## SUPPLEMENTAL METHODS FOR REPLICATION ANALYSIS

To replicate our observations for LRIG1 copy number and recurrence, we identified 18 publically available breast cancer datasets with both gene expression data and outcomes including samples with information on distant metastasis-free survival (DMFS), total sample size of 1576), and with overall survival, total sample size 791.

We used the NCBI 36.1 start and end locations of LRIG1 (chr3:66,511,911-66,633,535) and SLC25A26 (chr3:66,376,316-66,512,041) in order to find the associated U133Plus2 probesets that aligned with our most significant MIP probes.

The sample set is described in *Table 1*.

**Table 1. Gene expression data set description**

| Dataset | Samples | OS | RFS | DMFS | Title |
|---|---|---|---|---|---|
| GSE11121 | 200 | 0 | 0 | 200 | The humoral immune system has a key prognostic impact in node-negative breast cancer |
| GSE12093 | 136 | 0 | 0 | 136 | The 76-gene Signature Defines High-Risk Patients that Benefit from Adjuvant Tamoxifen Therapy |
| GSE12276 | 204 | 0 | 204 | 0 | Expression data from primary breast tumors |
| GSE1456 | 159 | 159 | 159 | 0 | Gene expression of breast cancer tissue in a large population-based cohort of Swedish patients |
| GSE16391 | 55 | 0 | 55 | 0 | GGI: a potential predictor of relapse for endocrine-treated breast cancer patients in the BIG 1-98 trial |
| GSE16446 | 120 | 107 | 0 | 107 | Multifactorial Approach to Predicting Resistance to Anthracyclines |
| GSE17705 | 298 | 0 | 0 | 298 | Endocrine Sensitivity Index Validation Dataset |
| GSE17907 | 55 | 0 | 0 | 39 | Molecular profiling of ERBB2-amplified breast cancers |
| GSE19615 | 115 | 0 | 0 | 115 | Integrated genomic and function characterization of the 8q22 gain |
| GSE2034 | 286 | 0 | 286 | 0 | Breast cancer relapse free survival |
| GSE20685 | 327 | 327 | 0 | 83 | Microarray-based molecular subtyping of breast cancer |
| GSE21653 | 266 | 0 | 252 | 0 | A gene expression signature identifies two prognostic subgroups of basal breast cancer |
| GSE4922_UPPSALA | 249 | 0 | 249 | 0 | Genetic Reclassification of Histologic Grade Delineates New Clinical Subtypes of Breast Cancer |

| | | | | | |
|---|---|---|---|---|---|
| GSE6532_GPL570 | 87 | 0 | 87 | 87 | Definition of clinically distinct molecular subtypes in estrogen receptor positive breast carcinomas using genomic grade |
| GSE6532_GPL96 | 327 | 0 | 127 | 249 | Definition of clinically distinct molecular subtypes in estrogen receptor positive breast carcinomas using genomic grade |
| GSE7390 | 198 | 198 | 198 | 198 | Strong Time Dependence of the 76-Gene Prognostic Signature |
| GSE9195 | 77 | 0 | 77 | 77 | Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen |
| GSE18229 (UNC337) | 337 | 254 | 255 | 0 | Phenotypic and Molecular Characterization of the Claudin-low Intrinsic Subtype of Breast Cancer |

| | | | | |
|---|---|---|---|---|
| Total | 3496 | 1045 | 1949 | 1589 |

We found 3 probe sets for *LRIG1* and 1 for *SLC25A26*, *see Table 2*.

**Table 2. Probe sets for LRIG1 and SLC25A26**

| ***SLC25A26*** | | | |
|---|---|---|---|
| probe.set | chr | start | end |
| 225862_at | 3 | 66376316 | 66512037 |
| ***LRIG1*** | | | |
| probe.set | chr | start | end |
| 211596_s_at | 3 | 66511910 | 66634041 |
| 236173_s_at | 3 | 66548059 | 66633398 |
| 238339_x_at | 3 | 66546059 | 66633398 |

We used the mean expression of the available probesets as a gene-level summary. In order to adjust for batch effects, we categorized the expression values within each dataset into low (low 30%), high (high 30%) and neutral. Since the available datasets have data either from the U133A or U133Plus2 platforms, not all probe sets are represented in all datasets. Table 3 shows the number of samples with data for *LRIG1* and *SLC25A26*.

**Table 3. Datasets with available data for SLC25A26 and LRIG1**

| Dataset | Total | SLC25A26 | LRIG1 |
|---|---|---|---|
| GSE11121 | 200 | No | Yes |
| GSE12093 | 136 | No | Yes |
| GSE12276 | 204 | Yes | Yes |
| GSE1456 | 159 | No | Yes |
| GSE16391 | 55 | Yes | Yes |
| GSE16446 | 120 | Yes | Yes |
| GSE17705 | 298 | No | Yes |
| GSE17907 | 55 | Yes | Yes |
| GSE19615 | 115 | Yes | Yes |
| GSE2034 | 286 | No | Yes |
| GSE20685 | 327 | Yes | Yes |
| GSE21653 | 266 | Yes | Yes |
| GSE4922_UPPSALA | 249 | No | Yes |
| GSE6532_GPL570 | 87 | Yes | Yes |
| GSE6532_GPL96 | 327 | No | Yes |
| GSE7390 | 198 | No | Yes |
| GSE9195 | 77 | Yes | Yes |
| GSE18229 (UNC337) | 337 | No | No |
| Total | 3496 | 1306 | 3159 |

# SUPPLEMENTAL TABLES

**Supplemental Table 1.** *LRIG1* loss by race/ethnicity and tumor subtype

| | *LRIG1* no loss | *LRIG1* loss | *p* value* |
|---|---|---|---|
| All Cases | | | |
| NHW | 660 (92.3) | 55 (7.7) | |
| Black | 109 (87.2) | 16 (12.8) | |
| Hispanic | 108 (87.8) | 15 (12.2) | 0.08 |
| Luminal A | | | |
| NHW | 281 (94.9) | 15 (5.1) | |
| Black | 29 (93.5) | 2 (6.5) | |
| Hispanic | 44 (97.8) | 1 (2.2) | 0.67 |
| Luminal B | | | |
| NHW | 96 (92.3) | 8 (7.7) | |
| Black | 16 (84.2) | 3 (15.8) | |
| Hispanic | 15 (83.3) | 3 (16.7) | 0.37 |
| HER2+ | | | |
| NHW | 121 (88.3) | 16 (11.7) | |
| Black | 29 (85.3) | 5 (14.7) | |
| Hispanic | 26 (86.7) | 4 (13.3) | 0.57 |
| TNBC | | | |
| NHW | 104 (87.4) | 15 (12.6) | |
| Black | 26 (83.9) | 5 (16.1) | |
| Hispanic | 19 (82.6) | 4 (17.4) | 0.57 |

Abbreviations: HER2, human epidermal growth factor receptor 2; NHW, non-Hispanic white; TNBC, triple-negative breast cancer
*Fisher's exact test

**Supplemental Table 2.** Region and probe level *LRIG1* copy number status (loss, gain, or normal) and clinicopathologic characteristics

| Characteristic‡ | LRIG1$^{1-5}$ | | LRIG1$^{6-11}$ | |
|---|---|---|---|---|
| Race | loss (%) | gain (%) | loss (%) | gain (%) |
| Non-hispanic white (n=715) | **93 (13)*** | 155 (21.7) | **130 (18.2)*** | 73 (10.2) |
| Black (n=125) | **26 (20.8)** | 28 (22.4) | **35 (28)** | 11 (8.8) |
| Hispanic (n=123) | **25 (20.3)** | 20 (16.3) | **27 (22)** | 12 (9.8) |
| Stage | | | | |
| I (n=304) | 52 (17.1) | 57 (18.8) | 67 (22) | 26 (8.6) |
| II (n=662) | 93 (14) | 146 (22.1) | 124 (18.7) | 71 (10.7) |
| Age at diagnosis (y) | | | | |
| <50 (n=394) | 55 (14) | 79 (20.1) | 80 (20.3) | 33 (8.4) |
| ≥50 (n=555) | 87 (15.7) | 116 (20.9) | 110 (19.8) | 59 (10.6) |
| Intrinsic subtype# | | | | |
| Luminal A (n=373) | 46 (12.3) | 85 (22.8) | 63 (16.9) | 37 (9.9) |
| Luminal B (n=145) | 21 (14.5) | 28 (19.3) | 27 (18.6) | 19 (13.1) |
| HER2+ (n=203) | 32 (15.8) | 42 (20.7) | 47 (23.2) | 23 (11.3) |
| TNBC (n=174) | 36 (20.7) | 26 (14.9) | 45 (25.9) | 13 (7.5) |
| ER status | | | | |
| Negative (n=293) | **57 (19.5)*** | **43 (14.7)*** | **70 (23.9)*** | 22 (7.5) |
| Positive (n=666) | **84 (12.6)** | **158 (23.7)** | **120 (18)** | 75 (11.3) |
| HER2 status | | | | |
| Negative (n=768) | 113 (14.7) | 161 (21) | 145 (18.9) | 74 (9.6) |
| Positive (n=203) | 32 (15.8) | 42 (20.7) | 47 (23.2) | 23 (11.3) |
| HER2/ER status | | | | |
| Her2+/ER+ (n=115) | 14 (12.2) | **32 (27.8)** | 26 (22.6) | 17 (14.8) |
| Her2+/ER– (n=84) | 16 (19) | **10 (11.9)** | 20 (23.8) | 6 (7.1) |
| Lymph node status | | | | |
| Negative (n=565) | **102 (18.1)*** | 112 (19.8) | 124 (21.9) | 54 (9.6) |
| Positive (n=383) | **40 (10.4)** | 83 (21.7) | 66 (17.2) | 38 (9.9) |
| Endocrine therapy | | | | |
| Yes (n=422) | 56 (13.3) | 94 (22.3) | 73 (17.3) | **51 (12.1)** * |
| No (n=522) | 85 (16.3) | 100 (19.2) | 115 (22) | **40 (7.7)** |
| Chemotherapy | | | | |
| None (n=480) | 74 (15.4) | 102 (21.2) | 94 (19.6) | **51 (10.6)*** |
| Anthracycline (n=323) | 42 (13) | 62 (19.2) | 66 (20.4) | **21 (6.5)** |
| Anthracycline/taxane (n=114) | 21 (18.4) | 26 (22.8) | 24 (21.1) | 18 (15.8) |
| Nuclear grade† | | | | |
| I (n=92) | 11 (12) | 15 (16.3) | 15 (16.3) | 8 (8.7) |

| | | | | |
|---|---|---|---|---|
| II (n=477) | 67 (14) | 116 (24.3) | 87 (18.2) | 53 (11.1) |
| III (n=336) | 58 (17.3) | 59 (17.6) | 80 (23.8) | 29 (8.6) |
| Tumor size (cm) | | | | |
| < 2 (n=566) | 85 (15) | 113 (20) | 109 (19.3) | 50 (8.8) |
| ≥ 2 (n=369) | 55 (14.9) | 79 (21.4) | 78 (21.1) | 39 (10.6) |

Abbreviations: As illustrated in Figure 1, $LRIG1^{1-5}$ includes 5 probes from chromosome position 66,532,949–66,596,637; $LRIG1^{6-11}$ includes 6 probes from chromosome position 66,512,700–66,515,666; ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; TNBC, triple-negative breast cancer.

*Indicates significant at ≤ 0.05 as loss vs. no loss or gain vs. no gain

‡ Numbers do not add up to the column totals due to missing values.

# Tumor subtype was determined using ER, PR, Ki67, and HER2 as defined in Materials and Methods.

† Nuclear grade was determined by the modified Black's method.

## Supplemental Table 3

## Multivariate Cox models for patient subgroups (A, no chemotherapy; B, chemotherapy; C, no treatment [no chemo/no endocrine])

**Supplemental Table 3A**. Multivariate Cox proportional hazards model for risk of distant metastasis and overall survival for copy number imbalance in LRIG1 in patients receiving ***no chemotherapy*** patients, n=480.

| | **Distant Metastasis** | | **Overall survival** | |
|---|---|---|---|---|
| Patient or tumor characteristic | HR (95% CI) | *p* value | HR (95% CI) | *p* value |
| Age at diagnosis (y) | | | | |
| < 50 | Reference | | Reference | |
| ≥ 50 | 1.63(0.99–2.69) | 0.054 | 0.66 (0.44–0.99) | 0.044 |
| Tumor size (cm) | | | | |
| < 2 | Reference | | Reference | |
| ≥ 2 | 2.57 (1.63–4.04) | <0.0001 | 1.90 (1.43–2.53) | <0.0001 |
| Lymph node status | | | | |
| Negative | Reference | | Reference | |
| Positive | 1.51 (0.93–2.45) | 0.09 | 1.29 (0.95–1.77) | 0.1 |
| *LRIG1* | | | | |
| Copy normal | Reference | | Reference | |
| **Loss** | **3.07 (1.76–5.35)** | **<0.0001** | **1.69 (1.09–2.62)** | **0.018** |
| Gain | 1.11 (0.40–3.09) | 0.84 | 0.89 (0.45–1.77) | 0.75 |

**Supplemental Table 3B.** Multivariate Cox proportional hazards model for risk of distant metastasis and overall survival for copy number imbalance in LRIG1 in patients receiving ***chemotherapy***, n=437.

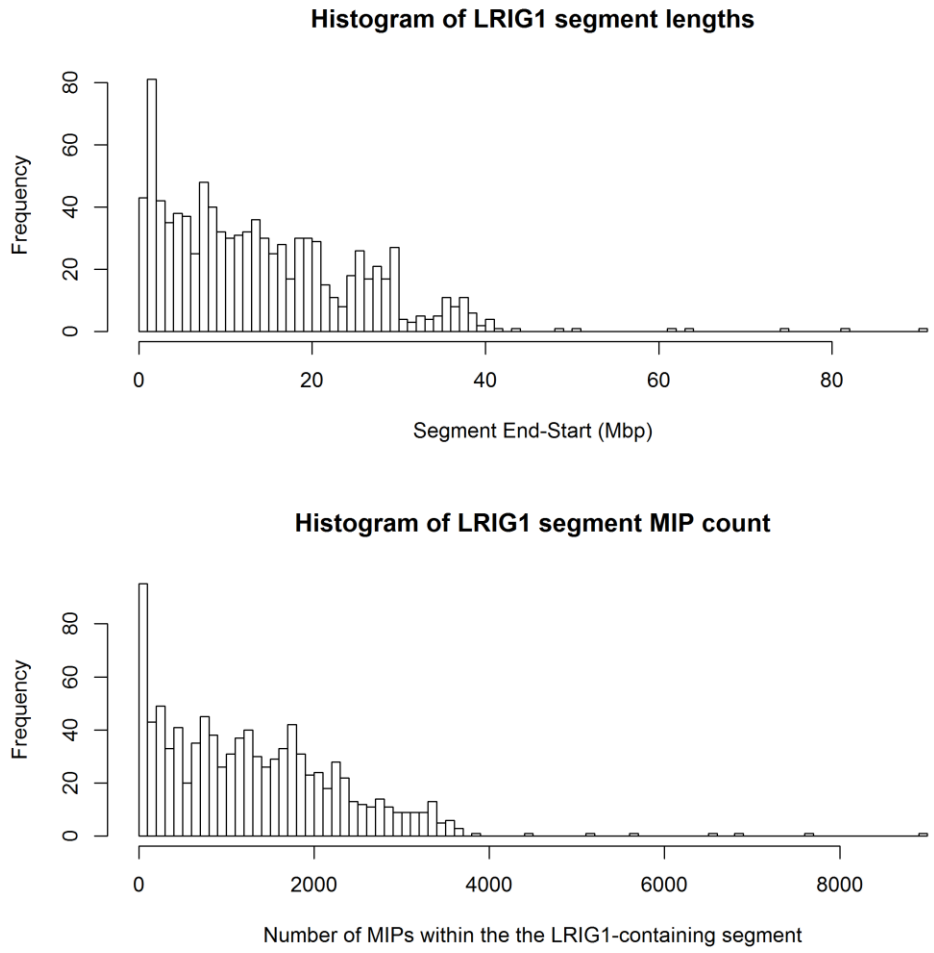
| | **Distant Metastasis** | | **Overall survival** | |
|---|---|---|---|---|
| Patient or tumor characteristic | HR (95% CI) | *p value* | HR (95% CI) | *p* value |
| Age at diagnosis (y) | | | | |
| < 50 | Reference | | Reference | |
| ≥ 50 | 1.35(0.91–2.01) | 0.14 | 0.93(0.65–1.34) | 0.71 |
| Tumor size (cm) | | | | |
| < 2 | Reference | | Reference | |
| ≥ 2 | 1.68 (1.16–2.44) | 0.006 | 1.42 (0.99–2.01) | 0.053 |
| Lymph node status | | | | |
| Negative | Reference | | Reference | |
| Positive | 2.19 (1.45–3.29) | 0.0002 | 1.98 (1.36–2.91) | 0.0004 |
| *LRIG1* | | | | |
| Copy normal | Reference | | Reference | |
| **Loss** | **1.53 (0.87–2.67)** | **0.14** | **1.66 (0.99–2.78)** | **0.053** |
| Gain | 0.81 (0.26–2.56) | 0.72 | 1.07 (0.39–2.92) | 0.89 |

**Supplemental Table 3C.** Multivariate Cox proportional hazards model for risk of distant metastasis and overall survival for copy number imbalance in LRIG1 in patients receiving **no *chemotherapy or endocrine treatment***, n=225.

| | **Distant Metastasis** | | **Overall survival** | |
|---|---|---|---|---|
| Patient or tumor characteristic | HR (95% CI) | *p value* | HR (95% CI) | *p* value |
| Age at diagnosis (y) | | | | |
| < 50 | Reference | | Reference | |
| ≥ 50 | 1.81(0.96–3.39) | 0.07 | 0.67 (0.42–1.07) | 0.094 |
| Tumor size (cm) | | | | |
| < 2 | Reference | | Reference | |
| ≥ 2 | 3.36 (1.77–6.36) | <0.0001 | 2.37 (1.61–3.50) | <0.0001 |
| Lymph node status | | | | |
| Negative | Reference | | Reference | |
| Positive | 1.41 (0.58–3.41) | 0.44 | 1.08 (0.59–1.96) | 0.8 |
| *LRIG1* | | | | |
| Copy normal | Reference | | Reference | |
| Loss | **3.73 (1.80–7.72)** | **<0.0001** | **1.92 (1.07–3.46)** | **0.03** |
| Gain | - | 1 | 0.69 (0.24-1.95) | 0.48 |

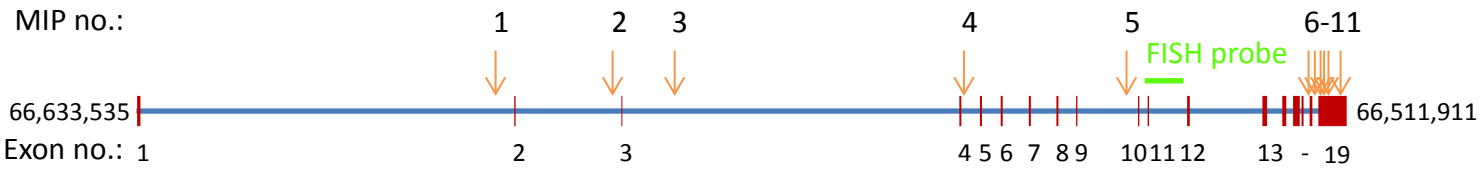
## SUPPLEMENTAL FIGURES

**Supplemental Figure 1**.

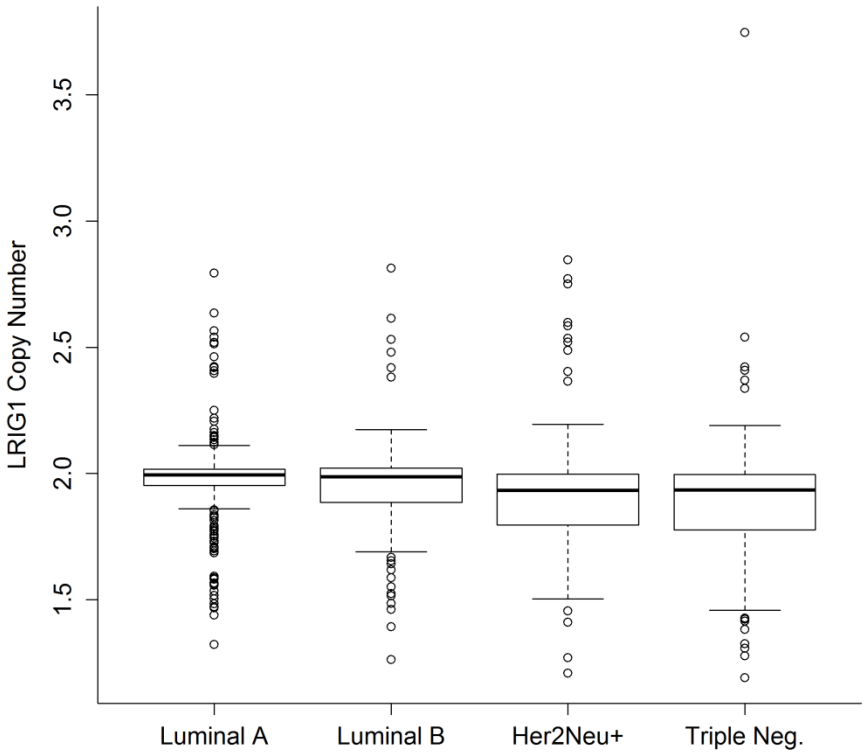



**Supplemental Figure 1. LRIG1 segment lengths and probe count**. The top histogram shows the distribution of the segment lengths in Mega-base pairs. On average, the segment the include LRIG1 is quite big compared to the length of LRIG1 (122kbp), median/mean values are 12.05/14.16 Mbp. The bottom histogram shows the distribution of MIPs that were averaged in each segment. The median/mean numbers of MIPs per segment are 1183/1313.

**Supplemental Figure 2. LRIG1 gene and MIP probe location**

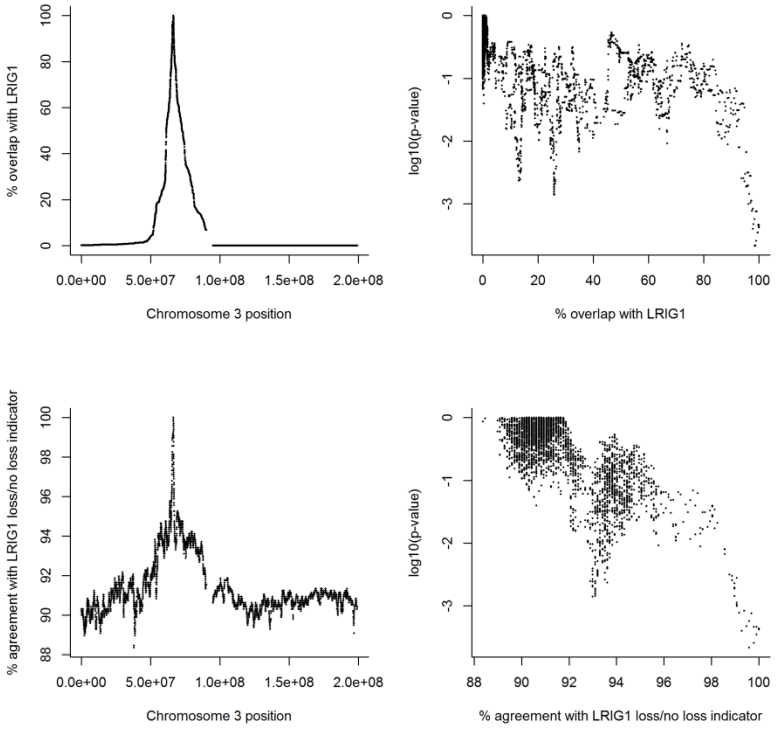


**Supplemental Figure 3**.



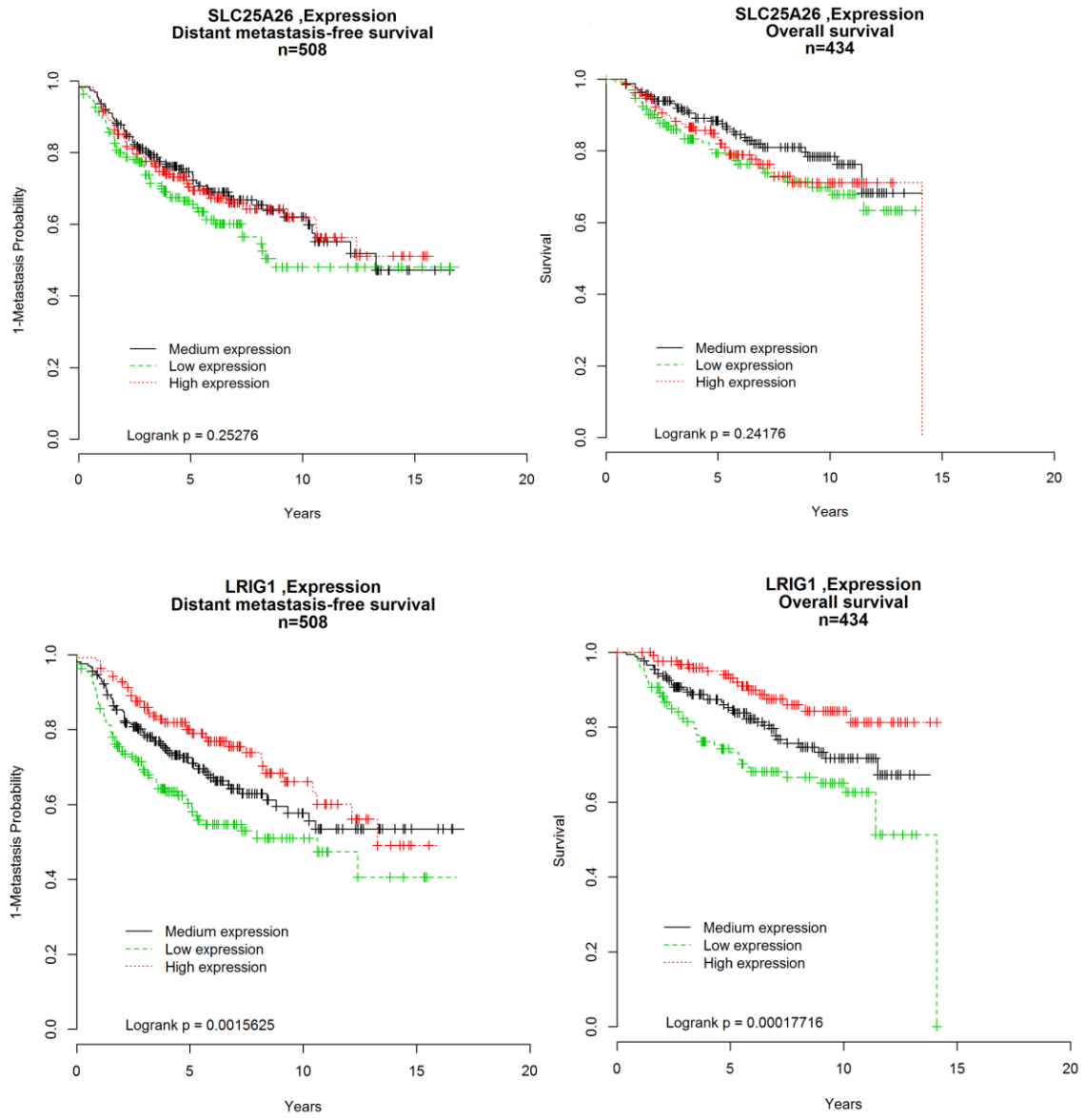

**Supplemental Figure 3**. Box plots of the entire cohort (n=971) of the copy number at the LRIG1 locus for each intrinsic subtype (Kruskal-Walis p=3e-12).

**Supplemental Figure 4.**




**Supplemental Figure 4**. **Computed average overlap with LRIG1 across all samples**. Here, we computed, for a particular sample and a particular MIP the average overlap of MIPs in a segment that includes LRIG1. The average overlap means the percentage of samples for which a given MIP belongs to the same locus as LRIG1 and thus shares the same smoothed CN value Thus, the % overlap with LRIG1 would be 100% if the MIP belongs to a segment that always includes LRIG1 (for all samples). This only happens for the 11 MIPs in the LRIG1 region. The top-left panel shows the % overall across chr3; because of the large length of the segments, there is a large area with MIPs that have non-zero overlap with LRIG1. The top-right panel plots the % overlap of all MIPs in chromosome 3 versus the log10 p-value for the association with recurrence. The only MIPs showing low p-values are also the ones with very high overlap with LRIG1. Another way to show the importance of LRIG1 is shown in the bottom panels. Instead of using % overlap, we use % agreement of the loss/no loss calls for each MIP with the LRIG1 loss/no loss calls. The bottom-right panel shows that only a few MIPs with very high (>99%) agreement (correlation) with LRIG1 losses show small p-values. All those MIPs belong either to LRIG1 or to SLC25A26.

**Supplemental Figure 5.**




**Supplemental Figure 5**. Unlike its neighbor LRIG1, we observed no evidence for an association between the level of expression of SLC25A26 and DMFS or OS, though notably the available samples with probe data for SLC25A26 were smaller than that for LRIG1.