Polygenic hazard score showed protective effect on progression-free survival of Korean colorectal cancer patients

Soojin Cha, Manu Shivakumar, Sang-Hyuk Jung, Aesun Shin, Min Jung Kim, Seung-Bum Ryoo, Seung-Yong Jeong, Kyu Joo Park, Jae Hwan Oh, Dokyoon Kim, Ji Won Park

Perelman School of Medicine, University of Pennsylvania
Sung Kyun Kwan University (SKKU)
Seoul National University Hospital (SNUH)
National Cancer Center

Backgrounds

- Colorectal cancer (CRC) is one of the common cancer worldwide.
- Considered germline variants contributing to the background of tumorigenesis, progression, and mortality, identifying survival-associated genetic components may help CRC patients' health care.
- However, the **prognostic effects of germline variants** remain unclear in CRC.
- In this study, we investigated the effect of germline variants on **progression-free survival (PFS)** of Korean CRC patients using **polygenic hazard score (PHS)**.

Methods

- We collected Korean CRC patient samples from Seoul National University Hospital (SNUH, N=911) and National Cancer Center (NCC, N=1,282).
- We performed genotyping of the samples using Korean-chip.
- After quality control and imputation, 7,362,480 variants remained.
- In the discovery phase, we estimated the effect of each variant on PFS in the NCC cohort using gwasurvivr adjusted for age, sex, tumor stage (1–4), and principal component 1–4.
- To select stronger germline variants on PFS, we performed lasso-based high-throughput logistic regression analyses on each variant of early-stage patient samples using R package snpnet.
- We did clump variants within linkage disequilibrium (LD) structure ($r^2 > 0.5$) using plink and used the remaining 326 variants as components of PHS.
- To construct PHS, we did weighted sum of beta for each variant (Fig 1).
- In the validation phase, after performing Cox assumption test for each variant on PFS (P<0.1), we did Cox proportional hazard regression analysis of the PHS in the SNUH samples.
- We estimated the predictive power using Harrell's C-index.

Methods

$$\mathrm{PHS} = \sum_{i}^{n} Xi\beta i$$

Xi: vector of a patients' genotype for the selected SNPs
βi: corresponding parameter estimates from a Cox PH regression

Fig 1. Equation of PHS

Results

- Our PHS (N SNPs=286) showed the protective effect on PFS in the SNUH CRC samples regardless of age, sex, and tumor stage.

SNUH N samples (event/tot)	PHS from	covariates	HR	LCI	UCI	P
230/911	NCC	none	0.84	0.74	0.95	0.0063
	NCC	age	0.84	0.74	0.95	0.0062
	NCC	age, sex	0.84	0.73	0.95	0.0062
	NCC	age, sex, stage	0.84	0.74	0.95	0.0075

- With PHS, clinical predictors increased the predictability of PFS on CRC patients compared to without the PHS.

Predictor	C-index (SE)
Age	0.529 (0.020)
Sex	0.497 (0.017)
Stage	0.698 (0.016)
Age, Sex, Stage	0.709 (0.017)
PHS	0.554 (0.019)
Age, Sex, Stage, PHS	0.723 (0.016)

- When stratifying patients according to tumor stage, late-stage patients showed significant association with PFS adjusted for age and sex.

Tumor stage	SNUH N samples (event/tot)	PHS from	HR	LCI	UCI	P
1	16/172	NCC	1.21	0.69	2.15	0.51
2	38/249	NCC	0.78	0.55	1.10	0.15
3	82/324	NCC	0.77	0.63	0.94	0.01
4	94/155	NCC	0.98	0.78	1.23	0.85
Early (1/2)	54/421	NCC	0.85	0.64	1.14	0.29
Late (3/4)	176/479	NCC	0.83	0.72	0.96	0.01

Results

- Late-stage CRC patients with higher PHS showed significantly better PFS than patients with lower PHS (stage 3 and 4, P=0.012, Fig 2).

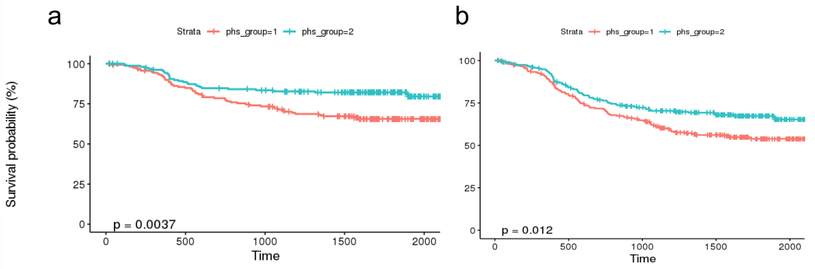

Fig 2. Kaplan-Meier plot of the late-stage CRC patients (a, stage 3; b, stage 3–4) stratified by PHS.

Conclusion

- The PHS (N SNPs = 286) showed the protective effect on PFS in the Korean CRC patients irrespective to age, sex and tumor stage.
- Prediction model with PHS and clinical variables showed better predictive performance compared to those without PHS (C-index = 0.723 vs. 0.709).
- Late-stage patients seem to be more associated with the PHS.
- Patients with the higher PHS showed better PFS compared to patients with the lower PHS in the late-stage group.

Future direction

- We will investigate the relationship of PHS and clinical factors including inflammatory components of CRC.
- By annotating the variants, we will see the functional meaning of the PHS on PFS of CRC patients.
- For generalize, we will validate the impact of our PHS in the external dataset.