SUPPLEMENTARY MATERIAL

OPEN ACCESS

ISSN: 1875-0362

Predictive Analysis of Methylation Patterns in Oral Squamous Cell Carcinoma (OSCC) Using Machine Learning

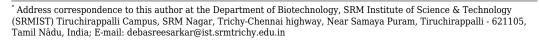


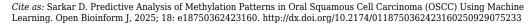
Debasree Sarkar^{1,*}

¹Department of Biotechnology, SRM Institute of Science & Technology (SRMIST) Tiruchirappalli Campus, SRM Nagar, Trichy-Chennai highway, Near Samaya Puram, Tiruchirappalli - 621105, Tamil Nādu, India; India

© 2025 The Author(s). Published by Bentham Open.

This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Public License (CC-BY 4.0), a copy of which is available at: https://creativecommons.org/licenses/by/4.0/legalcode. This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.







Received: June 24, 2025 Revised: August 7, 2025 Accepted: August 27, 2025 Published: October 07, 2025



Send Orders for Reprints to reprints@benthamscience.net

TABLE S1. Details of the top 10 CpG sites used as features in the best performing minimalistic model, including the summary test statistic for the DMR(stat), mean difference in M-values across the DMR (diff), p-value for the DMR before any correction for multiple testing (rawpval), and FDR-adjusted p-value using the Benjamini-Hochberg method (fdr), along with the genomic co-ordinates and overlapping genes (if any)

DMR	Statistical parameters				Genomic Co-ordinates and Overlapping Gene Information					
	stat	diff	rawpval	fdr	chr	pos	strand	UCSC_RefGene_Name	UCSC_RefGene_Accession	UCSC_RefGene_Group
cg00702837	-7.8553132	-0.2392556	4.41E-12	1.48E-10	chr1	46090494	-	CCDC17	NM_001114938	TSS1500
cg19045284	-5.3743876	-0.2374391	4.95E-07	2.21E-06	chr2	26572101	+	SELI	NM_033505	Body
cg09991426	-7.603319	-0.2511719	1.53E-11	3.41E-10	chr2	83316213	+	_	_	_
cg06672958	-10.120889	-0.2268153	4.85E-17	3.25E-15	chr5	106830972	+	EFNA5	NM_001962	Body
cg10093438	-7.0258164	-0.2315154	2.54E-10	3.40E-09	chr7	138089004	-	_	_	_
cg11523381	-6.3814464	-0.1777731	5.37E-09	4.50E-08	chr11	63751153	-	_	_	_
cg18275339	-6.1698024	-0.2855646	1.43E-08	1.06E-07	chr11	91760338	+	_	_	_
cg26199225	-7.0466065	-0.1897635	2.30E-10	3.40E-09	chr14	100673716	-	_	_	_
cg21387117	-6.8195749	-0.1653235	6.82E-10	6.52E-09	chr17	3288348	-	_	_	_
cg09614762	-6.8631874	-0.176285	5.54E-10	6.18E-09	chrX	65038012	-	_	_	_

stat represents the mean or combined test statistic for all CpGs in the region, often derived from moderated t-statistics from the underlying linear modelling; diff represents the average methylation difference between the two groups, where Positive values \rightarrow Hypermethylation & Negative values \rightarrow Hypomethylation; rawpval represents the original/uncorrected p-value for the DMR before any correction for multiple testing, where lower values indicate stronger evidence of differential methylation;

fdr represents the False Discovery Rate (FDR)-adjusted p-value using the Benjamini-Hochberg method, used to assess statistical significance of DMRs while controlling for multiple comparisons;

chr is the chromosome number; pos is the genomic location.

DISCLAIMER: The above article has been published, as is, ahead-of-print, to provide early visibility but is not the final version. Major publication processes like copyediting, proofing, typesetting and further review are still to be done and may lead to changes in the final published version, if it is eventually published. All legal disclaimers that apply to the final published article also apply to this ahead-of-print version.