

PhD Thesis Abstracts

English and Italian Versions

Abstract

Endocrine therapy resistance in ER α -positive breast cancer is a complex regulatory phenotype that rarely arises from alterations confined to a single molecular layer. Instead, resistance emerges through coordinated changes across transcriptional programs, chromatin state, and epigenetic regulation, motivating integrative multi-omics approaches that constrain interpretation by requiring concordant evidence across complementary assays. This thesis pursues two objectives: (i) to investigate ER α -associated regulatory dependencies in luminal breast cancer using integrative functional genomics, including the development of a pathway-first framework for cross-model methylation–expression concordance and (ii) to contribute general-purpose methodology for supervised analysis and interpretation of longitudinal multi-omics data with an intrinsic multi-way (tensor) structure.

In Part I, BRPF1 is investigated as a chromatin-associated coregulator connected to ER α signalling. By integrating ER α and BRPF1 ChIP-seq, BRPF1 ChIP-seq after ER α knockdown, ATAC-seq under BRPF1 perturbation, and RNA-seq, the analyses support an ER α -dependent recruitment of BRPF1 at a subset of shared regulatory loci and show that BRPF1 inhibition is associated with widespread accessibility losses enriched in ER α -linked regions, accompanied by coherent transcriptional attenuation. Together, these multi-layer readouts motivate BRPF1 as a candidate regulatory dependency within ER α -associated circuitry.

In Part II, endocrine resistance across heterogeneous ER α -positive cell line models is investigated by integrating EPIC DNA methylation profiling with RNA-seq. Because global methylation structure is dominated by cell line identity, the analysis shifts from CpG-level heterogeneity to pathway-level convergence, combining Hallmark GSEA, consensus leading-edge gene cores, and promoter methylation summaries to test whether recurrent functional programs show *enriched* epigenetic support in the expected inverse direction. Direction-preserving filters further assess robustness by enforcing increasing cross-model consistency at both the pathway and gene-core levels, yielding conservative but interpretable sets of recurrent pathway signatures and their associated consensus gene cores.

In Part III, `TensorPLS` is introduced as an R package for supervised analysis of three-way longitudinal omics datasets represented as tensors. The workflow combines tensor-native preprocessing and Tucker-3 imputation with PLS-based discrimination, and provides mode-resolved interpretability, including time point contribution analysis and time-resolved feature importance through VIP2D, with permutation-based robustness assessment. Evaluated on longitudinal case–control data from the TEDDY cohort across multiple omics modalities, `TensorPLS` illustrates how preserving tensor structure in preprocessing and interpretation enables structured feature prioritisation and identification of temporal windows associated with class separation.

Taken together, the chapters demonstrate that combining cross-layer evidence with structure-aware analytical workflows improves interpretability and supports robust prioritisation of candidate mechanisms and biomarkers for subsequent validation.

Riassunto

La resistenza alla terapia endocrina nel carcinoma mammario ER α -positivo rappresenta un fenotipo regolatorio complesso che raramente deriva da alterazioni confinate a un singolo livello molecolare. Al contrario, la resistenza emerge attraverso cambiamenti coordinati nei programmi trascrizionali, nello stato della cromatina e nella regolazione epigenetica, rendendo necessari approcci multi-omici integrativi che ne vincolino l'interpretazione richiedendo evidenze concordanti tra saggi complementari. Questa tesi persegue due obiettivi: (i) investigare dipendenze regolatorie associate a ER α nel carcinoma mammario luminale mediante genomica funzionale integrativa, incluso lo sviluppo di un framework pathway-first per la concordanza metilazione–espressione tra modelli distinti, e (ii) contribuire con una metodologia general-purpose per l'analisi supervisionata e l'interpretazione di dati multi-omici longitudinali caratterizzati da una struttura intrinsecamente multi-way (tensoriale).

Nella Parte I, BRPF1 viene studiato come coregolatore associato alla cromatina e connesso alla segnalazione di ER α . Integrando dati di ChIP-seq per ER α e BRPF1, ChIP-seq di BRPF1 dopo knockdown di ER α , ATAC-seq in condizioni di perturbazione di BRPF1 e RNA-seq, le analisi supportano un reclutamento di BRPF1 dipendente da ER α in un sottoinsieme di loci regolatori condivisi e mostrano che l'inibizione di BRPF1 è associata a diffuse perdite di accessibilità, arricchite in regioni collegate a ER α , accompagnate da una coerente attenuazione trascrizionale. Nel complesso, questi risultati multi-livello identificano BRPF1 come una potenziale dipendenza regolatoria all'interno del circuito associato a ER α .

Nella Parte II, la resistenza endocrina in modelli eterogenei di linee cellulari ER α -positive viene studiata integrando il profilo di metilazione del DNA ottenuto con array EPIC e dati RNA-seq. Poiché la struttura globale della metilazione è dominata dall'identità della linea cellulare, l'analisi si sposta dall'eterogeneità a livello di CpG verso la convergenza a livello di pathway, combinando Hallmark GSEA, nuclei consenso di geni leading-edge e riassunti della metilazione a livello di promotore per verificare se programmi funzionali ricorrenti mostrino un *arricchimento* di supporto epigenetico nella direzione inversa attesa. Filtri che preservano la direzionalità valutano inoltre la robustezza imponendo una crescente coerenza tra modelli sia a livello di pathway sia a livello dei nuclei genici, producendo insiemi conservativi ma interpretabili di firme di pathway ricorrenti e dei relativi nuclei consenso di geni associati.

Nella Parte III, **TensorPLS** viene presentato come un pacchetto R per l'analisi supervisionata di dataset omici longitudinali tridimensionali rappresentati come tensori. Il workflow combina preprocessing nativo per tensori e imputazione Tucker-3 con discriminazione basata su PLS, e fornisce interpretabilità risolta per modalità, inclusa l'analisi del contributo dei time point e l'importanza delle feature nel tempo tramite VIP2D, con valutazione della robustezza basata su permutazioni. Valutato su dati longitudinali caso–controllo della coorte TEDDY attraverso diverse modalità omiche, **TensorPLS** illustra come il mantenimento della struttura tensoriale nel preprocessing e nell'interpretazione permetta una prioritizzazione strutturata delle feature e l'identificazione di finestre temporali associate alla separazione tra classi.

Nel complesso, i capitoli dimostrano che la combinazione di evidenze cross-layer con workflow analitici sensibili alla struttura migliora l'interpretabilità e supporta una prioritizzazione robusta di meccanismi candidati e biomarcatori da sottoporre a successive validazioni.