

A Two-Phase ANN Method for Genome-Wide Detection of Hormone Response Elements

Maria Stepanova¹, Feng Lin², and Valerie C.-L. Lin³

¹ Bioinformatics Research Centre

² School of Computer Engineering

³ School of Biological Sciences

Nanyang Technological University, Singapore 637551

{mari0004,asflin,cllin}@ntu.edu.sg

Abstract. Steroid hormone receptors compose a subgroup of regulatory proteins which tend to recognize partially symmetric response elements on DNA. Identification of the members of a gene regulatory machine conducted by steroid hormones could provide better understanding of nature and development of diseases. We present an approach based on a succession of neural networks, which can be used for highly specific detection of binding signals. It exploits the capability of a feed-forward neural network to model datasets with high confidence, while a recurrent network grants putative response elements with biologically meaningful structures. We have used a novel method to train such a two-phase artificial neural network with a set of experimentally validated response elements for steroid hormone receptors. We have demonstrated that sequence-based prediction followed by structure-based classification of putative binding sites allows to eliminate large amount of false positives. An implementation of the neural network with Field-Programmable Gate Array is also briefly described.

1 Introduction

The super-family of steroid and thyroid hormone receptor proteins includes receptors for steroid hormones, thyroid hormones, vitamin D and vitamin A (retinoic acid) [1]. Steroid hormones are involved in vital physiological processes, ranging from establishment and maintenance of pregnancy [2] to regulation in the genesis, progression and treatment of different cancers [3]. Activated by their hormone molecules, steroid hormone receptors usually bind to their target DNA – hormone response elements (HREs) – as partially symmetric homodimers [4]. HRE sequence, however, allows for a certain amount of flexibility in its nucleotide composition. Thus, recognition of these binding sites on DNA is one of the significant research topics in computational biology.

A comprehensive review on recognition of different strategies for transcription factor binding sites (TFBS) is given by Wasserman and Sandelin [5]. The most popular public resource is the P-Match tool based on a database of Position Weight Matrices (PWMs) for eukaryotic transcription factors, TRANSFAC [6];

and another open access is the database of TFBS profiles, JASPAR [7]. Unfortunately, due to high diversity of the eukaryotic transcription factor families, PWM-based representation of general TFBS patterns has been proved not very effective [8,9]; the poor quality of datasets for particular TFs of interest in these databases is another concern. For certain vertebrate transcription factors, however, weight matrix approach used as a part of a complex prediction system may perform well, as demonstrated by Bajic et al. for estrogen response element [10]. Research has also been done for more accurate recognition of HREs or HRE-like patterns. The best result to date is implemented Hidden Markov Model (HMM) for recognition of nuclear receptor binding sites on DNA [11]. However, the reported results suffer from a dataset which included all nuclear receptors together. Another promising idea, which involves Gibbs sampling model for partially symmetric structures, was reported by Favorov et al. [12], but only preliminary results for bacterial motifs have been reported.

In this paper, we present an artificial neural network (ANN) based approach for prediction of hormone receptor binding sites on DNA. The main idea of this approach is to predict a HRE-like DNA sequence first, and then estimate its probability of being a dimeric hormone response element based on its half-site structure. For this purpose, a two-phase neural network has been developed. A feed-forward neural network returns a list of potential HREs, followed by a recurrent neural classifier which predicts a possible dimeric structure for each of them. We found this approach very promising for HRE prediction in practice, with high sensitivity and reliable specificity. Furthermore, we analyzed the efficiency and cost of implementation of such a neural network. The expensive training process is eased via use of hardware acceleration, thanks to the outstanding performance of Field-Programmable Gate Arrays (FPGA).

2 Methods

The entire system for HRE recognition has been implemented in a form of two-phase sequence analysis. The first phase is the sequence-based prediction of putative hormone response elements. It is performed by a feed-forward neural network, and used for selection of HRE-like sequences on the entire DNA sequence of interest. The output of this module is a list of sequences with posterior probability of being functional hormone response element higher than a predefined threshold. This network is trained with use of a set of experimentally verified HREs by an adapted back-propagation method. The second phase is the structure-based validation. It is performed by a recurrent neural network, and corresponds to an attempt to classify a putative HRE (predicted at the previous stage) into one of groups each representing a dimeric response element structure – direct, inverted or palindromic repeat. The rationale behind that comes from experimental observations of dimeric protein-DNA interactions: a sequence is unlikely involved in homodimeric DNA-protein binding even if it is marked as a HRE-like in case if it cannot be reliably assigned with any known HRE structure.

2.1 Data Preparation

One can easily achieve very high sensitivity and specificity in identification of HRE with just a few sequences, but that result would be unreliable. A sufficient and accurate representative dataset is essential for the model to be trustable. In our study, the HRE data was collected from more than 200 literature sources and our in-house wet-lab experiments. Such a collection of HREs has no analogs in the current public and commercial databases of TFBS profiles.

In the dataset, while a few of the regulatory elements are derived from genes of fish and birds, most of the sites are mammalian and 90% of all sites are from human or rodent genomic DNA. The collection contains response elements for androgen (218 AREs), progesterone (66 PREs) and glucocorticoid (377 GREs) hormone receptors, which have been reported to share the same DNA sequence [13]. This observation has also been confirmed with use of our dataset [9]. Thus, in our current work, we use the joint unit consisting of all the three hormone receptors of interest unless stated otherwise.

2.2 Modeling of HREs with a Feed-Forward Neural Network

For sequence-based modeling of hormone response elements on DNA, a feed-forward neural network (FFNN) has been developed. It obtains the preprocessed (with use of one-hot notations for nucleotides) DNA sequence from an encrypting module, and returns the two posterior probabilities, each corresponding to either HRE-like or non-HRE sequence. Final output of the network – whether the sequence is a HRE – is subject to adjustment by a threshold. Training of the network is performed with use of experimentally validated HRE sequences for positive patterns, and neutral DNA sequences for negative ones.

From the output of the preprocessing module, the 15bp-long DNA sequence is converted to a 60-vector. The neural network theory [14] suggests that for a reliable learning result, the number of degrees of freedom, i.e. weights to be fitted, must be at most half of the number of constrains (the inputs accompanied by desired outputs), in order to avoid over-fitting. Therefore, in the case of one hidden layer and a dataset of about 700 positive (experimentally validated) and 7000 negative (extracted from neutral DNA) HREs, we should limit the number of hidden layer neurons to about 50. In the case of two hidden layers, the maximum number of neurons on each layer is approximately 40. The amount of negative patterns can even be increased to more than ten-fold over positive ones (as we have much more neutral DNA in comparison of functional DNA), but it may increase the risk of bias of the ANN model towards the most abundant pattern.

The bipolar sigmoid function is used for activation of neuron synaptic inputs. A series of cross-validation tests allowed us to fix its coefficients (individual for each neuron).

For training of the feed-forward neural model, a back-propagation learning algorithm is implemented. The weight adjustment for each neuron is represented by the following equation:

$$w^{t+1} = w^t + \alpha^t \times \delta \times x \quad (1)$$

where w^t is a vector of weights for a particular neuron at the t^{th} step of learning, α^t is the learning parameter at the t^{th} step ($0 < \alpha^t < 1, \forall t > 0$), and the delta value for each neuron of the ANN with one hidden layer and one output layer is calculated as follows:

$$\delta^{output} = \left. \frac{\partial f(u^t)}{\partial u^t} \right|_{u^t=w^t x} \times (d^t - o^t) \quad (2)$$

$$\delta^{back-propagated} = \left. \frac{\partial f(u^t)}{\partial u^t} \right|_{u^t=w^t x} \times \sum_{k=1}^K w_{h \rightarrow k} \delta_k^{output} \quad (3)$$

where d^t and o^t represent the desired and current outputs of the neuron respectively; x is the input to the layer being considered (either hidden or output), $u^t = w^t x$ is the synaptic input to the neuron, and $f(u^t)$ is the activation function of the neuron. For the back-propagated delta value, K is a number of neurons on the output layer, $w_{h \rightarrow k}$ is the weight coefficient of the connection between h^{th} neuron of hidden layer and k^{th} neuron of the output layer, δ_k^{output} is a delta value for the k^{th} neuron of the output layer calculated as shown by formula (2). The back-propagation is terminated when

- i.* Error tolerance for the accuracy of 99.99% is satisfied, or
- ii.* The desired number of back-propagation cycles or the error plateau is reached.

Learning rate α is adjusted depending on whether current level of training error is decreasing (then α is increased by 10%), or it has jumped over a minimum and it is better to move back slightly slower (then α is decreased by 20%).

2.3 Classification of Respective HREs with a Recurrent Neural Network

For structure-based recognition of various HREs, we designed a recurrent neural network (RNN). It employs a quasi-heuristic approach for classification of partially overlapped datasets, and requires domain knowledge for its design.

With a few exceptions, steroid hormone receptors bind to their response elements in a dimeric form. Thus, the structure of a response element can be treated as a repeat. However, the half-sites of a response element can occur in different orientations, and each still can interact with the zinc-fingers of a hormone receptor protein's DNA-binding domain [15,16]. The three possible structures of HREs include direct repeat, inverted repeat, and palindromic repeat of a consensus half-site. Therefore, the recurrent neural network is expected to classify a potential hormone response element into one of these HRE-like dimeric structures.

The applicability of recurrent neural networks for biological sequence analysis has recently been reviewed [17], using examples of motif detection and prediction of subcellular localization of peptides. It shows that though the network architecture reflects the presence of bias, recurrent neural networks do provide access to biologically significant patterns.

Recurrent neural networks work in an unsupervised manner [18]. Theoretically, the goal is to design a network that stores a specific set of equilibrium points such that, when an initial condition for a neural system is provided, the network eventually comes to stability at one of these equilibrium points. Thus, the network is used for classification, and has been proven to be more flexible and powerful than distance-based clustering. The network is recursive, so its output is fed back as the input iteratively. With a properly designed RNN, each input is expected to be transformed into one of the stable states of the network in a finite number of iterations [14].

In our study, three consensus of the respective HRE structures, and seven non-HRE sequences taken from the papers by Thackray et al. [19] and Lieberman et al. [20] are used as ten equilibrium points for the RNN. If the network comes to a stable state other than the ten desired equilibrium states, the input DNA sequence is 'unclustered'; otherwise, a certain cluster labeled 1 corresponding to a direct repeat (DR), 2 to an inverted repeat (IR), 3 to an palindromic repeat (ER) of the HRE half-site, and 7-10 to a non-HRE pattern is attributed to the input. Besides, as the sigmoid activation function returns non-integer values, the squared Euclidian distance between the final stable state and the nearest equilibrium point is estimated; for a successful classification, it must be 50% greater than that of the second nearest equilibrium point. In most cases, machine precision of the equilibrium points are reached within 100 iterations.

2.4 Hardware Acceleration

After numerous tests, we have reached quite a good accuracy of the system, but its very long processing time made it impossible to be used for genomic sizes which are often hundreds of million of base pairs. The bottleneck is identified at the operation of the recurrent ANN in the system. While for a given training set even with thousands of sequences, the back-propagation training of the feed-forward ANN is done once and forever, the recurrent neural network requires tens of iterations for each HRE-like input out of possible millions for genomic scales. We implemented the neural network on a 4-way IBM X260 server with four 3.16GHz CPUs, 3.25GB RAM, and 667MHz system bus. It took about 20 minutes to screen only 1Mb of DNA with eight parallel threads for RNN operation. Therefore, we propose hardware acceleration to make our system affordable in practice.

A full description of the acceleration with the Field-Programmable Gate Arrays is presented in another paper [21]. Briefly, referring to Fig.1, this FPGA-implemented RNN unit communicates with the rest of system implemented in a PC via a local bus. The fabricated RNN contains 60 neurons, which operate sequentially on ten identical neuron units. We used Verilog Hardware Description Language (Verilog HDL) in the circuit design for FPGA. The Alpha Data Virtex-4LX160 board was used for FPGA design. A Windows application reads DNA sequences predicted as HRE-like by a preceding feed-forward neural network from a text file, and sends the 32bit-long input vector to a configured FPGA board. It also obtains the output from the board, and proceeds to the decision making module.

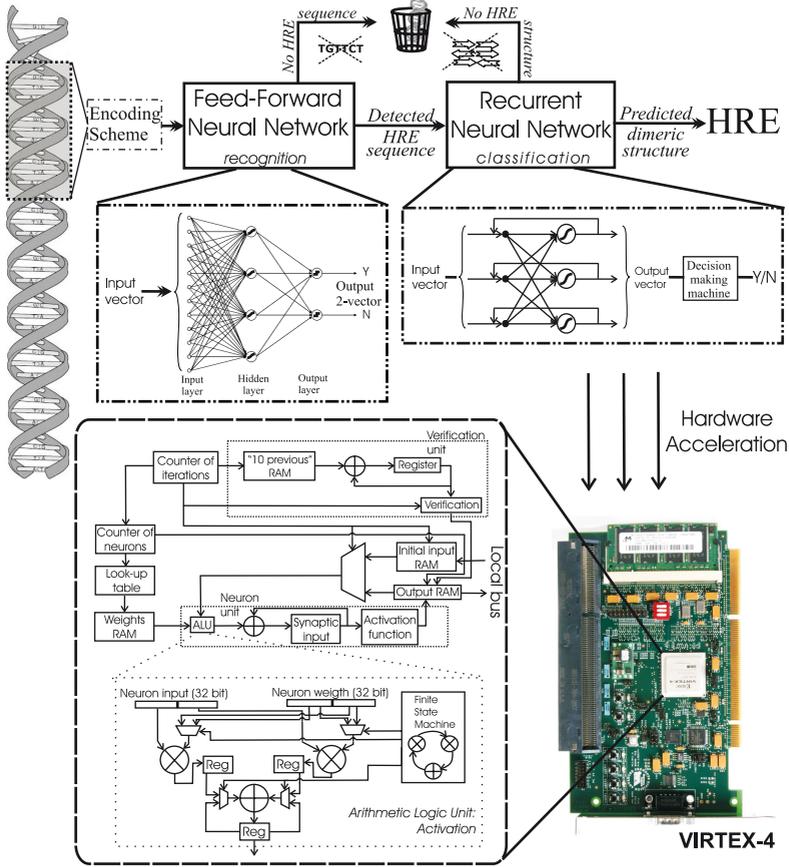


Fig. 1. A hardware-accelerated hybrid neural system for HRE prediction

Arithmetic calculations of the RNN module was implemented with use of fixed-point signed number representation, where 29 out of 32 bits were used to represent a fractional part. The sigmoid activation function was replaced by a polyline approximation, which allowed us to decrease area usage substantially with the resulting accuracy changing insignificantly.

The entire on-chip RNN system could be subdivided into two constituents: the neuron units and the control unit. The lower part of Fig. 1 shows configuration of the ANN from a point of view of arithmetic operation. Neuron units accumulate synaptic inputs of the neurons, activate them, and return current output vector into the memory.

The control unit includes the counter of iterations which defines whether an input must be put into the system for the first time (initial input). It also regulates sequence of processing of neurons by neuron units, and involves verification unit for the current output. The control unit terminates the recurrent processes if the maximum number of iterations or a stable state (no oscillations larger

then 0.2% amongst outputs of the last ten iterations) is reached. The counter of neurons is reset before each iteration starts; then, it regulates the sequence of the ten neuron units involved for calculation of sixty actual neurons, and decides which values of weights must be selected from memory.

With the RNN implementation on Virtex-4 hardware, we achieved almost ten-fold speedup for a recurrent neural network. With use of more advanced digital technologies, such as flexible on-board clock management and involvement of hyper-transport techniques instead of 66MHz PCI bus, it is possible to realize an even faster hybrid neural solution.

3 Results

3.1 HRE Classification

First, the collected dataset of experimentally validated HREs was used to test the capabilities of our RNN-based classification system. The results of classification of HREs are given in the Table 1. The difference between preferred structures of the DNA response elements for the three steroid hormones of interest – progesterone, androgen, and glucocorticoid – was found to be significant (p-value 0.007), with ARE slightly different from the two other subgroups of binding sites - it is less tolerated to direct repeats, and in most cases prefers to bind to palindromic response elements.

Table 1. Results of classification of training set of HREs with RNN

| HRE | PRE | GRE | ARE | U |
|----------------------|-------------|--------------|--------------|--------------|
| Direct repeat | 35 or 53.0% | 225 or 59.7% | 94 or 43.1% | 354 or 53.6% |
| Inverted repeat | 1 or 1.5% | 3 or 0.8% | 1 or 0.5% | 5 or 0.8% |
| Palindromic repeat | 26 or 39.4% | 114 or 30.2% | 115 or 52.8% | 255 or 38.6% |
| No dimeric structure | 4 or 6.1% | 35 or 9.3% | 8 or 3.7% | 47 or 7.1% |
| Σ | 66 | 377 | 218 | 661 |

3.2 Genome-Wide HRE Recognition

The sensitivity and specificity values were tracked on each step of machine processing; for calculations, ten-fold cross-validation is used.

The first step of the machine learning is to activate the trained feed-forward neural network, which recognizes HRE-like patterns. At this stage, the sensitivity was found to be as high as 98%, that is, 15 among 661 HREs are misclassified; the specificity is 1:5.8Kb which is measured on basis of 1Mb neutral DNA sequence.

In the next step, the recurrent neural network allows to increase specificity level to 1:7.29Kb, while the sensitivity stays favorable in comparison with

existing HRE prediction methods (for a review, see [9]) – it is as high as 92% (6 or 9% of PREs, 37 or 10% of GREs, and 9 or 4% AREs are now misclassified, or total 52 HREs).

Available vertebrate genomes were downloaded from the NCBI data repository, processed, and used for system performance measurements as well. To sum up, the average frequencies of prediction of HREs are as follows:

- Human genome (*Homo Sapiens*, #36.2) – 1:8.15Kb
- Chimpanzee genome (*Pan Troglodytes*, #2.1) – 1:8.13Kb
- Mouse genome (*Mus Musculus*, #36.1) – 1:7.69Kb
- Rat genome (*Rattus Norvegicus*, #4.1) – 1:7.11Kb
- Cow genome (*Bos Taurus*, #3.1) – 1:6.35Kb
- Dog genome (*Canis Familiaris*, #2.1) – 1:8.43Kb
- Opossum genome (*Monodelphis Domestica*, #2.1) – 1:7.36Kb
- Chicken genome (*Gallus Gallus*, #2.1) – 1:9.81Kb
- Zebrafish genome (*Danio Rerio*, #2.1) – 1:8.95Kb

We can see from the above that high accuracy of prediction is achieved by the combination of feed-forward and recurrent neural networks, which is crucial for the system to be used in practice. Sensitivity of 92% can be combined with random expectation level of 1 prediction per 7.29Kb of neutral DNA and 1:8.15Kb of human genomic DNA. The parameters, though, can be adjusted for a particular task.

4 Discussion and Conclusion

Due to the importance for regulation of vital processes, the group of steroid hormones has been studied extensively over decades, but very few computational methods have been available to aid experimentalists to find hormone receptor regulatory signals. In this paper, we propose an approach which combines feed-forward and recurrent neural networks for recognition of a subgroup of spaced, symmetrically structured DNA motifs, and demonstrate its performance using the examples of steroid hormone response elements. Its high level of accuracy provides an access to a powerful method for *de novo* HRE prediction, and further analysis of hormone-regulated genes as well.

The proposed approach benefits from the advantages of both feed-forward and recurrent neural networks; they employ different strategies of machine learning and allow to reveal different features of the patterns. The feed-forward neural network provides a very flexible tool to model almost any dataset of interest, but its flexibility may result in rather low specificity values. This is where the recurrent neural network may cut in; it is able eliminate most of the false positive HRE-like findings, comparing them against the symmetrically structured HREs.

On the other hand, training a highly accurate feed-forward neural network and especially elaborating all the RNN-based classifications with software implementation needs prohibitively long time. The hardware acceleration with reconfigurable FPGA computing provides us with a solution.

With use of two successive neural networks, we managed to model the HRE training set and separate it from the neutral DNA sequences quite reliably, but some outliers were detected. They were found through non-consensus binding sites for progesterone, androgen and glucocorticoid receptors in the promoters and gene regions for a number of genes: rabbit uteroglobin gene [22], chicken lysozyme gene [23], porcine uteroferrin gene [24], pro-opiomelanocortin gene [25], murine *c-myc* gene [26], late leader of the control region of the human polyomavirus BK [27], gene promoter of two milk protein genes (β -casein and whey acidic protein) [28], human Na/K ATPase α 1 gene promoter [29], and mouse sex-limited protein enhancer [30]. The first three are progesterone-regulated genes, the next five are glucocorticoid primary targets, and the last one is associated with androgen activity. Unless they are experimental artifacts, the possible explanation could lie in the area of complex protein-DNA interaction which is beyond DNA sequence similarity itself; it probably could be a secondary structure of DNA, or location of surrounding nucleosomes. Nevertheless, more sensitive procedures should be implemented, involving other conditions which could be related to successful formation of protein-DNA complex *in vivo*.

The dramatic progress in experimental identification of transcription factor binding sites is obvious. Thus, the availability of accurate algorithms for *in silico* binding site prediction is of great importance. Our proposed model for steroid receptor binding sites prediction can be used for determination of androgen, progesterone and glucocorticoid primary target genes with high reliability. It can also be used for detection of steroid hormone response elements *de novo*, and for evaluation of known HREs as well. Finally, the proposed model can be potentially involved for prediction of any other structured DNA motifs of interest.

Acknowledgements. This work is supported in part by two research grants, ARC02/2004 and RG50/06, from Ministry of Education, Singapore.

References

1. Larsen, P., Kronenberg, H., Melmed, S., Polonsky, K.: Williams textbook of endocrinology. Saunders, Philadelphia (2003)
2. Conneely, O.M.: Perspective: Female Steroid Hormone Action. *Endocrinology* 142(6), 2194–2199 (2001)
3. Ko, Y.J., Balk, S.P.: Targeting Steroid Hormone Receptor Pathways in the Treatment of Hormone Dependent Cancers. *Curr. Pharm. Biotechnol.* 5(5), 459–470 (2004)
4. Danielsen, M., Hinck, L., Ringold, G.M.: Two amino acids within the knuckle of the first zinc finger specify DNA response element activation by the glucocorticoid receptor. *Cell* 57(7), 1131–1138 (1989)
5. Wasserman, W.W., Sandelin, A.: Applied bioinformatics for the identification of regulatory elements. *Nat. Rev. Genet.* 5(4), 276–287 (2004)
6. Kel, A.E., Gossling, E., Reuter, I., et al.: MATCH: A tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res.* 31(13), 3576–3579 (2003)

7. Sandelin, A., Alkema, W., Engstrom, P., et al.: JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res.* 32(DB), D91–D94 (2004)
8. Rahmann, S., Muller, T., Vingron, M.: On the power of profiles for transcription factor binding site detection. *Stat. Appl. Genet. Mol. Biol.* 2(1): Article 7 (2003)
9. Stepanova, M., Lin, F., Lin, V.: Establishing a Statistic Model for Recognition of Steroid Hormone Response Elements. *Comput. Biol. Chem.* 30(5), 339–347 (2006)
10. Bajic, V.B., Tan, S.L., Chong, A., et al.: Dragon ERE Finder version 2: A tool for accurate detection and analysis of estrogen response elements in vertebrate genomes. *Nucleic Acids Res.* 31(13), 3605–3607 (2003)
11. Sandelin, A., Wasserman, W.W.: Prediction of nuclear hormone receptor response elements. *Mol. Endocrinol.* 19(3), 595–606 (2005)
12. Favorov, A.V., Gelfand, M.S., Gerasimova, A.V., et al.: A Gibbs sampler for identification of symmetrically structured, spaced DNA motifs with improved estimation of the signal length. *Bioinformatics* 21(10), 2240–2245 (2005)
13. Khorasanizadeh, S., Rastinejad, F.: Nuclear-receptor interactions on DNA-response elements. *Trends Biochem. Sci.* 26(6), 384–390 (2001)
14. Hagan, M., Demuth, H., Beale, M.: *Neural Network Design*. PWS Publishing company, Boston (1996)
15. Evans, R.M.: The steroid and thyroid hormone receptor superfamily. *Science* 240(4854), 889–895 (1988)
16. Nelson, C.C., Hendy, S.C., Shukin, R.J., et al.: Determinants of DNA sequence specificity of the androgen, progesterone, and glucocorticoid receptors: evidence for differential steroid receptor response elements. *Mol. Endocrinol.* 13(12), 2090–2107 (1999)
17. Hawkins, J., Boden, M.: The Applicability of Recurrent Neural Networks for Biological Sequence Analysis. *IEEE ACM T. Comput. BI.* 2(3), 243–253 (2005)
18. Hastie, T., Tibshirani, R., Friedman, J.: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York (2001)
19. Thackray, V.G., Lieberman, B.A., Nordeen, S.K.: Differential gene induction by glucocorticoid and progesterone receptors. *J. Steroid Biochem. Mol. Biol.* 66(4), 171–178 (1998)
20. Lieberman, B.A., Bona, B.J., Edwards, D.P., Nordeen, S.K.: The constitution of a progesterone response element. *Mol. Endocrinol.* 7(4), 515–527 (1993)
21. Stepanova, M., Lin, F., Lin, V.C.: A Hopfield Neural Classifier and Its FPGA Implementation for Identification of Symmetrically Structured DNA Motifs. *J. VLSI Sig. Proc. Syst.* (in press)
22. Jantzen, K., Fritton, H.P., Igo-Kemenes, T., et al.: Partial overlapping of binding sequences for steroid hormone receptors and DNaseI hypersensitive sites in the rabbit uteroglobin gene region. *Nucleic Acids Res.* 15(11), 4535–4552 (1987)
23. von der Ahe, D., Renoir, J.M., Buchou, T., et al.: Receptors for glucocorticosteroid and progesterone recognize distinct features of a DNA regulatory element. *Proc. Natl. Acad. Sci. USA.* 83(9), 2817–2821 (1986)
24. Lamian, V., Gonzalez, B.Y., Michel, F.J., Simmen, R.C.: Non-consensus progesterone response elements mediate the progesterone-regulated endometrial expression of the uteroferrin gene. *J. Steroid Biochem. Mol. Biol.* 46(4), 439–450 (1993)
25. Drouin, J., Trifiro, M.A., Plante, R.K., et al.: Glucocorticoid receptor binding to a specific DNA sequence is required for hormone-dependent repression of pro-opiomelanocortin gene transcription. *Mol. Cell Biol.* 9(12), 5305–5314 (1989)

26. Ma, T., Copland, J.A., Brasier, A.R., Thompson, E.A.: A novel glucocorticoid receptor binding element within the murine c-myc promoter. *Mol. Endocrinol.* 14(9), 1377–1386 (2000)
27. Moens, U., Subramaniam, N., Johansen, B., et al.: A steroid hormone response unit in the late leader of the noncoding control region of the human polyomavirus BK confers enhanced host cell permissivity. *J. Virol.* 68(4), 2398–2408 (1994)
28. Welte, T., Philipp, S., Cairns, C., et al.: Glucocorticoid receptor binding sites in the promoter region of milk protein genes. *J. Steroid Biochem. Mol. Biol.* 47(1-6), 75–81 (1993)
29. Kolla, V., Robertson, N.M., Litwack, G.: Identification of a mineralocorticoid/glucocorticoid response element in the human Na/K ATPase alpha1 gene promoter. *Biochem. Biophys. Res. Commun.* 266(1), 5–14 (1999)
30. Verrijdt, G., Schauwaers, K., Haelens, A., et al.: Functional interplay between two response elements with distinct binding characteristics dictates androgen specificity of the mouse sex-limited protein enhancer. *J. Biol. Chem.* 277(38), 35191–35201 (2002)