Computational Biology Methods in Protein Research: Identifying Chemokine-Receptor Pairs

Matteo Pellegrini
Protein Pathways, Los Angeles USA

1 Introduction

During the past few decades, computational methods in protein research focused on sequence based analyses. The growing number of protein sequences accumulated in Genbank allowed scientists to ask many questions about how the proteins were evolutionarily related. These methods all relied on sequence comparison algorithms: Needleman Wunsch, Smith-Waterman, BLAST, Hidden Markov Models to name a few **(1-4)**. In general, all these methods allowed scientists to group proteins into families that share sequence similarity.

This exercise is very useful to identify the biochemical functions of proteins. This is due to the fact, that as sequences diverge within a family, protein functions can often remain intact. Therefore, the central paradigm has been that if one can identify the function of one protein in a family, one can understand the function of many of the other proteins within the family as well.

Although this research has by means been exhausted, in the current era where the entire genomes of organisms can be sequenced, the focus of computational biology is gradually shifting. A great deal of research in the past few years has focused on what many consider the next grand challenge: now that we know the protein components of cells, and understand the biochemical functions of many of them, can we begin to understand how these functions are coupled **(5)**. In other words, the function of each protein depends on the concentrations of other molecules within the cell, through a vast network of molecular interactions. Can we begin to elucidate the network of interactions in order to predict how modulating the concentrations of one molecular species affect other proteins?

Although the general question is still too difficult to tackle, here we present two methodologies that begin to ask a more detailed but related question: can we predict the pairings of chemokines with their receptors (*see* **Note 1**). Chemokine-receptor pairs are of particular interest since they have been extensively studied and a great deal is known about them **(6,7)**. In general, chemokines control diverse biological processes by activating G-protein-coupled receptors on the cell surface. These processes include angiogenesis, hematopoesis and organogenesis among many others. Understanding the

interactions between chemokines and their receptors can shed a great deal of light on the molecular basis of these events.

Here we present two separate methodologies to predict the pairings between chemokines and their receptors. The first method relies on the premise that in certain cases chemokines and their receptors will be transcriptionally co-regulated **(8)**. This may occur in cases where the recptor-ligand pairs are part of autocrine signaling loops. The strategy will be to identify which chemokine-receptor pairs are potentially forming autocrine loops by finding pairs that are co-regulated using expression microarray data.

The second method attempts to reconstruct chemokine-receptor pairs using phylogeny **(9)**. The central assumption is that the pairs must co-evolve in order to maintain their binding specificity. The method consists of first constructing extensive phylogenetic trees to capture the evolutionary distance between members of the chmokine and receptor families independently. By comparing the trees it is possible to demonstrate that they share many similarities: ligands that cluster in one tree bind receptors that are clustered in the other. It is therefore possible in principle to identify receptors of orphan ligands using this methodology.

2 Materials

1. List of known chemokine receptor pairs
2. Microarray expression data sets
3. PSI-BLAST
4. ClustalW
5. PHYLIP

3 Methods

3.1 List of Chemokine Receptor Pairs

There are many public databases of protein interactions that contain information on chemokine receptor pairs, such as DIP and BIND **(10,11)**. One of the most comprehensive is the Database of Ligand Receptor Pairs (DLRP) **(8)**. The contents of this database may be obtained at http://dip.doe-mbi.ucla.edu/dip/DLRP.cgi. We list the set of pairs in Table I.

3.2 Identifying Co-expressed Chemokine-Receptor Pairs

Relationships between chemokines and their receptors may be uncovered using expression data **(8)**. Below we describe how one may obtain and analyze expression data

to determine which pairs of ligands and receptors are most likely transcriptionally coregulated. When a receptor-ligand pair is co-expressed in data from one cell tyope this indicates that the ligand potentially binds the receptor to form an autocrine loop.

3.2.1 Microarray Expression Datasets

There is a growing wealth of public data on microarray expression experiments. Unfortunately, unlike sequence or structure databases, there is not yet a single repository for microarray data. One of the most extensive sources is the Stanford Microarray database (http://genome-www5.stanford.edu/MicroArray/SMD/). This database allows users to download expression microarray datasets for human and other organisms. As of the writing of this article there were 18 human datsests available, ranging over a wide variety of tissue types, including cancerous tissues.

3.2.2 Correlations in Expression between Chemokines and their Receptors

The first step to establish that a chemokine and its receptor are expressed in a correlated fashion is to extract the values of the expression intensities for each across a dataset. Expression values are usually measured as absolute intensities or as log ratios. In either case the correlation between the measured values of two genes across multiple experiments may be computed using the Pearson correlation coefficient:

$$C = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2}\sqrt{\sum_i (y_i - \bar{y})}},$$

where C is the correlation coefficient, x_i are the values of gene x in experiment i, $\bar{x}$ is the average value of gene x across all the experiments and y_i are the values measured for gene y (*see* **Note 2**).

3.2.3 Estimating the Statistical Significance of the Correlation

The correlations computed in the previous section are useful for rank ordering chemokine-receptor pairs according to the likelihood that they are co-regulated. However, it is important to assess the probability that the pair is actually co-expressed. One method to accomplish this is to ask what the probability of observing the above correlation is from the distribution of correlations between all the pairings of the genes on the microarray. This distribution usually takes the form of a Gaussian distribution. An example of this distribution is seen in figure 1. Once the distribution is computed, one can calculate the fraction of pairs that have a greater pairwise correlation than the one observed for the specific chemokine-receptor pair. This probability is a measure of the likelihood of observing the correlation by chance, and hence allows us to gauge whether the observation is meaningful.

3.2.4 Coexpressed Chemokine-Receptor Pairs

In Table II we list the most significantly coexpressed chemokine-receptor pairs from Graeber et.al. **(8)**. The expression data came form a leukemia dataset generated at the Whitehead/MIT Center for Genome Research (http://waldo.wi.mit.edu/MPR/) **(14)**. Among this list are known chemokine-receptor pairs that might be involved in autocrine signaling loops.

3.3 Identifying Co-Evolving Chemokine-Receptor Pairs

Relationships between chemokines and their receptors may also be determined using phylogenetic techniques **(9)**. The underlying assumption is that if a ligand and its receptor are to maintain a high affinity for each other as they evolve subject to point mutations, the preserved mutations must be correlated between the two. That is, it is likely that a mutation in the receptor binding-site is conserved only if a compensating mutation is made on the ligand's epitope. Therefore if we capture the degree of evolution that occurs among chemokine receptors in a phylogenetic tree, it is likely to mirror the tree produced from the chemokines that bind these receptors. The mate of an orphan receptor or chemokine may be inferred from a comparison of the two trees.

3.3.1 Constructing Phylogenetic Trees

To construct phylogenetic trees we have used the program CLUSTALW **(12)**. The program is available on the web at http://www.ebi.ac.uk/clustalw/. The program takes as an input a list of sequences. We have constructed such a list for the receptors and the ligands listed in Table I. The sequences must be in the FASTA format as seen in Figure 2. The list of sequences may then be submitted to the program though a web form or, if the program is downloaded, though a command line prompt. CLUSTALW allows the user to alter many parameters that affect the topology of the phylogenetic tree, but a detailed discussion of these parameters is beyond the scope of this work. To generate the trees in Figures 3 and 4 we used the default parameters (*see* **Note 3**). With these parameters the program generates a phylogentic tree in the standard "New Hampshire" format.

3.3.2 Displaying the Phylogenetic Tree

The phylogenetic tree that is produced by CLUSTALW may be displayed using the program PHYLIP **(13)**. PHYLIP is a suite of programs for phylogentic and sequence analysis. We have used the program DrawTree from this suite to generate Figures 3 and 4. This program takes as its input a phylogenetic tree in the standard "New Hampshire" format, generated for example by CLUSTALW, and generates an unrooted tree. The program allows the user to alter parameters that modify the appearance of the unrooted tree, such as the font and direction of the labels.

3.3.3 Comparing the Chemokine Receptor and Chemokine Phylogentic Trees

The phylogenetic tree for the chemokine receptors is illustrated in figure 3, while that for the chemokines is in figure 4. In the chemokine tree, each leaf is labeled not only by the

chemokine name, but also by the receptors that the chemokine binds. As is evident form the figure, there are more chemokines than receptors and therefore the chemokine tree is inherently more complex. At first site there is not much evidence that the two trees are in any way correlated, however, upon closer inspection one notices some similarities. For instance the chemokines that bind CCR1, CCR3 and CCR8 form a tight cluster in the chemokine tree. Similarly, CCR1, CCR3 and CCR8 also form a tight cluster in the receptor tree. Similarly, all the ligands for ILR8 and CXCR4 form a tight cluster in the chemokine tree, while the receptors ILR8 and CXCR4 are close by on the receptor tree. Other similarities can be found between the trees of the receptors and the ligands even though there are many relationships that do not hold up. For instance, the chemokines that bind GPR9 and CCR6 are nearby on the chemokine tree but their receptors are not close by on the receptor tree.

Given the level correspondence between the ligand and receptor phylogenetic trees, it should be possible in some cases to identify the receptor of an orphan ligand and vice versa. To do this one would identify the position of the orphan ligand on the chemokine tree and look to see which receptors are bound by ligands nearby on this tree. Although this would certainly not identify with high confidence the receptor of an orphan ligand, it could certainly provide valuable clues that could be verified experimentally.

3.3.4 Determining the Statistical Significance of the Correspondence between two Trees

Is the observation that the chemokine tree and the chemokine receptor tree show similarities statistically significant? If we were to generate many phylogenetic trees at random would we find that they were more or less similar than the actual trees of the chemokines and chemokine receptors?

One can answer these questions by comparing the distance matrices associated with each phylogenetic tree. The distance matrices may be output by the command line version of the CLUSTALW program with the appropriate settings. The distance matrix is what the program actually uses to calculate the phylogentic tree, and is an intermediate representation of the data between the multiple alignment of the input sequences and the final tree. The distance matrix contains all the pairwise distances between the input sequences to CLUSTALW; in our case the sequences of the chemokines or the receptors. Distances are a measure the evolutionary time of divergence between two sequences.

To compare two distances matrices one must look at the correlation between the distances of pairs of receptors and pairs of the associated chemokines. One may then calculate the correlations between these pairs of distances using the formula of section 3.2.2. This correlation is then compared to the correlations that one observes when pairs of receptors are randomly associated with pairs of ligands. One must calculate the distribution of correlations for the random couplings of receptors and ligands, and determine how many standard deviations from the mean the correlation for the true couplings is. The greater the number of standard deviations from the mean, the more significant is the similarity between the receptor and ligand tree. For instance, the correlation between the chemokine and chemokine receptor tree calculated by Goh et. al. is 0.44, while the standard deviation from the mean of randomly paired trees is 11.2. The

probability of observing this correlation by chance from the random distribution is 2.3 x 10^{-29}, therefore the trees are far more similar than one would expect by chance.

4 Notes

1. Interactions between proteins, as in the case of ligand-receptor associations, occur over a broad range of affinities. The strongest protein-protein associations have equilibrium constants in the nanomolar range, while weaker, yet still biologically important ones, have constants in the micromolar range. This broad range of affinities complicates any attempts to predict protein interactions from expression or protein sequence data. One should consider only interactions that have a measurable biological outcome, although it is difficult to determine how to do this in many cases.
2. The correlation coefficients that one calculates between genes based on their expression levels measured across multiple experiments are critically dependent on the nature of the experiments. Therefore one should not always expect to observe the same correlations between genes measured for instance in a set of cancerous cells versus a set of tissues from normal cells. One should not always expect the observation of a significant correlation to be tissue independent.
3. The construction of phylogenetic trees is very sensitive to the particular methodology used. There is no optimal procedure for constructing multiple alignments between all the input sequences as this critical first step depends on the substitution matrix, the alignments parameters, the gap parameters and the multiple sequence algorithm used. There are also a variety of methodologies to compute distance matrices from a multiple sequence alignment, each one yielding slightly different results. Finally a tree is an approximate and non-optimal representation of a distance matrix, and therefore each tree-building algorithm yields somewhat different results. Nonetheless, the information captured in a phylogenetic tree is an extremely useful and often robust representation of the evolutionary relationships between protein sequences.

5 References

1. Needleman, S.B., Wunsch, C.D. (1970) A general method applicable to the search for similarities in the amino acid sequences of two proteins. J Mol Biol **48**, 443-453.

2. Smith, T.F., Waterman, M.S. (1981) Identification of common molecular subsequences. J Mol Biol **147**, 195-197.

3. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z, Miller, W., Lipman, D.J. (1999) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acid Res **25**, 3389-3402.

4. Krogh, A., Brown, M., Mian, I.S., Sjolander, K., Haussler, D. (1994) Hidden Markov models in computational biology. Applications to protein modeling. J Mol Biol **235**, 1501-1531.

5. Pellegrini, M. (2001) Computational methods for protein function analysis. Current Opinion in Chemical Biology, **5**, 46–50

6. Baggiolini, M., Dewald, B., Moser, B. (1997). Human chemokines: an update. Annu. Rev. Immunol. **15**, 675-705.

7. Murphy, P. M., Baggiolini, M., Charo, I. F., Hebert, C. A., Horuk, R., Matsushima, K., Miller, L. H., Oppenheim, J. J., and Power, C. A. (2000). International union of pharmacology. XXII. Nomenclature for chemokine receptors, Pharmacological Reviews **52**, 145-76.

8. Graeber, T.G., Eisenberg, D. (2001) Bioinformatic identification of potential autocrine signaling loops in cancers from gene expression profiles nature genetics. Nature Genetics **29**, 295-300.

9. Goh, C.S., Bogan, A.A., Joachimiak, M., Walther, D. and Cohen, F.E. (2000) Co-evolution of Proteins with their Interaction Partners. J. Mol. Biol. **299**, 283-293.

10. Xenarios, I., Salwinski, L., Duan, X.J., Higney, P., Kim, S.M., Eisenberg, D. (2002) DIP, the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions. Nucleic Acids Res. **30**, 303–305.

11. Bader, G.D., Hogue, C.W. (2000) BIND--a data specification for storing and describing biomolecular interactions, molecular complexes and pathways. Bioinformatics. **16**(5), 465-77.

12. Thompson, J. D., Higgins, D. C. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acids Res. **22**, 4673-4680.

13. Felsenstein, J. (1993). PHYLIP (Phylogeny Inference Package), 3.5c edit., Department of Genetics, University of Washington, Seattle.

14. Golub, T.R., Slonim, D.K., Tamayo, P, Huard, C, Gaasenbeek, M, Mesirov, J.P., Coller, H, Loh, M.L., Downing, J.R., Caligiuri, M.A., Bloomfield, C.D., Lander, E.S. (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **28**6, 531–537.

Table I : Chemokine-Receptor Pairs*

Ligand	Receptor
SCYA18 small inducible cytokine subfamily A (Cys-Cys), member 18, pulmonary and activation-regulated	
SCYA11 small inducible cytokine subfamily A (Cys-Cys), member 11 (eotaxin)	CCR3 chemokine (C-C motif) receptor 3
	GPR9 G protein-coupled receptor 9 (weak)
	CCR5 chemokine (C-C motif) receptor 5 (weak)
	CCBP2 chemokine binding protein 2 (weak)
SCYA24 small inducible cytokine subfamily A (Cys-Cys), member 24	CCR3 chemokine (C-C motif) receptor 3
SCYA23 small inducible cytokine subfamily A (Cys-Cys), member 23	CCR1 chemokine (C-C motif) receptor 1
SCYA14 small inducible cytokine subfamily A (Cys-Cys), member 14	CCR1 chemokine (C-C motif) receptor 1
	CCBP2 chemokine binding protein 2
SCYA16 small inducible cytokine subfamily A (Cys-Cys), member 16	
SCYA26 small inducible cytokine subfamily A (Cys-Cys), member 2	CCR3 chemokine (C-C motif) receptor 3
SCYA3 small inducible cytokine A3 (homologous to mouse Mip-1a)	CCR1 chemokine (C-C motif) receptor 1
	CCR5 chemokine (C-C motif) receptor 5
SCYA4 small inducible cytokine A4 (homologous to mouse Mip-1b)	CCR5 chemokine (C-C motif) receptor 5
	CCR8 chemokine (C-C motif) receptor 8 (weak)
	CCR1 chemokine (C-C motif) receptor 1 (weak)
	CCBP2 chemokine binding protein 2
SCYA15 small inducible cytokine subfamily A (Cys-Cys), member 15	CCR1 chemokine (C-C motif) receptor 1
	CCR3 chemokine (C-C motif) receptor 3
SCYA1 small inducible cytokine A1 (I-309, homologous to mouse Tca-3)	CCR8 chemokine (C-C motif) receptor 8
SCYA21 small inducible cytokine subfamily A (Cys-Cys), member 21	CCR7 chemokine (C-C motif) receptor 7
	GPR9 G protein-coupled receptor 9 (weak)
SCYA5 small inducible cytokine A5 (RANTES)	CCR1 chemokine (C-C motif) receptor 1
	CCR3 chemokine (C-C motif) receptor 3
	CCR4 chemokine (C-C motif) receptor 4
	CCR5 chemokine (C-C motif) receptor 5
	FY Duffy blood group (non-signalling)
	CCBP2 chemokine binding protein 2
SCYA2 small inducible cytokine A2 (monocyte chemotactic protein 1, homologous to mouse Sig-je)	CCR2 chemokine (C-C motif) receptor 2
	FY Duffy blood group (non-signalling)
	CCR1 chemokine (C-C motif) receptor 1 (weak)
	CCBP2 chemokine binding protein 2
SCYA8 small inducible cytokine subfamily A (Cys-Cys), member 8 (monocyte chemotactic protein 2)	CCR2 chemokine (C-C motif) receptor 2
	CCR3 chemokine (C-C motif) receptor 3
	CCR5 chemokine (C-C motif) receptor 5
	CCR1 chemokine (C-C motif) receptor 1
	CCBP2 chemokine binding protein 2
SCYA27 small inducible cytokine subfamily A (Cys-Cys), member 27	GPR2 G protein-coupled receptor 2
SCYA7 small inducible cytokine A7 (monocyte chemotactic protein 3)	CCR1 chemokine (C-C motif) receptor 1
	CCR2 chemokine (C-C motif) receptor 2
	CCR3 chemokine (C-C motif) receptor 3
	CCR5 chemokine (C-C motif) receptor 5 (antagonist)

	CCBP2 chemokine binding protein 2 (weak)
SCYA13 small inducible cytokine subfamily A (Cys-Cys), member 13	CCR2 chemokine (C-C motif) receptor
	CCR3 chemokine (C-C motif) receptor 3
	CCR1 chemokine (C-C motif) receptor 1
	CCR5 chemokine (C-C motif) receptor 5 (weak)
	CCBP2 chemokine binding protein 2
SCYA17 small inducible cytokine subfamily A (Cys-Cys), member 17	CCR4 chemokine (C-C motif) receptor 4
	CCR8 chemokine (C-C motif) receptor 8 (weak)
SCYA20 small inducible cytokine subfamily A (Cys-Cys), member 20	CCR6 chemokine (C-C motif) receptor 6
SCYA19 small inducible cytokine subfamily A (Cys-Cys), member 19	CCR7 chemokine (C-C motif) receptor 7
SCYA25 small inducible cytokine subfamily A (Cys-Cys), member 25	CCR9 chemokine (C-C motif) receptor 9
SCYA22 small inducible cytokine subfamily A (Cys-Cys), member 22	CCR4 chemokine (C-C motif) receptor 4
IL8 interleukin 8	IL8RA interleukin 8 receptor, alpha
	IL8RB interleukin 8 receptor, beta
	FY Duffy blood group (non-signalling)
SCYB6 small inducible cytokine subfamily B (Cys-X-Cys), member 6 (granulocyte chemotactic protein 2)	IL8RA interleukin 8 receptor, alpha
	IL8RB interleukin 8 receptor, beta
MIG monokine induced by gamma interferon	GPR9 G protein-coupled receptor 9
SDF1 stromal cell-derived factor 1	CXCR4 chemokine (C-X-C motif), receptor 4 (fusin)
SCYB11 small inducible cytokine subfamily B (Cys-X-Cys), member 11	GPR9 G protein-coupled receptor 9
PF4 platelet factor 4	
SCYB5 small inducible cytokine subfamily B (Cys-X-Cys), member 5 (epithelial-derived neutrophil-activating peptide 78)	IL8RB interleukin 8 receptor, beta
	IL8RA interleukin 8 receptor, alpha (weak)
SCYB10 small inducible cytokine subfamily B (Cys-X-Cys), member 10	GPR9 G protein-coupled receptor 9
PPBP pro-platelet basic protein (includes platelet basic protein, beta-thromboglobulin, connective tissue-activating peptide III, neutrophil-activating peptide-2)	IL8RB interleukin 8 receptor, beta
	FY Duffy blood group (non-signalling)
GRO1 GRO1 oncogene (melanoma growth stimulating activity, alpha)	IL8RB interleukin 8 receptor, beta
	FY Duffy blood group (non-signalling)
GRO2 oncogene	IL8RB interleukin 8 receptor, beta
GRO3 oncogene	IL8RB interleukin 8 receptor, beta
SCYB13 small inducible cytokine B subfamily (Cys-X-Cys motif), member 13 (B-cell chemoattractant)	BLR1 Burkitt lymphoma receptor 1, GTP-binding protein
SCYB14 small inducible cytokine subfamily B (Cys-X-Cys), member 14 (BRAK)	
SCYC1 small inducible cytokine subfamily C, member 1 (lymphotactin)	CCXCR1 chemokine (C motif) XC receptor 1
SCYC2 small inducible cytokine subfamily C, member 2	CCXCR1 chemokine (C motif) XC receptor 1
SCYD1 small inducible cytokine subfamily D (Cys-X3-Cys), member 1 (fractalkine, neurotactin)	CX3CR1 chemokine (C-X3-C) receptor 1

*This list is obtained from DLRP **(8)**

Table II: Co-expressed Chemokine Receptor Pairs

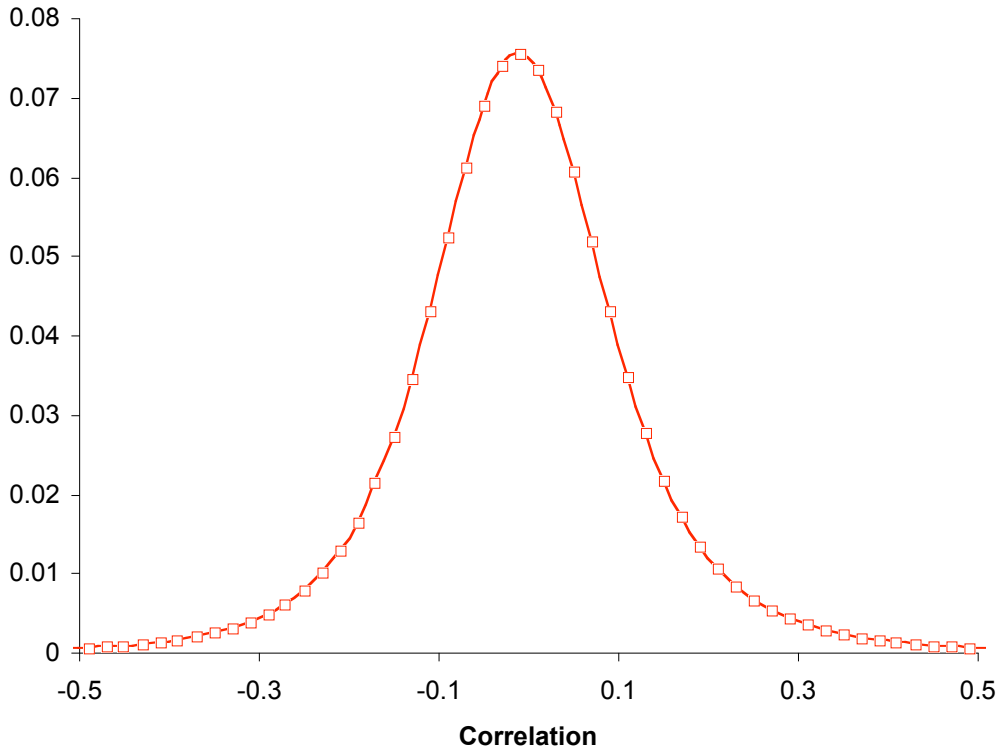
Chemokine	Receptor	Correlation	P
CCL2	FY	0.7	0.0014
CCL4	CCR5	0.61	0.0086
FGF4	FGFR2	0.54	0.022
EFNB1	EPHB4	0.51	0.032
IL10	IL10RA	0.48	0.043
CCL8	CCR5	0.47	0.048
JAG1	NOTCH4	0.34	0.15

Figure 1: Typical distribution of correlations between all pairs of genes within a microarray expression dataset. In this case the correlations are computed between all the pairs of yeast genes in about 300 microarray experiments obtained from the Stanford Microarray database.

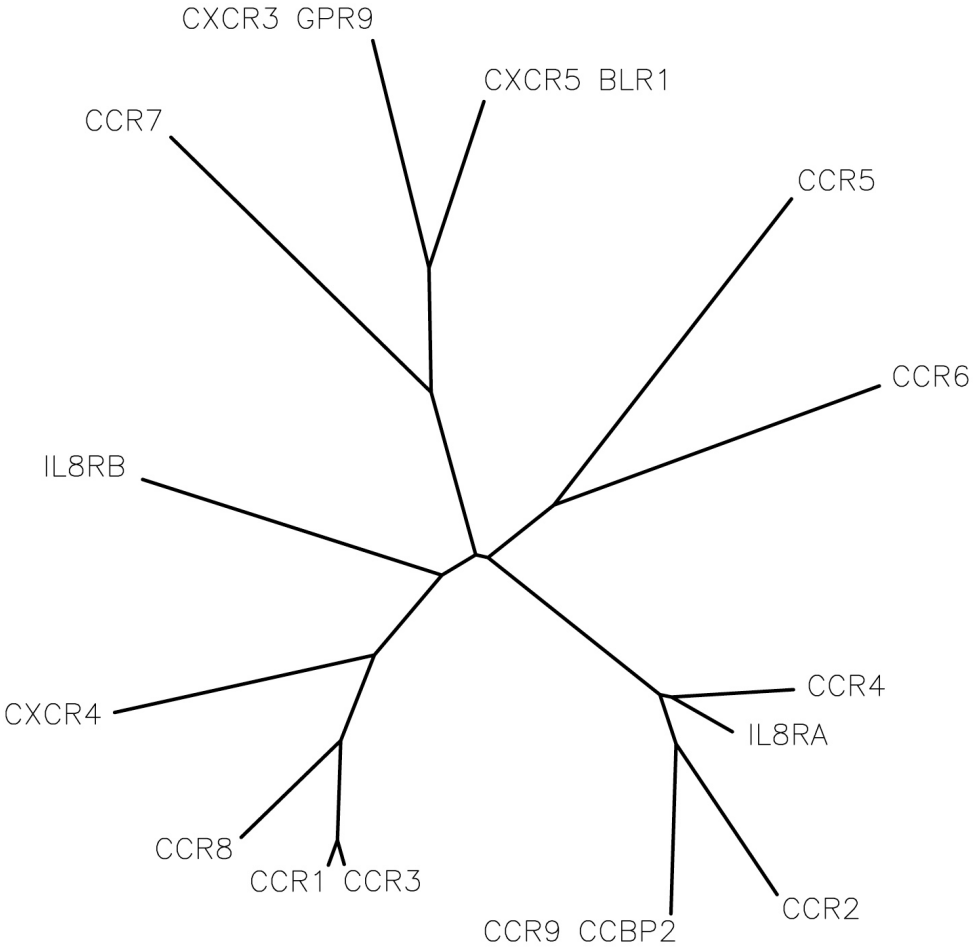
Figure 2: Chemokine receptor sequences in the FASTA format used by the program CLUSTALW to construct phylogentic trees.

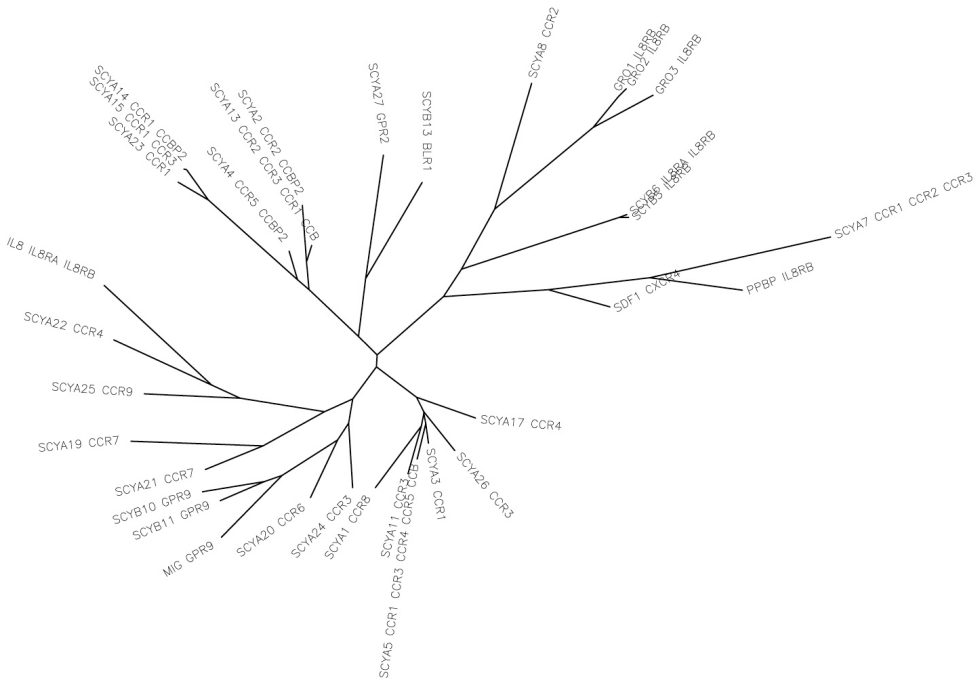
Figure 3: Phylogenetic tree of chemokine receptors. The tree was computed using the sequences of the receptors listed in Table I, using the program CLUSTALW and displayed using the program DrawTree from the PHYLIP package.

Figure 4: Phylogenetic tree of chemokines. The tree was computed using the sequences of the chemokines listed in Table I, using the program CLUSTALW and displayed using the program DrawTree from the PHYLIP package. Next to each chemokine, we list its receptors.


0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
-0.5
-0.3
-0.1
0.1
0.3
0.5
Correlation

>CCR3
MTTSLDTVETFGTTSYYDDVGLLCEKADTRALMAQFVPPLYSLVFTVGLLGNVVVVMILIKYRRLRIMTN
IYLLNLAISDLLFLVTLPFWIHYVRGHNWVFGHGMCKLLSGFYHTGLYSEIFFIILLTIDRYLAIVHAVF
ALRARTVTFGVITSIVTWGLAVLAALPEFIFYETEELFEETLCSALYPEDTVYSWRHFHTLRMTIFCLVL
PLLVMAICYTGIIKTLLRCPSKKKYKAIRLIFVIMAVFFIFWTPYNVAILLSSYQSILFGNDCERSKHLD
LVMLVTEVIAYSHCCMNPVIYAFVGERFRKYLRHFFHRHLLMHLGRYIPFLPSEKLERTSSVSPSTAEPE
LSIVF
>CCR5
MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKRLKSMTDIYLL
NLAISDLFFLLTVPFWAHYAAAQWDFGNTMCQLLTGLYFIGFFSGIFFIILLTIDRYLAVVHAVFALKAR
TVTFGVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTCSSHFPYSQYQFWKNFQTLKIVILGLVLPLLVM
VICYSGILKTLLRCRNEKKRHRAVRLIFTIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQ
VTETLGMTHCCINPIIYAFVGEKFRNYLLVFFQKHIAKRFCKCCSIFQQEAPERASSVYTRSTGEQEISV
GL


CXCR3 GPR9
CXCR5 BLR1
CCR7
CCR5
CCR6
IL8RB
CXCR4
CCR4
IL8RA
CCR8
CCR1 CCR3
CCR2
CCR9 CCBP2


SCYA14 CCR1 CCBP2
SCYA15 CCR1 CCR3
SCYA23 CCR1
SCYA2 CCR2 CCBP2
SCYA13 CCR2 CCR3 CCR1 CCB
SCYA4 CCR5 CCBP2
SCYA27 GPR2
SCYB13 BLR1
SCYA8 CCR2
GRO1 IL8RB
GRO2 IL8RB
GRO3 IL8RB
SCYA7 CCR1 CCR2 CCR3
PPBP IL8RB
SDF1 CXCR4
IL8 IL8RA IL8RB
SCYA22 CCR4
SCYA25 CCR9
SCYA19 CCR7
SCYA21 CCR7
SCYB10 GPR9
SCYB11 GPR9
MIG GPR9
SCYA20 CCR6
SCYA24 CCR3
SCYA1 CCR8
SCYA11 CCR3
SCYA5 CCR1 CCR3 CCR4 CCR5 CCB
SCYA3 CCR1
SCYA26 CCR3
SCYA17 CCR4