**Supplementary Materials**

Fig. S1. Phenotype-phenotype associations

Fig. S2. Volume by mutation.

Fig. S3. Location and cystic structure VAF separated by individual mutation.

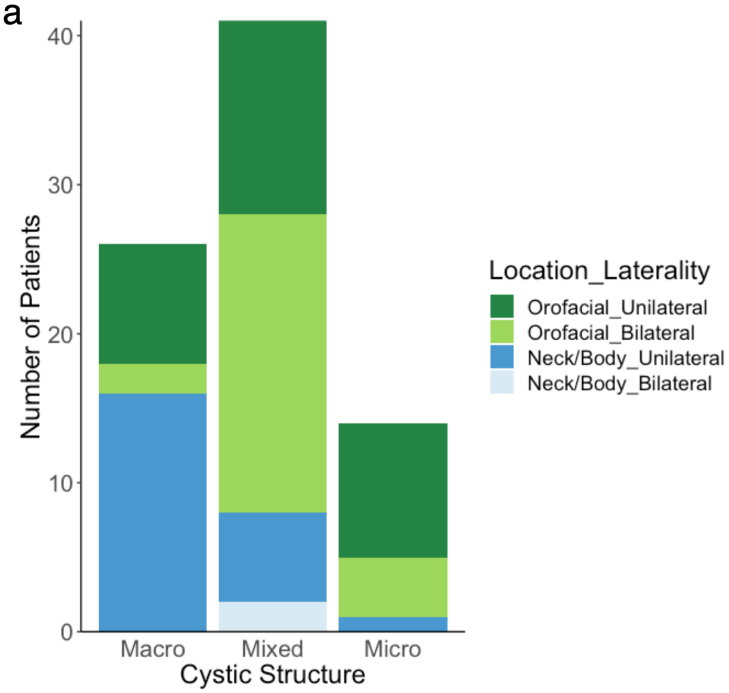
Fig. S4. Head and neck LM staging system and genotype/stage results.

Table S1. Phenotype data and sequencing results by individual. (See separate .xlsx file)

Table S2. Genes included in smMIP panel.

Table S3. Correlation of variant frequency with cell growth assays from Dogruluk et al. 2015 (31).

a




b

**Number of Procedures**

| | 0-1 | 2-5 | ≥6 |
|---|---|---|---|
| **Neck/Body** | 15 | 8 | 2 |
| **Orofacial** | 16 | 16 | 24 |

**p = 0.006**

**Number of Procedures**

| | 0-1 | 2-5 | ≥6 |
|---|---|---|---|
| **Macro** | 17 | 9 | 0 |
| **Mixed** | 10 | 10 | 21 |
| **Micro** | 4 | 5 | 5 |

**p < 0.0001**

**Number of Procedures**

| | 0-1 | 2-5 | ≥6 |
|---|---|---|---|
| **Unilateral** | 28 | 19 | 6 |
| **Bilateral** | 3 | 5 | 20 |

**p < 0.00001**

**Figure S1: Phenotype-phenotype associations.** (a) Histogram demonstrating relationship between location, laterality, and cystic structure. Unilateral neck and body lesions are predominantly macrocystic while microcystic lesions are predominantly in the orofacial region. (b) Contingency tables for number of procedures and location, laterality, and cystic structure demonstrating and increased number of procedures for orofacial, microcystic, and bilateral LMs.

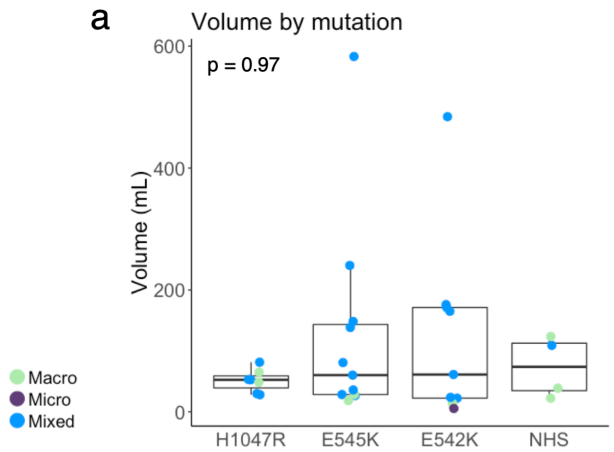



**Figure S2: Volume by mutation.** (a) Dot-box plot of volume by mutation (n = 31). *P* value indicates result of Kruskal-Wallis comparison. Boxes are defined by the 1st quartile inferiorly, median, and 3rd quartile superiorly with whiskers extending to the farthest non-outlier point (defined as within 3/2 times the interquartile range). Dot colors correspond to cystic structure: green – macrocystic, purple – microcystic, and blue – mixed cystic. Abbreviations: mL – milliliters.

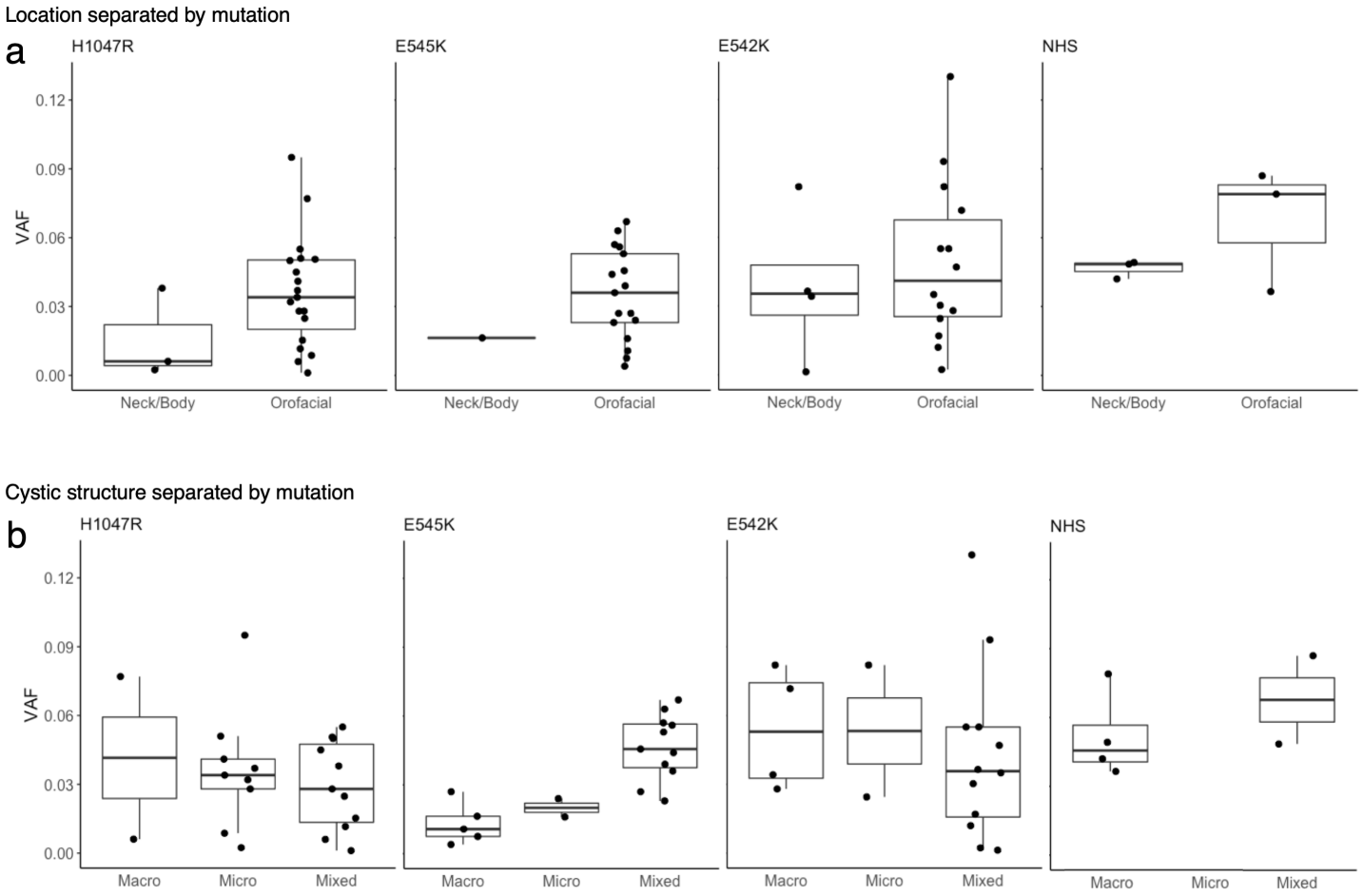



**Figure S3: Location and cystic structure VAF separated by individual mutation.** Dot-box plots of VAF by (a) location and (b) cystic structure separated by individual mutations (H1047R, n = 22; E545K, n = 18; E542K, n = 18; NHS, n = 6). Boxes are defined by the 1$^{st}$ quartile inferiorly, median, and 3$^{rd}$ quartile superiorly with whiskers extending to the farthest non-outlier point (defined as within 3/2 times the interquartile range). Abbreviations: NHS – non-hotspot mutations. VAF – variant allele fraction.

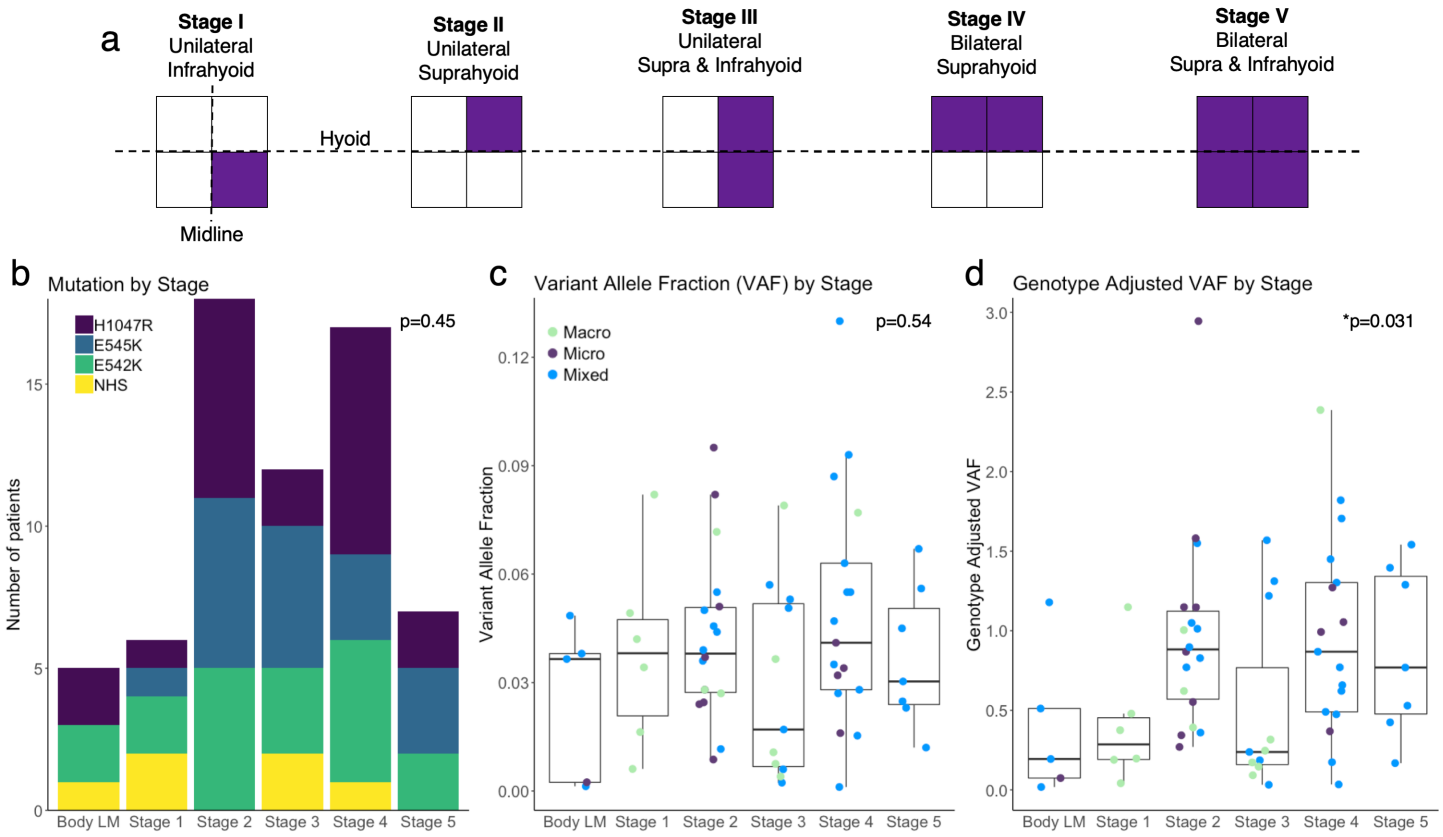



**Figure S4: Head and neck LM staging system and genotype/stage results.** (a) Schematic exemplifying the five de Serres clinical stages for head and neck LMs. (b) Histogram showing variation of mutations across stage. *P* value is for Fisher exact test. (c) Dot-box plot of VAF plotted by stage and (d) genotype adjusted VAF plotted by stage. *P* values are for Kruskal-Wallis tests. Dot colors correspond to cystic structure (pale green: macrocystic, purple: microcystic, blue: mixed cystic).

Abbreviations: LM – lymphatic malformation. NHS – non-hotspot mutations. VAF – variant allele fraction.

**Table S2. Genes included in smMIP panel.**

| Gene | # of MIPs |
|---|---|
| ACVRL1 | 24 |
| AKT3 | 23 |
| ARAF | 33 |
| CCBE1 | 20 |
| CCM2 | 28 |
| CELSR1 | 127 |
| CTNNB1 | 31 |
| DCHS1 | 123 |
| ENG | 30 |
| EPHB4 | 43 |
| FAT4 | 159 |
| FGFR1 | 42 |
| FLT4 | 72 |
| FOXC2 | 18 |
| GATA2 | 21 |
| GDF2 | 14 |
| GJC2 | 15 |
| GLMN | 30 |
| GNA11 | 18 |
| GNAQ | 14 |
| HGF | 35 |
| HRAS | 13 |
| IDH1 | 19 |
| IDH2 | 24 |
| ITGA9 | 54 |
| KIF11 | 47 |
| KRAS | 10 |
| KRIT1 | 34 |
| MAP2K1 | 19 |
| MAP3K3 | 35 |
| NRAS | 8 |
| PDCD10 | 9 |
| PDGFRB | 60 |
| PIEZO1 | 133 |
| PIK3CA | 42 |
| PIK3R6 | 38 |
| PORCN | 26 |
| PTEN | 16 |
| PTPN14 | 48 |
| RASA1 | 48 |
| SMAD4 | 23 |
| SOX18 | 14 |
| TEK | 49 |
| VEGFC | 17 |

**Table S3. Correlation of variant frequency with cell growth assays from Dogruluk et al., 2015 (31).**

| *PIK3CA* Variant | Count (COSMIC, all cancers) | Relative Frequency (total *PIK3CA* variants = 13969) | IL3-independent Ba/F3 growth fold change (mean) [Fig.2b (29)] | EGF/Insulin-independent MCF10A growth fold change (mean) [Fig.2d (29)] |
|---|---|---|---|---|
| H1047R | 4266 | 30.54 | 48.4 | 16.5 |
| E545K | 3174 | 22.72 | 52.3 | 10.8 |
| E542K | 1953 | 13.98 | 51.2 | 6.1 |
| H1047L | 583 | 4.17 | 46.0 | 5.9 |
| E545A | 269 | 1.93 | 20.9 | 5.3 |
| Q546K | 241 | 1.73 | 30.1 | 7.1 |
| E545G | 204 | 1.46 | 21.0 | 5.7 |
| N345K | 202 | 1.45 | 36.2 | 4.2 |
| C420R | 168 | 1.20 | 32.7 | 5.3 |
| H1047Y | 119 | 0.85 | 5.5 | 2.3 |
| M1043I | 106 | 0.76 | 2.7 | 1.6 |
| G1049R | 102 | 0.73 | 20.8 | 9.0 |
| Q546R | 102 | 0.73 | 24.1 | 5.5 |
| E545Q | 73 | 0.52 | 6.1 | 2.8 |
| M1043V | 62 | 0.44 | 5.7 | 3.1 |
| E453K | 61 | 0.44 | 8.4 | 3.4 |
| Q546P | 46 | 0.33 | 23.0 | 5.8 |
| N1044K | 41 | 0.29 | 8.0 | 3.1 |
| K111N | 39 | 0.28 | 2.4 | 1.1 |
| P539R | 32 | 0.23 | 6.1 | 2.6 |
| Q546E | 24 | 0.17 | 12.6 | 3.2 |
| T1025T | 22 | 0.16 | 1.4 | 1.6 |
| E542V | 18 | 0.13 | 7.3 | 3.6 |
| N345I | 15 | 0.11 | 3.9 | 2.6 |
| H701P | 8 | 0.06 | 1.1 | 0.5 |
| A1020V | 4 | 0.03 | 0.9 | 1.3 |
| I31M | 2 | 0.01 | 1.2 | 1.4 |
| Pearson correlation with COSMIC frequency | | | r = 0.72<br>*P* = 2.45E-5 | r = 0.82<br>*P* = 1.33E-7 |