Robust Single Sample Consensus Molecular Subtype Classification for Primary and Metastatic Colorectal Cancer

Andrew J. Sedgewick[1], Tushar Chandra[1], Timothy Taxter[1], Justin Guinney[1]
[1] Tempus AI, Inc., Chicago, IL

Published Abstract Number: 5059

INTRODUCTION

Consensus Molecular Subtypes (CMS) represent a well-established molecular stratification framework for colorectal cancer (CRC). Existing methods for CMS classification rely on a representative input cohort as a preprocessing step (CMScaller) or have difficulty generalizing to metastatic samples (CMSclassifier). We developed Tempus CMS to overcome these limitations. Our method normalizes gene expression data from input samples to a static reference cohort to enable single sample classification of both primary and metastatic tumors. We evaluated the performance of this classifier on a large, de-identified CRC cohort comprising 8,489 samples from primary and metastatic sites.

METHODS

To normalize input data, our method shifts and scales each gene expression value based on the mean and standard deviation of the expression of that gene in the reference cohort (n=2,787 primary and metastatic CRC). CMS calls are then generated via nearest template prediction similar to CMScaller (Eide et al., 2017). The analysis cohort was selected from clinical biopsies within 30 days of primary or metastatic diagnosis excluding reference cohort samples. CMS calls were assessed by comparing subtype prevalence to reported rates and by testing for known enrichments of pathways and genomic markers.

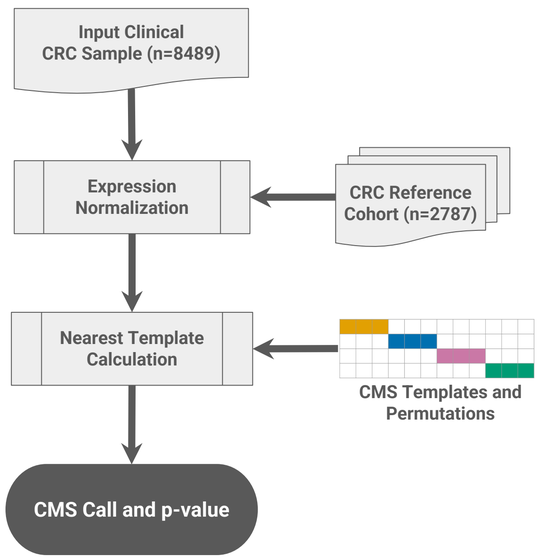

Figure 1. Workflow

SUMMARY

- Tempus CMS enables single sample classification of primary and metastatic CRC samples by normalizing gene expression data from input samples to a static reference cohort
- Application of Tempus CMS in a large cohort recapitulated previously described biology of CMS subtypes in CRC samples
- CMS subtype prevalence by sample site supports future tumor heterogeneity and evolution studies of paired primary and metastatic samples
- This tool can be used to support clinical studies requiring robust molecular stratification

RESULTS

Summary of CRC Cohorts

Characteristic	Analysis , N = 8,489	Reference , N = 2,787
Stage		
Stage 1	41 (0.6%)	15 (1.0%)
Stage 2	336 (4.6%)	94 (6.6%)
Stage 3	1,045 (14%)	150 (10%)
Stage 4	5,845 (80%)	1,171 (82%)
Unknown	1,222	1,357
Sample Site		
Liver	1,715 (20%)	494 (18%)
Lower GI	5,090 (60%)	1,625 (58%)
Other	1,684 (20%)	668 (24%)
MSI Status		
MSI	591 (7.0%)	227 (8.1%)
MSS	7,898 (93%)	2,560 (92%)
CMS		
CMS1	1,071 (13%)	379 (14%)
CMS2	2,465 (29%)	798 (29%)
CMS3	1,209 (14%)	424 (15%)
CMS4	2,780 (33%)	913 (33%)
No Call	964 (11%)	273 (9.8%)

Table 1. Characteristics of the clinical CRC samples used for development and validation of Tempus CMS.

Expression patterns of CMS genes found in all tissue types

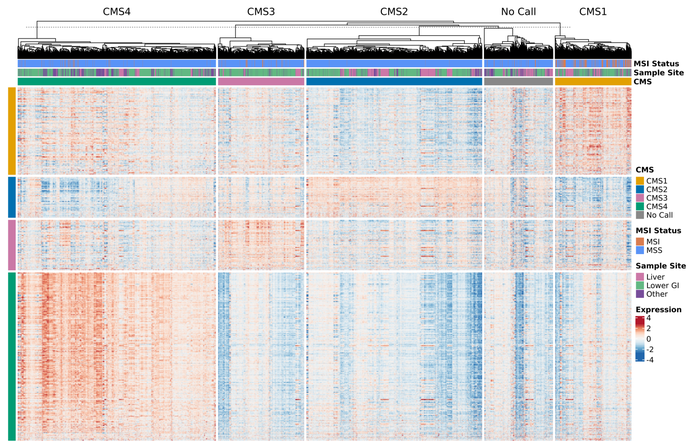

Figure 1. Normalized RNA expression of CMS template genes in analysis cohort. Expression values are log2(TPM+1) shifted and scaled by mean and standard deviation of the expression of each gene in the reference cohort.

Sample site and purity affect CMS calls

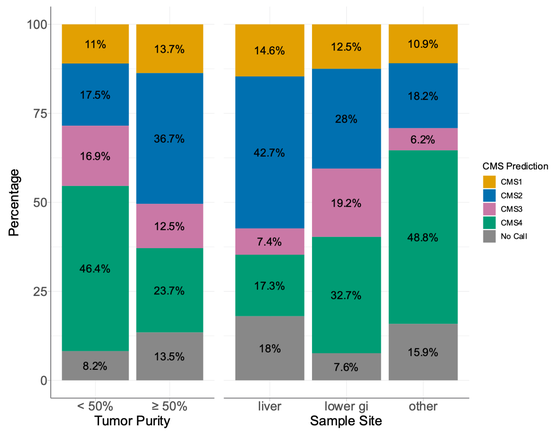

Figure 2. Subtype composition of purity and biopsy site subgroups in analysis cohort.

Tempus CMS shows expected pathway enrichment

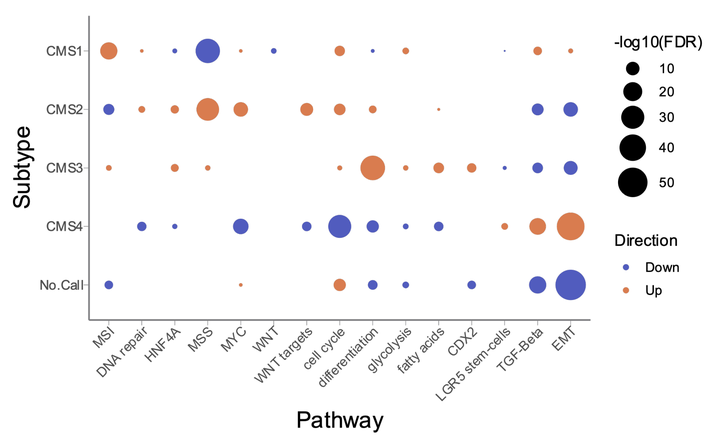

Figure 3. Pathway enrichment of RNA expression data by CMS group (FDR < 0.05) calculated using Camera (Wu and Smyth 2012). Pathway definitions from CMScaller (Eide et al., 2017) except for WNT targets (Guinney et al., 2015).

Subtypes are enriched in DNA markers

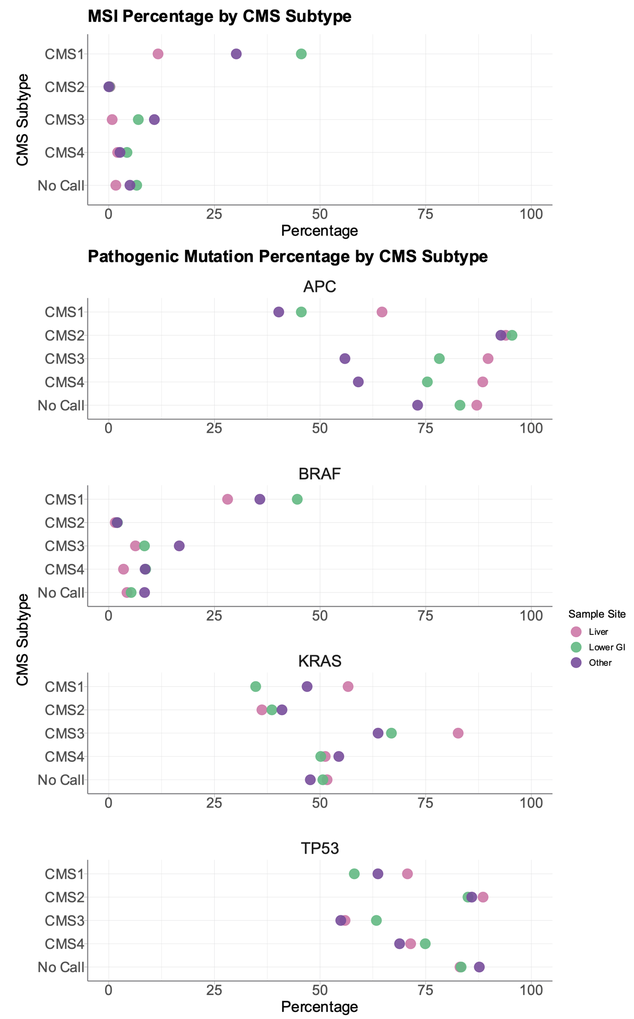

Figure 4. Prevalence of MSI status and pathogenic mutations detected by the Tempus xT assay within each CMS group and biopsy site (dot color).

REFERENCES

Eide, P.W., Bruun, J., Lothe, R.A. *et al.* CMScaller: an R package for consensus molecular subtyping of colorectal cancer pre-clinical models. *Sci Rep* 7, 16618 (2017). doi: 10.1038/s41598-017-16747-x
Guinney, J., Dienstmann, R., Wang, X. *et al.* The consensus molecular subtypes of colorectal cancer. *Nat Med* 21, 1350–1356 (2015). doi: 10.1038/nm.3967
Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic Acids Res* 40(17), e133-e133 (2012). doi: 10.1093/nar/gks461

ACKNOWLEDGMENTS

We thank Dana DeSantis from the Tempus Science Communications team for poster development.