E*-value of $2.0 \times 10^{-58}$. When we measured its binding to the 5-$HT_{5A}$ receptor, Fabahistin had a $K_i$ of 130 nM (Fig. 2c and Table 1). This is another example of a drug whose new, off-target affinity is much better than that for its canonical $H_1$ receptor target[36]. Its activity against 5-$HT_{5A}$ and related serotonergic receptors[37] may have implications for Fabahistin's role as an Alzheimer's disease therapeutic.

## Off-targets as side-effect mediators

Some of the new off-targets may contribute to a drug's adverse reactions. Motilium is an antiemetic and dopamine $D_{1/2}$ antagonist that achieves peak plasma concentrations of 2.8 μM[38] after intravenous administration. This formulation was withdrawn owing to adverse cardiovascular effects, with the FDA citing cardiac arrest, arrhythmias and sudden death[39]. Although Motilium binds the hERG channel with a half-maximum inhibitory concentration ($IC_{50}$) of 5 μM[40], the 71–710 nM affinities observed here against $\alpha_{1A}$, $\alpha_{1B}$ and $\alpha_{1D}$ may also contribute to these cardiovascular effects (Fig. 2d, Table 1 and Supplementary Fig. 1).

Similarly, the micromolar activity against the β adrenergic receptors of the widely used selective serotonin reuptake inhibitor (SSRI) antidepressants Prozac and Paxil (Fig. 2e, Table 1 and Supplementary Fig. 1) may explain several of their adverse effects. The abrupt withdrawal of Paxil raises standing heart rate, a symptom of the SSRI discontinuation syndrome[41]. This is counterintuitive, as relieving blockade of serotonin reuptake should reduce synaptic serotonin, inconsistent with the cardiovascular syndrome[42]. β-blockade by these SSRIs may partially explain this effect because β-blockers induce a similar rebound tachycardia after abrupt withdrawal, owing to β receptor upregulation and sensitization. Despite its higher affinity for β receptors, Prozac has a longer half-life than Paxil, and its withdrawal does not induce SSRI discontinuation syndrome. Also, SSRIs and many β-blockers can induce sexual dysfunction[43]. Because both serotonergic and adrenergic signalling are involved in sexual response, the binding of Paxil and Prozac to the $\beta_1$-receptor may explain why they induce greater dysfunction than other SSRIs.



©2009 Macmillan Publishers Limited. All rights reserved

**Table 1 | Prediction and testing of new aminergic GPCR targets for drugs**

| Drug | Pharmacological action | *E*-value | Predicted target | $K_i$ (nM) |
|---|---|---|---|---|
| Sedalande | Neuroleptic | $8.3 \times 10^{-136}$ | $\alpha_1$ adrenergic blocker* | $\alpha_{1A}$, 1.2; $\alpha_{1B}$, 14; $\alpha_{1D}$, 7.0 |
| | | $1.7 \times 10^{-14}$ | 5-$HT_{1D}$ antagonist | 140 |
| Dimetholizine | Antihistamine; antihypertensive | $1.6 \times 10^{-129}$ | $\alpha_1$ adrenergic blocker* | $\alpha_{1A}$, 70; $\alpha_{1B}$, 240; $\alpha_{1D}$, 170 |
| | | $2.7 \times 10^{-113}$ | 5-$HT_{1A}$ antagonist | 110 |
| | | $7.4 \times 10^{-56}$ | Dopamine $D_2$ antagonist | 180 |
| Kalgut | Cardiotonic | $3.1 \times 10^{-79}$ | $\beta_3$ adrenergic agonist | $2.1 \times 10^3$ |
| Fabahistin | Antihistamine | $5.7 \times 10^{-57}$ | 5-$HT_{5A}$ antagonist | 130 |
| Prantal | Anticholinergic; antispasmodic | $5.5 \times 10^{-32}$ | δ-opioid agonist | $1.4 \times 10^4$ |
| *N*,*N*-dimethyltryptamine | Serotonergic hallucinogen | $3.1 \times 10^{-21}$ | 5-$HT_{1B}$ agonist | 130 |
| | | $1.2 \times 10^{-13}$ | 5-$HT_{2A}$ agonist† | 130 |
| | | $1.1 \times 10^{-7}$ | 5-$HT_{5A}$ antagonist | $2.1 \times 10^3$ |
| | | $5.0 \times 10^{-6}$ | 5-$HT_7$ modulator | 210 |
| Doralese | Adrenergic $\alpha_1$ blocker; antihypertensive; antimigraine | $2.8 \times 10^{-27}$ | Dopamine $D_4$ antagonist | 18 |
| Prozac | 5-HT reuptake inhibitor; antidepressant | $3.9 \times 10^{-15}$ | β adrenergic blocker* | $\beta_1$, $4.4 \times 10^3$ |
| Motilium | Antiemetic; peristaltic stimulant | $4.8 \times 10^{-11}$ | $\alpha_1$ adrenergic blocker* | $\alpha_{1A}$, 71; $\alpha_{1B}$, 530; $\alpha_{1D}$, 710 |
| Paxil | 5-HT reuptake inhibitor; antidepressant | $1.3 \times 10^{-7}$ | β adrenergic blocker* | $\beta_1$, $1.0 \times 10^4$ |

$K_i$ values are accurate ±20% at two significant figures.
* For the targets marked, the reference data set did not specify the receptor subtype, requiring a separate assay for each one. For instance, the MDDR contains an '$\alpha_1$ adrenergic blocker' set, for which it was necessary to test the $\alpha_{1A}$, $\alpha_{1B}$ and $\alpha_{1D}$ subtypes.
† 5-$HT_{2A}$ is a known target of DMT, but is shown here with its retrospective SEA *E*-value for comparison purposes.

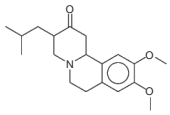
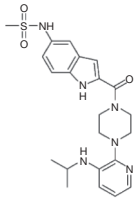
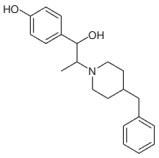
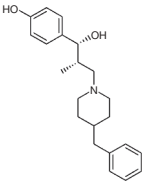
## Drug binding across major protein boundaries

Whereas many of the predicted off-targets occur among aminergic GPCRs, a target class for which cross-activity is well-known (see later)[44], four of the drugs bound to targets unrelated by sequence or structure to their canonical targets (Table 2). For instance, the reverse transcriptase (enzyme) inhibitor Rescriptor was predicted and shown to bind to the histamine $H_4$ receptor, a GPCR. These two targets share no evolutionary history, functional role, or structural similarity whatsoever. Intriguingly, although the $K_i$ value of Rescriptor for the $H_4$ receptor is high at 5.3 μM (Table 2 and Supplementary Fig. 1), this is within its steady-state plasma concentration (minimum plasma concentration averages 15 μM) and is consistent with the painful rashes associated with Rescriptor use[45]; likewise, $H_4$ dysregulation has been associated with atopic dermatitis[46]. Similarly, the vesicular monoamine transporter (VMAT) inhibitor[47] Xenazine binds two different GPCRs at sub-micromolar concentrations (Table 2 and Supplementary Fig. 1). Despite its use over the last 50 years, Xenazine has not been reported to bind to any



©2009 Macmillan Publishers Limited. All rights reserved

**Table 2 | Prediction and testing of new cross-boundary targets for drugs**

| Drug | | Canonical target | *E*-value | Predicted target | $K_i$ (nM) |
|---|---|---|---|---|---|
|  | Xenazine | VMAT2 (transporter) | $1.4 \times 10^{-61}$ | $\alpha_2$ adrenergic receptor (GPCR) | $\alpha_{2A}$, 960; $\alpha_{2C}$, $1.3 \times 10^3$ |
|  | Rescriptor | HIV-1 reverse transcriptase (enzyme) | $1.1 \times 10^{-30}$ | Histamine $H_4$ receptor (GPCR) | $5.3 \times 10^3$ |
|  | Vadilex | NMDAR (ion channel) | $5.1 \times 10^{-13}$<br>$2.0 \times 10^{-4}$ | μ-opioid receptor (GPCR)<br>5-HTT; serotonin transporter (transporter) | $1.4 \times 10^3$<br>77 |
|  | RO-25-6981 | NMDAR (ion channel) | $1.5 \times 10^{-8}$<br>$1.9 \times 10^{-6}$<br>$3.6 \times 10^{-6}$<br>$9.1 \times 10^{-5}$ | 5-HTT; serotonin transporter (transporter)<br>Dopamine $D_4$ receptor (GPCR)<br>NET; noradrenaline transporter (transporter)<br>κ-opioid receptor (GPCR) | $1.4 \times 10^3$<br>120<br>$1.3 \times 10^3$<br>$3.1 \times 10^3$ |

$K_i$ values are accurate ±20% at two significant figures. The MDDR database did not specify the $\alpha_2$ adrenergic receptor subtype, requiring a separate assay for each one ($\alpha_{2A}$, $\alpha_{2C}$).

GPCR. Finally, the selective ion-channel inhibitors Vadilex and RO-25-6981 were predicted and found to bind to GPCRs and to transporters to which they were previously unknown to bind (Fig. 2f, Table 2 and Supplementary Fig. 1). Although these ion-channel drugs have known polypharmacology (Fig. 3), a key point is that the new targets for these four drugs are unrelated to their main therapeutic targets except in the similarity of the ligands that modulate their activities.

More broadly, the protein target with the highest sequence similarity to any of a drug's known targets is rarely predicted by the SEA approach. Rather, the target predicted by ligand similarity is typically well down in the sequence-similarity ranking. Thus for Xenazine, the off-target $\alpha_2$ adrenergic receptor is the 78th most similar receptor to the known target VMAT2 and in fact has no significant similarity at all, with a PSI-BLAST *E*-value of 125 (Supplementary Table 6); and for Rescriptor, $H_4$ is the 167th most similar receptor to HIV-1 reverse transcriptase, and even for Prantal, the aminergic δ-opioid receptor is only the 26th most similar receptor to its known muscarinic $M_3$ target.

Certain caveats merit mention. Not all of the new off-targets predicted here would surprise specialists. For instance, Dimetholizine has antihypertensive activity and so its affinity for adrenergic receptors is not wholly unanticipated. Similarly, Kalgut is classified as a selective $\beta_1$ agonist, thought to have little activity on other adrenergic receptors[48]. Whereas the observation that it does bind to the $\beta_3$

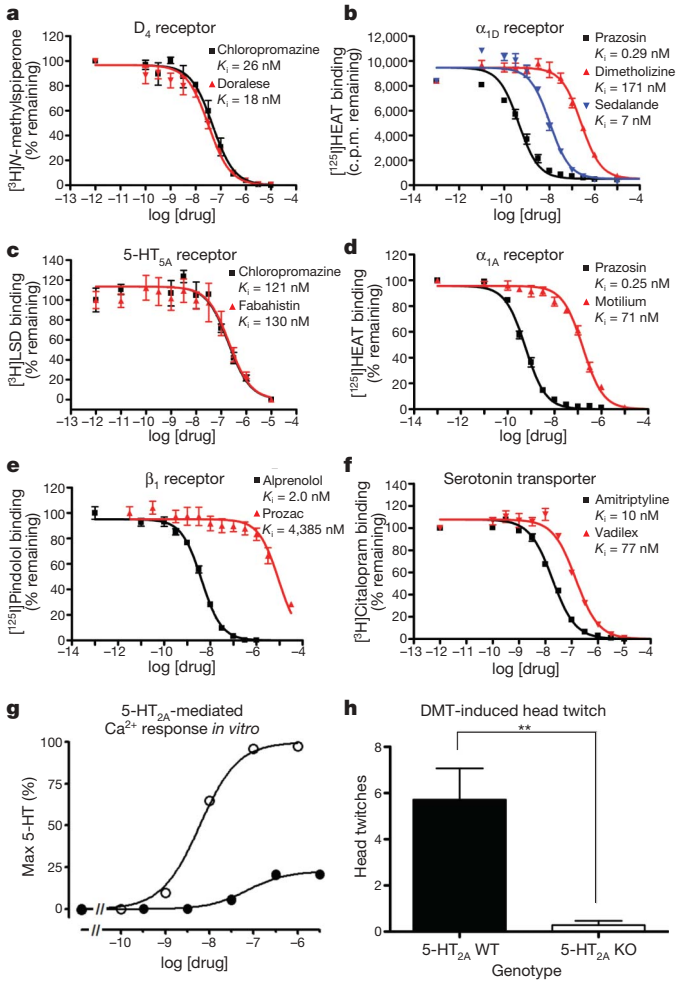



**Figure 2 | Testing new off-target activities. a–f**, Radioligand competition binding assays: Doralese at $D_4$ (**a**), Sedalande and Dimetholizine at $\alpha_{1D}$ (**b**), Fabahistin at 5-$HT_{5A}$ (**c**), Motilium at $\alpha_{1A}$ (**d**), Prozac at $\beta_1$ (**e**), and Vadilex (**f**) at the serotonin transporter. c.p.m., counts per minute; HEAT, 2-[β-(4-hydroxyphenyl)-ethyl-aminomethyl]tetralone; LSD, lysergic acid diethylamide. **g, h**, Investigating 5-$HT_{2A}$ as the target of DMT-induced hallucination: 5-$HT_{2A}$-mediated $Ca^{2+}$ response was measured after treating HEK 293 cells stably expressing the human 5-$HT_{2A}$ receptor with DMT (filled circles) or 5-HT (open circles) (**g**). The mean $EC_{50}$ ± s.e.m. of DMT was found to be 118 ± 29 nM (versus 5-HT's 6.6 ± 0.4 nM baseline, $n = 3$), with a maximum percentage relative (to 5-HT) efficacy ($E_{max}$) of 23 ± 0.4% ($n = 3$), confirming that DMT is a potent partial agonist at 5-$HT_{2A}$ receptors. DMT elicited head twitch behaviour only in 5-$HT_{2A}$ wild-type (WT) mice, confirming that it is a hallucinogenic 5-$HT_{2A}$ agonist (**h**). KO, knockout. $^{**}P < 0.01$. In **a–f**, error bars represent the s.e.m. (of three replicates) from one representative experiment; in **h**, error bars represent the s.e.m. (of seven mice per genotype).



©2009 Macmillan Publishers Limited. All rights reserved

receptor goes against this classification, structurally this seems easy to credit (Table 1 and Supplementary Fig. 1). Indeed, 10 of the 14 drugs reported here are active against aminergic GPCRs (Fig. 3), and so their cross-activities against other aminergic GPCRs have some precedent[44]. Finally, although most of the drugs were active at their predicted off-targets, one-third were not; these are examples of the false-positives to which this method is susceptible (Supplementary Table 7). Thus, the anxiolytics Valium and Centrax scored well against cholecystokinin B ligands, the antipsychotic Emilace was predicted to bind 5-$HT_4$, the anaesthetic Duocaine the κ-opioid receptor, the antihypertensive Doralese neurokinin receptors, and the narcotic Dromoran and the bradycardic Zatebradine scored well against the $D_2$ and $D_1$ receptors. None of these bound their predicted off-targets with affinities better than 10 μM. SEA ignores pharmacophores in its predictions, comparing drugs to ligand sets based on all shared chemical patterns. This is at once a strength, in that it is model-free, and a weakness, in that it may predict activity for drugs that share many features with the ligands of a target, and yet miss a critical chemotype.

### Predicting polypharmacology on a large scale

Notwithstanding these caveats, it is the model-free nature of these predictions that allows a comprehensive exploration of drug–target interactions, most of which remain unexplored. We have focused on a thin slice of pharmacological targets, one dominated by aminergic drugs (Fig. 3). Stepping back to view the larger picture, 364 additional off-targets for 158 drugs are predicted with *E*-values better than $1 \times 10^{-50}$, whereas 1,853 new off-targets are predicted with *E*-values better than $1 \times 10^{-10}$ (Fig. 1b). This compares to the only 972 off-target activities already annotated in the databases (Fig. 1a). The SEA and related chemoinformatics methods[16–20] provide tools to explore these associations systematically, both to understand drug effects and explore new opportunities for therapeutic intervention.

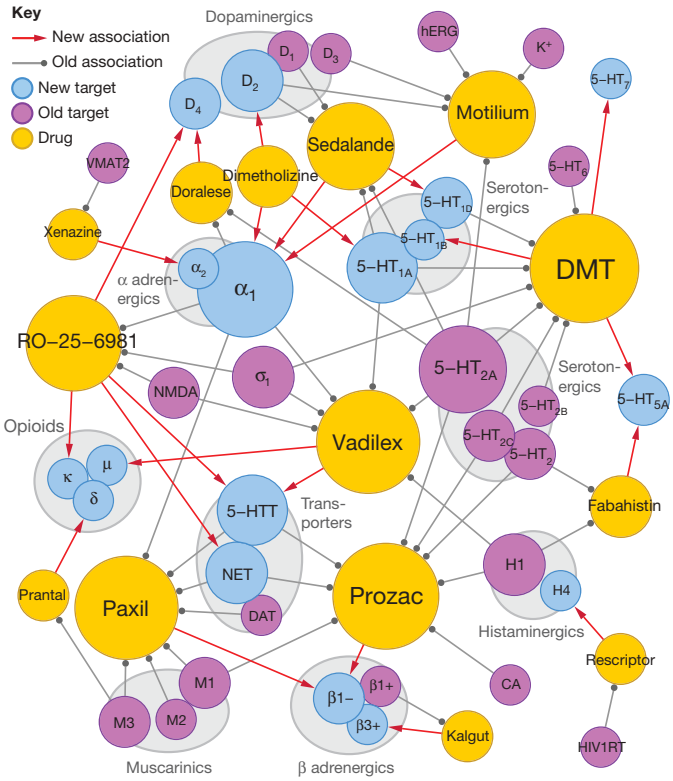



**Figure 3 | Discovered off-targets network.** Bipartite network where drugs (gold) are linked by grey edges to their known targets (violet) and by red arrows to their discovered off-targets (cyan). Grey edges denote binding at 1 μM or better, where these affinities are known. Node sizes increase with number of incident edges. Target abbreviations: 5-HT*x*, serotonin receptor type *x*; 5-HTT, serotonin transporter; β1+, $\beta_1$ adrenergic agonist; β1−, $\beta_1$ adrenergic antagonist; β3+, $\beta_3$ adrenergic agonist; $\sigma_1$, $\sigma_1$-receptor; CA, carbonic anhydrase; DAT, dopamine transporter; HIV1RT, HIV-1 reverse transcriptase; hERG, human Ether-a-go-go related gene channel; $K^+$, potassium channel; NET, noradrenaline transporter; NMDA, *N*-methyl-D-aspartate receptor; VMAT2, vesicular monoamine transporter 2.

## METHODS SUMMARY

**Prediction of off-targets.** A collection of 3,665 FDA-approved and investigational drug structures was computationally screened against a panel of more than 1,400 protein targets. The drug collection was extracted from the MDL Comprehensive Medicinal Chemistry database. Each target was represented solely by its set of known ligands, which were extracted from three sources of annotated molecules: the MDL Drug Data Report, the WOMBAT[26], and the StARlite databases. The two-dimensional structural similarity of each drug to each target's ligand set was quantified as an *E*-value using the SEA[21].

**Experimental testing.** Predicted off-targets with strong SEA *E*-values were evaluated for novelty against orthogonal databases and the literature. Those off-targets without precedent were subjected to radioligand competition binding assays using standard techniques[49] at the NIMH Psychoactive Drug Screening Program. The role of 5-$HT_{2A}$ agonism in DMT-induced hallucination was examined in cell-based and in knockout mouse models[34]. Derivatives of Sedalande were identified in the ZINC[50] database by substructure search, and their affinities for 5-$HT_{1D}$ tested using standard techniques[49].

**Drug–target networks and out-group analysis.** Comprehensive networks of known drug–target associations (by WOMBAT) and predicted off-targets (by SEA) were constructed. Furthermore, SEA off-target predictions were compared to those derived from naive Bayesian classifiers and from PSI-BLAST[21–23] comparisons of a drug's known protein target(s) against the panel of potential protein targets.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 7 April; accepted 14 September 2009.**
**Published online 1 November 2009.**

1. Ehrlich, P. The theory and practice of chemotherapy. *Folia Serologica* **7**, 697–714 (1911).
2. Peterson, R. T. Chemical biology and the limits of reductionism. *Nature Chem. Biol.* **4**, 635–638 (2008).
3. Nobeli, I., Favia, A. D. & Thornton, J. M. Protein promiscuity and its implications for biotechnology. *Nature Biotechnol.* **27**, 157–167 (2009).
4. Marona-Lewicka, D. & Nichols, D. E. Further evidence that the delayed temporal dopaminergic effects of LSD are mediated by a mechanism different than the first temporal phase of action. *Pharmacol. Biochem. Behav.* **87**, 453–461 (2007).
5. Marona-Lewicka, D. & Nichols, D. E. WAY 100635 produces discriminative stimulus effects in rats mediated by dopamine $D_4$ receptor activation. *Behav. Pharmacol.* **20**, 114–118 (2009).
6. Roth, B. L., Sheffler, D. J. & Kroeze, W. K. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. *Nature Rev. Drug Discov.* **3**, 353–359 (2004).
7. Rix, U. *et al.* Chemical proteomic profiles of the BCR-ABL inhibitors imatinib, nilotinib, and dasatinib reveal novel kinase and nonkinase targets. *Blood* **110**, 4055–4063 (2007).
8. Hopkins, A. L. Network pharmacology. *Nature Biotechnol.* **25**, 1110–1111 (2007).
9. Roth, B. L. Drugs and valvular heart disease. *N. Engl. J. Med.* **356**, 6–9 (2007).
10. Bajorath, J. Computational analysis of ligand relationships within target families. *Curr. Opin. Chem. Biol.* **12**, 352–358 (2008).
11. Oprea, T. I., Tropsha, A., Faulon, J. L. & Rintoul, M. D. Systems chemical biology. *Nature Chem. Biol.* **3**, 447–450 (2007).
12. Newman, D. J. Natural products as leads to potential drugs: an old process or the new hope for drug discovery? *J. Med. Chem.* **51**, 2589–2599 (2008).
13. Siegel, M. G. & Vieth, M. Drugs in other drugs: a new look at drugs as fragments. *Drug Discov. Today* **12**, 71–79 (2007).
14. Miller, J. R. *et al.* A class of selective antibacterials derived from a protein kinase inhibitor pharmacophore. *Proc. Natl Acad. Sci. USA* **106**, 1737–1742 (2009).
15. Walsh, C. T. & Fischbach, M. A. Repurposing libraries of eukaryotic protein kinase inhibitors for antibiotic discovery. *Proc. Natl Acad. Sci. USA* **106**, 1689–1690 (2009).
16. Young, D. W. *et al.* Integrating high-content screening and ligand-target prediction to identify mechanism of action. *Nature Chem. Biol.* **4**, 59–68 (2008).
17. Wagner, B. K. *et al.* Large-scale chemical dissection of mitochondrial function. *Nature Biotechnol.* **26**, 343–351 (2008).
18. Krejsa, C. M. *et al.* Predicting ADME properties and side effects: the BioPrint approach. *Curr. Opin. Drug Discov. Dev.* **6**, 470–480 (2003).
19. Campillos, M., Kuhn, M., Gavin, A. C., Jensen, L. J. & Bork, P. Drug target identification using side-effect similarity. *Science* **321**, 263–266 (2008).
20. Paolini, G. V., Shapland, R. H. B., van Hoorn, W. P., Mason, J. S. & Hopkins, A. L. Global mapping of pharmacological space. *Nature Biotechnol.* **24**, 805–815 (2006).



©2009 Macmillan Publishers Limited. All rights reserved

21. Keiser, M. J. *et al.* Relating protein pharmacology by ligand chemistry. *Nature Biotechnol.* **25,** 197–206 (2007).
22. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215,** 403–410 (1990).
23. Hert, J., Keiser, M. J., Irwin, J. J., Oprea, T. I. & Shoichet, B. K. Quantifying the relationships among drug classes. *J. Chem. Inf. Model.* **48,** 755–765 (2008).
24. Nigsch, F., Bender, A., Jenkins, J. L. & Mitchell, J. B. Ligand-target prediction using Winnow and naive Bayesian algorithms and the implications of overall performance statistics. *J. Chem. Inf. Model.* **48,** 2313–2325 (2008).
25. Schuffenhauer, A. *et al.* An ontology for pharmaceutical ligands and its application for in silico screening and library design. *J. Chem. Inf. Comput. Sci.* **42,** 947–955 (2002).
26. Olah, M. *et al.* in *Chemical Biology: From Small Molecules to Systems Biology and Drug Design* (eds Schreiber, S. L., Kapoor, T. M. & Wess, G.) 760–786 (Wiley-VCH, 2007).
27. Lomasney, J. W. *et al.* Molecular cloning and expression of the cDNA for the $\alpha_{1A}$-adrenergic receptor. The gene for which is located on human chromosome 5. *J. Biol. Chem.* **266,** 6365–6369 (1991).
28. Fontanilla, D. *et al.* The hallucinogen *N,N*-dimethyltryptamine (DMT) is an endogenous sigma-1 receptor regulator. *Science* **323,** 934–937 (2009).
29. Su, T. P., Hayashi, T. & Vaupel, D. B. When the endogenous hallucinogenic trace amine *N,N*-dimethyltryptamine meets the sigma-1 receptor. *Sci. Signal.* **2,** pe12 (2009).
30. Roth, B. L., Lopez, E., Patel, S. & Kroeze, W. K. The multiplicity of serotonin receptors: uselessly diverse molecules or an embarrasment of riches? *Neuroscientist* **6,** 252–262 (2000).
31. Smith, R. L., Canton, H., Barrett, R. J. & Sanders-Bush, E. Agonist properties of *N,N*-dimethyltryptamine at serotonin 5-$HT_{2A}$ and 5-$HT_{2C}$ receptors. *Pharmacol. Biochem. Behav.* **61,** 323–330 (1998).
32. Kohen, R. *et al.* Cloning, characterization, and chromosomal localization of a human 5-$HT_6$ serotonin receptor. *J. Neurochem.* **66,** 47–56 (1996).
33. Pierce, P. A. & Peroutka, S. J. Hallucinogenic drug interactions with neurotransmitter receptor binding sites in human cortex. *Psychopharmacology (Berl.)* **97,** 118–122 (1989).
34. Abbas, A. I. *et al.* PSD-95 is essential for hallucinogen and atypical antipsychotic drug actions at serotonin receptors. *J. Neurosci.* **29,** 7124–7136 (2009).
35. Kurland, A. A., Mc, C. K. & Michaux, W. W. Clinical trial of haloanisone (R-2028) with hospitalized psychiatric patients. *J. New Drugs* **2,** 352–360 (1962).
36. Gankina, E. M. *et al.* Effect of some antihistamine preparations on binding of $^3$H-mepyramine and $^3$H-cimetidine to histamine receptors in rat brain. *Pharm. Chem. J* **26,** 373–376 (1992).
37. Gankina, E. M. *et al.* The effect of antihistaminic preparations on the binding of labelled mepyramine, ketanserin and quinuclidinyl benzilate in the rat brain [in Russian with English abstract]. *Eksp. Klin. Farmakol.* **56,** 22–24 (1993).
38. Heykants, J. *et al.* On the pharmacokinetics of domperidone in animals and man. IV. The pharmacokinetics of intravenous domperidone and its bioavailability in man following intramuscular, oral and rectal administration. *Eur. J. Drug Metab. Pharmacokinet.* **6,** 61–70 (1981).
39. FDA. Talk Paper: FDA Warns Against Women Using Unapproved Drug, Domperidone, to Increase Milk Production < http://www.fda.gov/Drugs/DrugSafety/InformationbyDrugClass/ucm173886.htm> (7 June 2004).
40. Stork, D. *et al.* State dependent dissociation of HERG channel inhibitors. *Br. J. Pharmacol.* **151,** 1368–1376 (2007).
41. Michelson, D. *et al.* Interruption of selective serotonin reuptake inhibitor treatment. Double-blind, placebo-controlled trial. *Br. J. Psychiatry* **176,** 363–368 (2000).
42. Berger, M., Gray, J. A. & Roth, B. L. The extended pharmacology of serotonin. *Annu. Rev. Med.* **60,** 355–366 (2009).
43. Waldinger, M. D., Hengeveld, M. W., Zwinderman, A. H. & Olivier, B. Effect of SSRI antidepressants on ejaculation: a double-blind, randomized, placebo-controlled study with fluoxetine, fluvoxamine, paroxetine, and sertraline. *J. Clin. Psychopharmacol.* **18,** 274–281 (1998).
44. Peters, J. U., Schnider, P., Mattei, P. & Kansy, M. Pharmacological promiscuity: dependence on compound properties and target specificity in a set of recent Roche compounds. *ChemMedChem* **4,** 680–686 (2009).
45. Scott, L. J. & Perry, C. M. Delavirdine: a review of its use in HIV infection. *Drugs* **60,** 1411–1444 (2000).
46. Dijkstra, D. *et al.* Human inflammatory dendritic epidermal cells express a functional histamine H4 receptor. *J. Invest. Dermatol.* **128,** 1696–1703 (2008).
47. Mehvar, R., Jamali, F., Watson, M. W. & Skelton, D. Pharmacokinetics of tetrabenazine and its major metabolite in man and rat. Bioavailability and dose dependency studies. *Drug Metab. Dispos.* **15,** 250–255 (1987).
48. Masanori, I., Tetsuya, T., Tomihiro, I., Taku, N. & Shigeyuki, T. β1-adrenergic selectivity of the new cardiotonic agent denopamine in its stimulating effects on adenylate cyclase. *Biochem. Pharmacol.* **36,** 1947–1954 (1987).
49. Jensen, N. H. *et al. N*-desalkylquetiapine, a potent norepinephrine reuptake inhibitor and partial 5-$HT_{1A}$ agonist, as a putative mediator of quetiapine's antidepressant activity. *Neuropsychopharmacology* **33,** 2303–2312 (2008).
50. Irwin, J. J. & Shoichet, B. K. ZINC–a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **45,** 177–182 (2005).


**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** Supported by grants from the National Institutes of Health (NIH) supporting chemoinformatics (to B.K.S. and J.J.I.) and NIH grants and contracts supporting drug discovery and receptor pharmacology (to B.L.R). M.J.K., J.H. and C.L. were supported by fellowships from the National Science Foundation, the 6th FP of the European Commission, and the Max Kade Foundation, respectively. B.L.R. was also supported by a Distinguished Investigator Award from the NARSAD and the Michael Hooker Chair. We thank T. Oprea of Sunset Molecular for WOMBAT, Elsevier MDL for the MDDR, Scitegic for PipelinePilot, J. Overington of the European Bioinformatics Institute (EMBL-EBI) for StARlite, Daylight Chemical Information Systems Inc. for the Daylight toolkit, and J. Gingrich for 5-$HT_{2A}$ knockout mice.

**Author Contributions** B.K.S., J.J.I. and M.J.K. developed the ideas for SEA. M.J.K. wrote the SEA algorithms, undertook the calculations, and identified the off-targets reported here, typically vetted with J.J.I. and B.K.S., unless otherwise noted below. M.J.K. wrote the naive Bayesian classifier algorithms with assistance from J.H. With assistance from B.K.S. and J.J.I., C.L. identified off-targets for Fabahistin, K.L.H.T. identified off-targets for Prozac and Paxil, and D.D.E. identified the off-target for Rescriptor. V.S. and B.L.R. designed empirical tests of the predictions, analysed and interpreted data, and performed experiments. V.S., T.B.T., R.W., R.C.M., A.A., N.H.J. and M.B.K. performed empirical testing of the predictions. V.S., S.J.H. and R.A.G. generated materials for the experiments. M.J.K., B.K.S. and B.L.R. wrote the manuscript with contributions and review from V.S. All authors discussed the results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Correspondence and requests for materials should be addressed to B.K.S. (shoichet@cgl.ucsf.edu) or B.L.R. (bryan_roth@med.unc.edu).



©2009 Macmillan Publishers Limited. All rights reserved

## METHODS

**Ligand sets.** We extracted ligand sets from databases that annotate molecules by therapeutic or biological category. For example, the 2006.1 MDDR contains 518 molecules annotated as $\alpha_1$ adrenergic receptor blockers, which we grouped into a single '$\alpha_1$ adrenergic blocker' set.

As ligand reference sources, we used the three databases shown in Supplementary Table 8. The first was a subset of the 2006.1 MDDR, prepared as previously described[21,23]. The second was the 2006.2 WOMBAT database[26], processed as earlier. We collapsed WOMBAT targets across species and organized them into inhibitory, activating and unspecified-binding classes. All ligands with affinities worse than 1 μM to their targets were removed. This left 1,133 classes built from 191,943 ligands with median and mean of 37 and 169 ligands per target class. The third database was StARlite, which we also processed as above. We extracted StARlite annotations at the two highest confidence levels (5 and 7), discarded those with affinities worse than 1 μM, and organized them into target classes. This yielded 1,158 classes built from 111,329 ligands with a median and mean of 43 and 186 ligands per target class.

For drugs and bioactive molecules, we used the 2004 MDL Comprehensive Medicinal Chemistry database (CMC) of 7,517 compounds. Drugs were processed identically to the ligands above. Filtering by vendor availability (as reported in MDL 2006.3 Available Chemical Directory (ACD), MDL 2006.1 Screening Compounds Directory, and ZINC[50]) yielded 3,665 unique purchasable drugs.

The 1,216 drugs used to link protein targets in the drug-target networks (Fig. 1) were downloaded from the 2008 EPA Distributed Structure-Searchable Toxicity (DSSTox) Database at http://www.epa.gov/NCCT/dsstox/, and prepared as earlier.

**Ligand activity predictions.** We compared each drug individually against each set of ligands. Molecules were represented by two topological descriptors: 2048-bit Daylight[51] and 1024-bit folded ECFP_4 fingerprints[23]. We used the SEA[21,23] using each descriptor separately and chose top-scoring hits (that is, those with small *E*-values) from each such 'screen' independently.

The initial SEA screens of 3,665 CMC drugs against 246 MDDR targets yielded 901,590 drug–target comparisons, and we subjected these to retrospective literature analysis and prospective empirical testing. However, we later extended SEA screens to WOMBAT and StARlite databases, comprising 4,152,445 and 4,244,070 drug–target comparisons, respectively. We have not mined these expanded SEA screens for retrospective validations; instead we used them only to conduct prospective tests. Supplementary Table 8 records the screen (that is, database) from which each prediction in Tables 1 and 2 is derived.

To compare SEA predictions against those of naive Bayesian classifiers (Supplementary Table 1), we implemented a Laplacian-corrected naive Bayesian classifier with Avidon weighting, as previously described[23].

**Drug–target and target–target networks.** The drug–target networks in Fig. 1 are bipartite; along any given path, nodes alternate between protein targets and the drugs that link them. Targets are from WOMBAT and drugs are from EPA DSSTox. Red edges denote SEA predictions with *E*-values $\leq 10^{-10}$. Predictions already reported in WOMBAT at $K_i \leq 1$ μM are shown as additional grey edges. All networks were generated in Cytoscape 2.6.1 (ref. 52).

Figure 3 is a bipartite graph linking drugs from Tables 1 and 2 with protein targets. Grey edges link drugs to known targets from manual literature and database search. Grey edges denote binding at $\leq 1$ μM, except when no $K_i$ value was available; in these three cases (Xenazine–VMAT2, Prantal–$M_3$ and Fabahistin–$H_1$), the link was included for completeness.

**WOMBAT out-group analysis.** We mapped 204 MDDR activity classes to matching WOMBAT targets in two phases. In the first, we mapped 87 MDDR activity classes using EC numbers from the Schuffenhauer ontology[25] to those present in WOMBAT. We second mapped a further 118 non-enzyme MDDR activity classes by supervised sub-phrase matching (Supplementary Tables 9 and 10). Although this mapping is not guaranteed to be exhaustive, it is correct to the best of our knowledge.

We then extracted all molecules marked 'drug' in WOMBAT (746 unique). Using SEA, we compared them against the mapped MDDR classes only, and discarded all trivial hits (that is, those in which the drug was already annotated in that MDDR class as a ligand). We asked how many of these were, in retrospect, substantiated by the existing WOMBAT annotations, at affinities $\leq 1$ μM.

**Sequence similarity comparison.** We associated each drug in Fig. 3 with the human FASTA sequences of its known and its new protein targets, using http://www.uniprot.org. We ran these sequences by PSI-BLAST (BLAST version 2.2.14, default parameters)[21–23] against a subset of MDDR targets that we prepared as previously described[21]. For each SEA prediction in Fig. 3, we reported the best direct PSI-BLAST match (along with its *E*-value and ranking) of the new protein target to any of that drug's previously known protein targets (Supplementary Table 6). Our goal was to address whether starting with the best choice from a drug's known protein targets, we could recapitulate each SEA prediction solely by sequence similarity.

**Experimental testing.** Radioligand binding and functional assays were performed as previously described[49,53]. Detailed experimental protocols are available on the NIMH PDSP website at http://pdsp.med.unc.edu/UNC-CH%20Protocol%20Book.pdf.

**Mice.** All experiments were approved by the Institutional Animal Care and Use Committee at the University of North Carolina, Chapel Hill. Mice were housed under standard conditions: 12 h light/dark cycle and food and water ad libitum.

**Head twitch.** Littermate pairs of 5-$HT_{2A}$ wild-type and knockout mice were pretreated for 2 h with 75 mg $kg^{-1}$ pargyline, intraperitoneally, prepared in sterile saline (0.9% NaCl) (Sigma-Aldrich). Mice were then injected with sterile saline or 1.0 mg $kg^{-1}$ DMT, intraperitoneally, prepared in sterile saline and moved to a new cage. Head twitch behaviour, which consists of a rapid, rotational flick of the head about the axis of the neck, was counted over 15 min. We have determined that trained observers count the same number of head twitches whether blinded or unblinded to genotype (data not shown). We confirmed that this was the case with three littermate pairs, and the rest of the studies were performed by one unblinded observer[34].


51. James, C., Weininger, D. & Delany, J. *Daylight Theory Manual* (Daylight Chemical Information Systems Inc., 1992–2005).
52. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
53. Roth, B. L. *et al.* Salvinorin A: a potent naturally occurring nonnitrogenous κ opioid selective agonist. *Proc. Natl Acad. Sci. USA* **99**, 11934–11939 (2002).


©2009 Macmillan Publishers Limited. All rights reserved