applied sciences

MDPI

Article

ErbB4 Is a Potential Key Regulator of the Pathways Activated by NTRK-Fusions in Thyroid Cancer

Andrey Kechin [1,2,*], Viktoriya Borobova [1,2], Alexander Kel [1,3], Anatoliy Ivanov [4] and Maxim Filipenko [1]

[1] Institute of Chemical Biology and Fundamental Medicine, Novosibirsk 630090, Russia; v.borobova@g.nsu.ru(V.B.); alexander.kel@genexplain.com (A.K.); max@niboch.nsc.ru(M.F.)
[2] Faculty of Natural Sciences, Novosibirsk State University, Novosibirsk 630090, Russia
[3] geneXplain GmbH, Wolfenbüttel 38302 Germany
[4] Altai Regional Oncological Dispensary, Barnaul 656000, Russia; anatolij0199@yandex.ru
* Correspondence: aa_kechin@niboch.nsc.ru

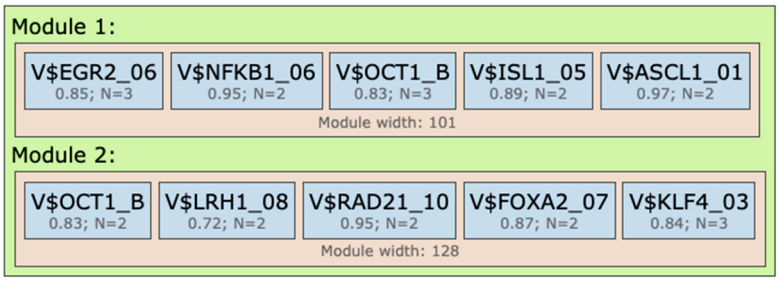
Citation: Kechin, A.; Borobova, V.; Kel, A.; Ivanov, A.; Filipenko, M. ErbB4 Is a Potential Key Regulator of the Pathways Activated by NTRK-Fusions in Thyroid Cancer. *Appl. Sci.* **2022**, *12*, 2506. https://doi.org/10.3390/app12052506

Academic Editor: Marco G. Alves

Received: 21 January 2022
Accepted: 23 February 2022
Published: 28 February 2022

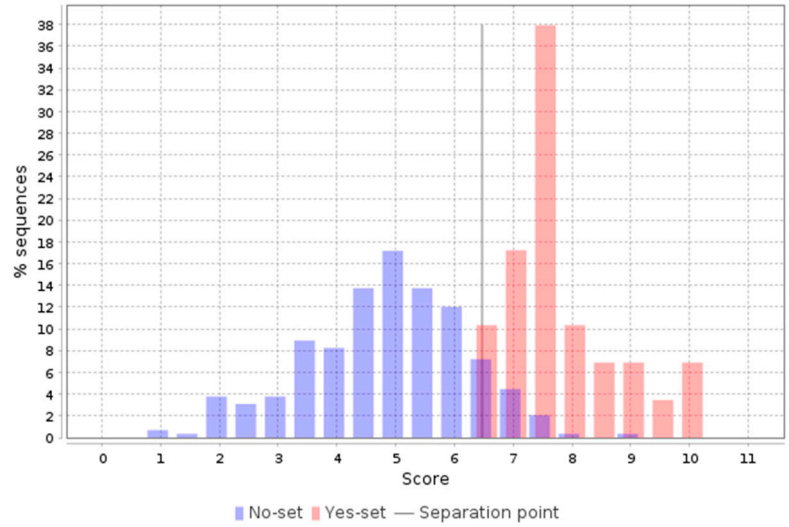
Supplement

(A)



Model score (-p*log10(pval)): 8.36
Wilcoxon p-value (pval): 2.06e-17
Penalty (p): 0.501
Average yes-set score: 7.77
Average no-set score: 4.87
AUC: 0.97

(B)



(C)

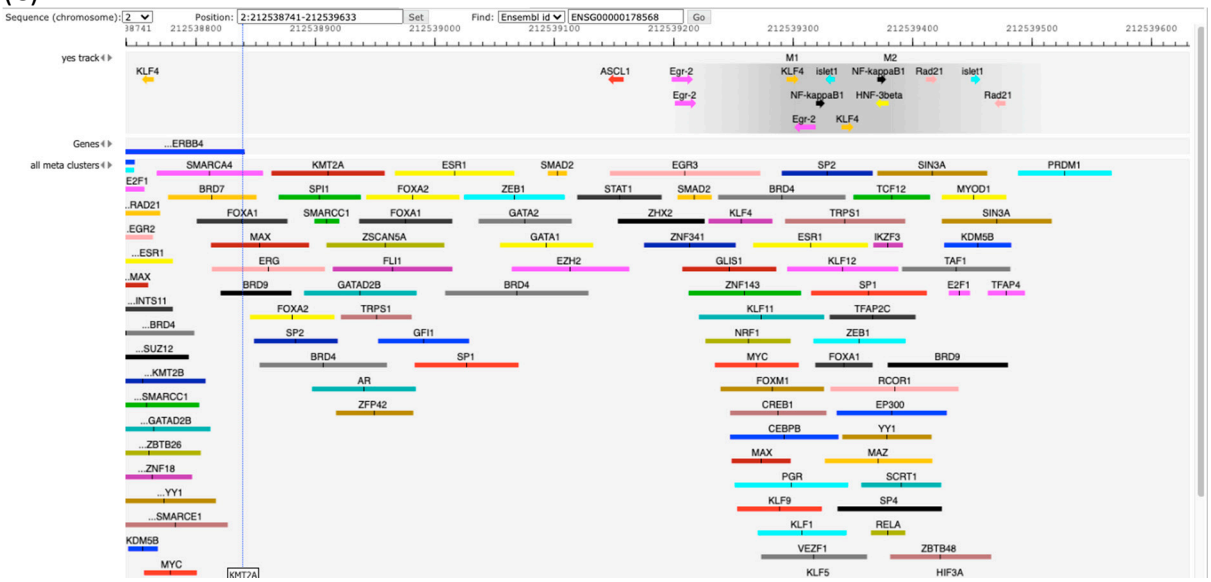


Figure S1. Composite modules (clusters) of binding sites for synergistically acting transcription factors (NFkB, Oct-1, Egr2, KLF4, HNF3B/FOXA2, and others) identified in the promoters of up-regulated genes involved in nervous system development in thyroid tumors carrying *NTRK1* and *NTRK3* gene fusions. (**A**) The structure of two composite modules identified in the promoters. Each module consists of 5 TRANSFAC position weight matrices (PWMs) with their optimized cut-offs and maximal number of top scoring sites (N), and the optimized module width. The high quality of the model is

characterized by the highly statistically significant level of the Wilcoxon test and the AUC value = 0.97. (**B**) Two histograms that characterize the promoter score distributions in the promoters of up-regulated genes (red bars) and in promoters of housekeeping genes (blue bars). There is a clear separation of these two distributions. (**C**) An example of the clusters of TF sites that belong to the composite modules identified by the algorithm in the promoter of the ERBB4 gene. The track "Gene" shows the beginning of the first exon of the gene, the dotted vertical line shows the position of the TSS. The track "yes track" represents the identified TF binding sites of the composite modules. The track "all meta clusters" shows the positions of the experimentally identified binding regions of various transcription factors from publicly available ChIP-seq data (from GTRD database []). One can see the co-localization of several predicted TF binding sites and respective experimentally identified TF binding regions for TF from the same families (Egr-2—EGR3; NF-kappaB1—RELA; KLF4—KLF4/1/5; HNF-3beta—FOXA2).