**Supplementary Table 5.** Continued

| Reference position | Gene | Chromosome | Coding sequence | Coverage | Allele change | Patient no. |
|---|---|---|---|---|---|---|
| 889026155[a] | CARD6 | 5 | 3 | 32 | insT | 2 |
| 891727995 | PAIP1 | 5 | 2 | 16 | insG | 2 |
| 901348948 | MAP3K1 | 5 | 13 | 24 | insC | 3 |
| 909928153 | ADAMTS6 | 5 | 3 | 23 | insG | 3 |
| 914509435 | SERF1B | 5 | 3 | 15 | insG | 3 |
| 914509482 | SERF1B | 5 | 3 | 49 | insA | 3 |
| 915509411[a] | GTF2H2 | 5 | 13 | 31 | insT | 2 |
| 919186632 | HEXB | 5 | 11 | 43 | insA | 3 |
| 922917852 | SCAMP1 | 5 | 7 | 37 | insA | 1 |
| 928534269 | EDIL3 | 5 | 7 | 25 | insA | 3 |
| 931867232 | CCNH | 5 | 7 | 26 | insT | 3 |
| 956697390 | EPB41L4A | 5 | 11 | 13 | insT | 1 |
| 966610635 | ZNF474 | 5 | 1 | 39 | delT | 1 |
| 972596345 | SLC12A2 | 5 | 8 | 35 | insT | 3 |
| 980635083[a] | SMAD5 | 5 | 6 | 105 | insC | 1 |
| 985314450[a] | LOC100288105 | 5 | 1 | 14 | delC | 4 |
| 985640033[a] | PCDHB9 | 5 | 1 | 32 | insA | 1 |
| 985844899[a] | PCDHGA8 | 5 | 1 | 27 | delC | 3 |
| 992330379 | SCGB3A2 | 5 | 1 | 14 | delA | 3 |
| 994446878[a] | TIGD6 | 5 | 1 | 136 | delT | 1 |
| 994476149 | HMGXB3 | 5 | 6 | 14 | delA | 3 |
| 998157358 | GRIA1 | 5 | 11 | 19 | insC | 3 |
| 1020539380 | FAM153B | 5 | 4 | 23 | insC | 3 |
| 1039337531 | C6orf114 | 6 | 1 | 30 | insA | 3 |
| 1052252527 | BTN2A2 | 6 | 2 | 44 | insG | 3 |
| 1054107191[a] | ZNF187 | 6 | 1 | 33 | insG | 4 |
| 1056096293[a] | FLJ45422 | 6 | 2 | 18 | insT | 2 |
| 1057247419[a] | MICA | 6 | 5 | 27 | delG | 3 |
| 1082305754 | DST | 6 | 45 | 18 | insT | 3 |
| 1088830738 | EYS | 6 | 6 | 19 | insT | 3 |
| 1093406718 | COL19A1 | 6 | 5 | 16 | insA | 3 |
| 1113248546 | MDN1 | 6 | 15 | 14 | insC | 3 |
| 1113280524 | MDN1 | 6 | 2 | 40 | insA | 3 |
| 1131602437[a] | FOXO3 | 6 | 2 | 64 | insG | 3 |
| 1133380782 | SLC22A16 | 6 | 4 | 20 | insA | 3 |
| 1135037748 | C6orf225 | 6 | 1 | 17 | delC | 3 |
| 1153093327 | SAMD3 | 6 | 7 | 14 | delC | 3 |
| 1154647636 | LOC643854 | 6 | 1 | 26 | insT | 3 |
| 1154648098 | LOC643854 | 6 | 1 | 20 | insC | 3 |
| 1159216432 | BCLAF1 | 6 | 2 | 13 | delT | 2 |
| 1161156444 | PBOV1 | 6 | 1 | 36 | insG | 3 |
| 1182019086 | RSPH3 | 6 | 6 | 43 | insA | 3 |
| 1200500350 | RSPH10B2 | 7 | 19 | 23 | insG | 3 |
| 1206053594 | VWDE | 7 | 19 | 26 | insA | 4 |
| 1221518316 | TAX1BP1 | 7 | 13 | 14 | insA | 3 |
| 1222659823[a] | KIAA0644 | 7 | 1 | 90 | delC | 4 |
| 1222659922[a] | KIAA0644 | 7 | 1 | 26 | insC | 3 |
| 1226974977 | BBS9 | 7 | 7 | 19 | insT | 1 |
| 1228643735[a] | DPY19L1 | 7 | 18 | 22 | delAT | 4 |
| 1228643736[a] | DPY19L1 | 7 | 18 | 50 | delT | 1 |
| 1262731853 | TYW1B | 7 | 8 | 142 | delA | 4 |
| 1262954278 | TRIM74 | 7 | 2 | 24 | insA | 2 |
| 1265555276 | TRIM73 | 7 | 2 | 84 | insT | 2 |
| 1266437261[a] | FLJ37078 | 7 | 14 | 43 | insC | 2 |
| 1266593512[a] | ZP3 | 7 | 8 | 51 | insG | 1 |
| 1266763110[a] | POMZP3 | 7 | 5 | 83 | delA | 4 |
| 1278946055 | C7orf62 | 7 | 1 | 20 | insC | 3 |
| 1283360469 | HEPACAM2 | 7 | 4 | 29 | insT | 3 |
| 1283589759 | CALCR | 7 | 9 | 26 | insT | 3 |
| 1290893801[a] | ZAN | 7 | 30 | 28 | insG | 3 |
| 1291366094 | MOGAT3 | 7 | 2 | 24 | insA | 3 |

**Supplementary Table 5.** Continued

| Reference position | Gene | Chromosome | Coding sequence | Coverage | Allele change | Patient no. |
|---|---|---|---|---|---|---|
| 1291722996[a] | EMID2 | 7 | 13 | 20 | insG | 4 |
| 1292538685 | LOC100289561 | 7 | 1 | 14 | insA | 3 |
| 1295252926 | MLL5 | 7 | 12 | 21 | insG | 3 |
| 1298402792 | NRCAM | 7 | 1 | 17 | insT | 3 |
| 1319055841[a] | KCP | 7 | 10 | 62 | insC | 1 |
| 1319073009 | KCP | 7 | 1 | 30 | delC | 2 |
| 1333766379 | LOC441294 | 7 | 1 | 46 | insA | 4 |
| 1334380185 | CTAGE4 | 7 | 1 | 39 | insA | 3 |
| 1334381975 | ARHGEF5L | 7 | 1 | 19 | insA | 1 |
| 1339923632[a] | KRBA1 | 7 | 12 | 76 | insC | 2 |
| 1339973995[a] | SSPO | 7 | 9 | 44 | insC | 1 |
| 1340003537[a] | SSPO | 7 | 60 | 15 | insC | 4 |
| 1340012514 | SSPO | 7 | 76 | 23 | delA | 2 |
| 1340015859 | SSPO | 7 | 83 | 14 | delC | 2 |
| 1340525483 | C7orf29 | 7 | 1 | 24 | delC | 1 |
| 1341211228[a] | ATG9B | 7 | 10 | 49 | insC | 1 |
| 1341434558 | SMARCD3 | 7 | 10 | 21 | delC | 3 |
| 1342197228 | GALNTL5 | 7 | 5 | 71 | delT | 4 |
| 1342442397[a] | MLL3 | 7 | 14 | 208 | insT | 4 |
| 1356372261[a] | XKR5 | 8 | 6 | 55 | delAG | 1 |
| 1374409954[a] | NEFL | 8 | 3 | 38 | delG | 4 |
| 1380219728[a] | UBXN8 | 8 | 7 | 83 | insT | 1 |
| 1380304215 | TEX15 | 8 | 1 | 23 | insA | 3 |
| 1388426070[a] | PLEKHA2 | 8 | 11 | 28 | delC | 2 |
| 1395399601[a] | PRKDC | 8 | 31 | 17 | insG | 1 |
| 1398930064 | PXDNL | 8 | 14 | 27 | insA | 3 |
| 1410692513[a] | YTHDF3 | 8 | 4 | 24 | insG | 1 |
| 1415952398 | C8orf34 | 8 | 2 | 32 | insG | 3 |
| 1445261384 | LAPTM4B | 8 | 2 | 16 | insC | 3 |
| 1490189877[a] | JRK | 8 | 1 | 12 | delCA | 3 |
| 1490189878[a] | JRK | 8 | 1 | 19 | delA | 2 |
| 1491176363 | ZNF623 | 8 | 1 | 29 | insT | 3 |
| 1492082552[a] | RECQL4 | 8 | 14 | 20 | delG | 3 |
| 1498992866 | LOC645969 | 9 | 1 | 155 | insT | 4 |
| 1527437913[a] | C9orf144B | 9 | 4 | 20 | delC | 4 |
| 1543290663[a] | FOXD4L5 | 9 | 1 | 39 | delG | 1 |
| 1546032104 | TRPM3 | 9 | 22 | 19 | insT | 3 |
| 1552818643 | VPS13A | 9 | 48 | 29 | insG | 3 |
| 1574648314 | COL15A1 | 9 | 13 | 29 | insC | 3 |
| 1586295095 | MUSK | 9 | 1 | 62 | insT | 3 |
| 1608846803[a] | ABO | 9 | 6 | 117 | insC | 4 |
| 1620006324 | GDI2 | 10 | 7 | 14 | insG | 3 |
| 1620254092 | IL2RA | 10 | 4 | 14 | insC | 3 |
| 1621795546[a] | ITIH5 | 10 | 14 | 23 | delC | 1 |
| 1633127998 | NSUN6 | 10 | 2 | 26 | insA | 3 |
| 1647389817 | ITGB1 | 10 | 13 | 22 | insA | 3 |
| 1652560241[a] | LOC340947 | 10 | 2 | 25 | delT | 1 |
| 1653671683[a] | LOC642424 | 10 | 3 | 117 | delT | 1 |
| 1657313101 | AGAP4 | 10 | 7 | 23 | delT | 2 |
| 1658942495 | FAM25G | 10 | 3 | 48 | insC | 3 |
| 1662197526 | LOC100287932 | 10 | 6 | 22 | insA | 4 |
| 1662338998[a] | AGAP6 | 10 | 1 | 50 | insC | 2 |
| 1666373362 | PCDH15 | 10 | 19 | 56 | insC | 3 |
| 1673760921 | TMEM26 | 10 | 6 | 25 | insT | 3 |
| 1685560504 | FAM149B1 | 10 | 7 | 26 | insT | 3 |
| 1701949711 | PANK1 | 10 | 3 | 30 | insA | 3 |
| 1708407108 | CCNJ | 10 | 3 | 18 | insC | 3 |
| 1708510568 | ZNF518A | 10 | 1 | 30 | insC | 1 |
| 1708668598 | DNTT | 10 | 2 | 18 | insA | 3 |
| 1709332414 | C10orf12 | 10 | 1 | 18 | insG | 3 |
| 1728973932[a] | PNLIPRP2 | 10 | 3 | 52 | insG | 1 |

**Supplementary Table 5.** Continued

| Reference position | Gene | Chromosome | Coding sequence | Coverage | Allele change | Patient no. |
|---|---|---|---|---|---|---|
| 1733216553 | BRWD2 | 10 | 8 | 28 | insT | 3 |
| 1737221823 | ZRANB1 | 10 | 1 | 17 | insA | 3 |
| 1738045786 | MMP21 | 10 | 7 | 33 | insG | 3 |
| 1748226082 | C11orf21 | 11 | 4 | 63 | insG | 1 |
| 1750053615 | RRM1 | 11 | 14 | 17 | insT | 3 |
| 1753342600 | SYT9 | 11 | 4 | 18 | insG | 3 |
| 1760006709[a] | SPON1 | 11 | 5 | 72 | insC | 4 |
| 1764016157 | SAAL1 | 11 | 7 | 40 | delT | 4 |
| 1771005317 | LUZP2 | 11 | 12 | 34 | insA | 3 |
| 1782417168 | TRAF6 | 11 | 6 | 24 | delG | 3 |
| 1782519665 | RAG2 | 11 | 1 | 21 | insA | 3 |
| 1792247476[a] | CREB3L1 | 11 | 12 | 40 | insG | 1 |
| 1802120506 | TCN1 | 11 | 7 | 13 | insA | 3 |
| 1802663567[a] | MS4A14 | 11 | 2 | 61 | delTT | 4 |
| 1802663568[a] | MS4A14 | 11 | 2 | 22 | delT | 3 |
| 1803663946[a] | TMEM216 | 11 | 3 | 54 | insA | 4 |
| 1804797590 | AHNAK | 11 | 3 | 12 | insG | 3 |
| 1805556025 | SLC22A10 | 11 | 1 | 17 | insC | 3 |
| 1810263379[a] | UNC93B1 | 11 | 7 | 53 | insG | 3 |
| 1810284280[a] | ALDH3B1 | 11 | 2 | 63 | insC | 2 |
| 1810287509[a] | ALDH3B1 | 11 | 6 | 18 | insC | 1 |
| 1810293595[a] | ALDH3B1 | 11 | 9 | 28 | insC | 4 |
| 1814065554 | LOC729523 | 11 | 1 | 22 | delT | 3 |
| 1826743977 | DLG2 | 11 | 5 | 23 | insT | 3 |
| 1832107207 | LOC642446 | 11 | 1 | 33 | delT | 4 |
| 1837197723[a] | CWC15 | 11 | 5 | 152 | insT | 1 |
| 1837299118[a] | SFRS2B | 11 | 1 | 36 | insC | 4 |
| 1850549218 | ATM | 11 | 49 | 24 | insT | 3 |
| 1852355678 | ZC3H12C | 11 | 2 | 25 | insC | 3 |
| 1854201323[a] | DIXDC1 | 11 | 7 | 16 | insC | 1 |
| 1860877259[a] | TREH | 11 | 15 | 28 | insG | 2 |
| 1861246651[a] | SLC37A4 | 11 | 3 | 37 | delC | 1 |
| 1861288156 | VPS11 | 11 | 2 | 13 | insC | 4 |
| 1867800518 | EI24 | 11 | 9 | 14 | insC | 4 |
| 1867851321 | CHEK1 | 11 | 5 | 44 | insC | 3 |
| 1888645169[a] | PRB3 | 12 | 4 | 34 | delG | 4 |
| 1888731023[a] | PRB1 | 12 | 3 | 136 | delC | 1 |
| 1891856090 | ATF7IP | 12 | 11 | 19 | insG | 3 |
| 1893735417[a] | MGST1 | 12 | 2 | 12 | delAA | 3 |
| 1893735418[a] | MGST1 | 12 | 2 | 18 | delA | 3 |
| 1898574937 | SLCO1B1 | 12 | 7 | 17 | insC | 3 |
| 1902256413 | BCAT1 | 12 | 5 | 22 | insG | 3 |
| 1913975525 | KIF21A | 12 | 10 | 20 | insT | 3 |
| 1914378775 | SLC2A13 | 12 | 10 | 17 | insA | 3 |
| 1927092176 | KRT6C | 12 | 1 | 15 | insG | 2 |
| 1930622534 | SUOX | 12 | 3 | 14 | insG | 3 |
| 1931678522 | TMEM194A | 12 | 9 | 23 | insG | 3 |
| 1932337710 | OS9 | 12 | 12 | 17 | insA | 3 |
| 1959863488 | LRRIQ1 | 12 | 26 | 12 | delA | 3 |
| 1962616654 | C12orf50 | 12 | 3 | 28 | insA | 3 |
| 1978598568[a] | TDG | 12 | 3 | 14 | insA | 3 |
| 1986789153 | LOC100287839 | 12 | 9 | 35 | insC | 3 |
| 1997115077 | RSRC2 | 12 | 10 | 28 | insG | 3 |
| 1999523126 | UBC | 12 | 1 | 29 | delT | 3 |
| 2009256491 | ZMYM5 | 13 | 5 | 14 | insC | 3 |
| 2012756904 | SACS | 13 | 9 | 20 | insT | 3 |
| 2012761230 | SACS | 13 | 9 | 23 | insT | 3 |
| 2017859185 | FLT1 | 13 | 4 | 36 | insA | 3 |
| 2022550582 | STARD13 | 13 | 5 | 85 | delT | 1 |
| 2026525487 | CSNK1A1L | 13 | 1 | 13 | insC | 2 |
| 2038965626 | RCBTB1 | 13 | 8 | 17 | insG | 3 |

**Supplementary Table 5.** Continued

| Reference position | Gene | Chromosome | Coding sequence | Coverage | Allele change | Patient no. |
|---|---|---|---|---|---|---|
| 2046563396 | PRR20 | 13 | 2 | 28 | delC | 2 |
| 2063234131 | KLF12 | 13 | 4 | 47 | insT | 3 |
| 2066482633 | MYCBP2 | 13 | 75 | 17 | insT | 3 |
| 2066508358 | MYCBP2 | 13 | 62 | 14 | insC | 3 |
| 2066632819 | MYCBP2 | 13 | 22 | 23 | delC | 3 |
| 2066717508 | MYCBP2 | 13 | 2 | 33 | delAA | 3 |
| 2066717509 | MYCBP2 | 13 | 2 | 33 | delA | 3 |
| 2088603872 | GPR18 | 13 | 1 | 20 | insT | 3 |
| 2105948032 | NDRG2 | 14 | 1 | 34 | delG | 1 |
| 2106009532 | FLJ10357 | 14 | 18 | 14 | delG | 3 |
| 2108927297 | DHRS4L2 | 14 | 6 | 41 | insA | 4 |
| 2109139875[a] | MDP-1 | 14 | 6 | 13 | delA | 1 |
| 2117359342 | AKAP6 | 14 | 1 | 20 | insA | 3 |
| 2117747539 | AKAP6 | 14 | 12 | 21 | insA | 3 |
| 2137979011 | DDHD1 | 14 | 10 | 22 | insC | 3 |
| 2148241015[a] | GPHB5 | 14 | 1 | 18 | insG | 4 |
| 2158414589[a] | C14orf169 | 14 | 1 | 19 | insC | 3 |
| 2159993929[a] | FAM164C | 14 | 1 | 14 | insA | 1 |
| 2160606560 | TTLL5 | 14 | 4 | 17 | insA | 3 |
| 2179419547 | SERPINA12 | 14 | 2 | 54 | insC | 3 |
| 2179491154 | SERPINA4 | 14 | 3 | 14 | insG | 3 |
| 2181450460 | PAPOLA | 14 | 5 | 33 | insC | 3 |
| 2202211427[a] | CHRFAM7A | 15 | 4 | 191 | delCA | 1 |
| 2202211428[a] | CHRFAM7A | 15 | 4 | 252 | delA | 4 |
| 2203996021[a] | CHRNA7 | 15 | 6 | 166 | delTG | 1 |
| 2203996022[a] | CHRNA7 | 15 | 6 | 50 | delG | 2 |
| 2204534873 | SCG5 | 15 | 5 | 24 | insC | 3 |
| 2212460825 | CASC5 | 15 | 10 | 14 | insA | 3 |
| 2220067652 | SLC12A1 | 15 | 5 | 21 | insA | 3 |
| 2237036677 | LOC100287371 | 15 | 3 | 32 | insG | 3 |
| 2243652079[a] | NR2E3 | 15 | 6 | 34 | delC | 1 |
| 2251295853 | KIAA1024 | 15 | 1 | 14 | insT | 3 |
| 2252413491 | ARNT2 | 15 | 14 | 24 | insC | 3 |
| 2256610001 | ZSCAN2 | 15 | 2 | 14 | insC | 3 |
| 2257065252 | PDE8A | 15 | 4 | 20 | delT | 3 |
| 2261248094 | FANCI | 15 | 2 | 19 | insC | 3 |
| 2261584966 | C15orf42 | 15 | 7 | 21 | insT | 3 |
| 2270957952[a] | LOC145814 | 15 | 4 | 23 | insC | 4 |
| 2271092254[a] | SYNM | 15 | 1 | 19 | insG | 3 |
| 2274046312[a] | C16orf35 | 16 | 12 | 89 | insG | 4 |
| 2274304546 | AXIN1 | 16 | 1 | 20 | delC | 3 |
| 2277509768[a] | NLRC3 | 16 | 7 | 81 | delG | 1 |
| 2285935013[a] | LOC729978 | 16 | 4 | 20 | delAT | 4 |
| 2285935014[a] | LOC729978 | 16 | 4 | 44 | delT | 1 |
| 2292434768 | NOMO2 | 16 | 24 | 22 | insG | 3 |
| 2294397443 | ACSM2A | 16 | 9 | 23 | delA | 3 |
| 2294883814 | DNAH3 | 16 | 53 | 18 | insC | 3 |
| 2304906998 | HSD3B7 | 16 | 6 | 48 | delC | 2 |
| 2332868773 | CLEC18C | 16 | 3 | 24 | insA | 3 |
| 2333553555[a] | HYDIN | 16 | 68 | 29 | delA | 4 |
| 2338969142[a] | CNTNAP4 | 16 | 1 | 82 | insT | 1 |
| 2351412465[a] | LOC100289580 | 16 | 2 | 67 | delC | 2 |
| 2354387432 | PRPF8 | 17 | 4 | 11 | insG | 3 |
| 2356396572[a] | P2RX5 | 17 | 3 | 13 | delG | 1 |
| 2359357840 | C17orf100 | 17 | 1 | 14 | insG | 2 |
| 2360272579[a] | SENP3 | 17 | 6 | 20 | delA | 4 |
| 2361527508[a] | PIK3R6 | 17 | 16 | 42 | insG | 1 |
| 2363416732 | C17orf48 | 17 | 3 | 19 | insA | 3 |
| 2371198121 | LGALS9C | 17 | 9 | 16 | insA | 4 |
| 2376394518[a] | SEBOX | 17 | 1 | 29 | insG | 4 |
| 2376430014[a] | SLC46A1 | 17 | 4 | 15 | delA | 1 |

**Supplementary Table 5.** Continued

| Reference position | Gene | Chromosome | Coding sequence | Coverage | Allele change | Patient no. |
|---|---|---|---|---|---|---|
| 2382300534 | CCL7 | 17 | 2 | 24 | insT | 3 |
| 2383802642[a] | MMP28 | 17 | 4 | 28 | insC | 4 |
| 2384283858 | TBC1D3C | 17 | 13 | 31 | insG | 1 |
| 2388631071 | KRT10 | 17 | 1 | 14 | delC | 3 |
| 2392844842[a] | PLCD3 | 17 | 10 | 24 | insC | 1 |
| 2393016586[a] | MAP3K14 | 17 | 4 | 16 | insG | 2 |
| 2407377091 | CLTC | 17 | 3 | 28 | insT | 3 |
| 2409792732 | MED13 | 17 | 2 | 24 | insA | 3 |
| 2411313186[a] | WDR68 | 17 | 5 | 42 | delG | 1 |
| 2412151377 | DDX5 | 17 | 8 | 23 | insT | 3 |
| 2434290839 | MYOM1 | 18 | 8 | 16 | insA | 3 |
| 2448603454 | RBBP8 | 18 | 14 | 22 | insC | 3 |
| 2451555232 | LOC100287386 | 18 | 2 | 31 | insA | 1 |
| 2471234235 | SLC14A2 | 18 | 4 | 39 | delC | 3 |
| 2492044315 | CDH19 | 18 | 11 | 24 | insA | 3 |
| 2501962862 | ZNF516 | 18 | 2 | 27 | delG | 2 |
| 2508173565[a] | SPPL2B | 19 | 7 | 26 | insC | 2 |
| 2510788089 | UHRF1 | 19 | 14 | 13 | insC | 3 |
| 2514792389 | MUC16 | 19 | 3 | 18 | insA | 3 |
| 2514803399 | MUC16 | 19 | 3 | 29 | insT | 3 |
| 2518236406 | ZNF799 | 19 | 4 | 25 | insA | 3 |
| 2521463907[a] | CYP4F8 | 19 | 4 | 79 | insC | 1 |
| 2522001621[a] | HSH2D | 19 | 5 | 71 | delA | 2 |
| 2538892348 | C19orf55 | 19 | 9 | 17 | delG | 2 |
| 2543188059 | ZNF780B | 19 | 2 | 24 | insC | 3 |
| 2543756504[a] | LTBP4 | 19 | 24 | 14 | insG | 1 |
| 2544255517[a] | CYP2F1 | 19 | 1 | 53 | insC | 4 |
| 2544853028 | CEACAM5 | 19 | 4 | 26 | insT | 3 |
| 2547650400[a] | CEACAM20 | 19 | 8 | 54 | delT | 1 |
| 2547930257[a] | CBLC | 19 | 8 | 18 | insC | 3 |
| 2552076265[a] | DHDH | 19 | 4 | 55 | insG | 2 |
| 2552600822 | ALDH16A1 | 19 | 10 | 73 | insC | 2 |
| 2554469302[a] | LOC147645 | 19 | 10 | 37 | insG | 4 |
| 2555437083[a] | ZNF480 | 19 | 1 | 51 | delG | 1 |
| 2555750854 | ZNF83 | 19 | 1 | 26 | insG | 3 |
| 2559350849 | ZSCAN5C | 19 | 1 | 50 | insA | 2 |
| 2560590155 | ZNF749 | 19 | 3 | 34 | insA | 1 |
| 2560866399 | ZNF671 | 19 | 4 | 14 | insA | 3 |
| 2561351770[a] | ZNF274 | 19 | 4 | 78 | insG | 2 |
| 2562070563 | DEFB126 | 20 | 2 | 20 | delCC | 3 |
| 2562070564 | DEFB126 | 20 | 2 | 20 | delC | 3 |
| 2567847490[a] | CHGB | 20 | 4 | 28 | delGA | 2 |
| 2567847491[a] | CHGB | 20 | 4 | 70 | delA | 2 |
| 2580083985 | CSRP2BP | 20 | 4 | 17 | insG | 3 |
| 2583130413[a] | NCRNA00153 | 20 | 7 | 49 | insG | 1 |
| 2606534089 | DDX27 | 20 | 4 | 21 | insA | 3 |
| 2608270909 | MOCS3 | 20 | 1 | 23 | insG | 3 |
| 2640900547[a] | KRTAP7-1 | 21 | 1 | 27 | delA | 4 |
| 2643647262[a] | SON | 21 | 12 | 40 | insA | 1 |
| 2643647273[a] | SON | 21 | 12 | 33 | delA | 4 |
| 2654166830 | TRAPPC10 | 21 | 21 | 24 | insT | 3 |
| 2656193953[a] | LOC100288508 | 21 | 5 | 14 | insC | 1 |
| 2670991039 | HORMAD2 | 22 | 2 | 23 | insG | 3 |
| 2681753989 | DNAJB7 | 22 | 1 | 66 | insA | 1 |
| 2683020374 | CYP2D6 | 22 | 5 | 18 | insG | 4 |
| 2701397266 | WWC3 | X | 7 | 15 | insG | 3 |
| 2708217926 | RBBP7 | X | 2 | 16 | insC | 3 |
| 2709924320 | CDKL5 | X | 4 | 26 | insC | 3 |
| 2711314268 | CXorf23 | X | 3 | 22 | delG | 3 |
| 2713575280 | PHEX | X | 19 | 29 | insG | 3 |
| 2736210225 | KDM6A | X | 17 | 17 | insC | 3 |

**Supplementary Table 5.** Continued

| Reference position | Gene | Chromosome | Coding sequence | Coverage | Allele change | Patient no. |
|---|---|---|---|---|---|---|
| 2736225869 | KDM6A | X | 24 | 16 | insA | 3 |
| 2737802282 | SLC9A7 | X | 7 | 21 | insC | 3 |
| 2739406485 | SSX1 | X | 6 | 45 | insT | 3 |
| 2739444455[a] | SSX9 | X | 2 | 11 | delC | 2 |
| 2741301867[a] | DGKK | X | 22 | 55 | insG | 1 |
| 2743920170 | SSX2B | X | 6 | 35 | insC | 3 |
| 2745406637 | WNK3 | X | 16 | 40 | insA | 3 |
| 2755460714 | OPHN1 | X | 8 | 18 | insC | 3 |
| 2757668459 | KIF4A | X | 28 | 24 | insG | 3 |
| 2758547967 | NONO | X | 6 | 22 | insT | 3 |
| 2761842615 | RLIM | X | 3 | 18 | insG | 3 |
| 2771580011 | HDX | X | 5 | 18 | insA | 3 |
| 2779112744 | PCDH11X | X | 2 | 18 | insT | 3 |
| 2788398590 | CENPI | X | 20 | 19 | insC | 3 |
| 2789376503 | TCEAL6 | X | 1 | 24 | insG | 2 |
| 2789554171 | NXF2 | X | 7 | 14 | insT | 3 |
| 2789554906 | NXF2 | X | 10 | 29 | insA | 3 |
| 2802179110 | IL13RA2 | X | 4 | 18 | insT | 3 |
| 2823639589 | ARHGEF6 | X | 18 | 16 | insA | 3 |
| 2840930428[a] | LCAP | X | 1 | 57 | insC | 2 |
| 2841794077 | MPP1 | X | 7 | 14 | insG | 3 |

[a]These indels commonly occurred in more than one HCC.

**Supplementary Table 6.** List of 81 Nucleotide Positions in 77 Genes With Indels at a Frequency of >20% of Reads in 4 Nontumorous Tissues From 4 Patients

| Reference position | Gene | Chromosome | Coding sequence | Coverage | Allele change | Patient no. |
|---|---|---|---|---|---|---|
| 36247083 | THRAP3 | 1 | 4 | 18 | delG | 1 |
| 75174421 | SLC44A5 | 1 | 16 | 19 | delT | 1 |
| 114430296 | TRIM33 | 1 | 20 | 23 | delC | 1 |
| 133132499 | YY1AP1 | 1 | 7 | 37 | insT | 1 |
| 133844355 | RHBG | 1 | 9 | 48 | delC | 4 |
| 201121914 | CAPN2 | 1 | 3 | 14 | delC | 1 |
| 201173637 | TP53BP2 | 1 | 13 | 22 | delG | 1 |
| 247664301 | C2orf43 | 2 | 4 | 18 | delA | 2 |
| 319394229 | SNRNP200 | 2 | 37 | 20 | delA | 1 |
| 322653290 | AFF3 | 2 | 14 | 35 | delA | 1 |
| 331834049 | RANBP2 | 2 | 20 | 16 | delG | 1 |
| 332901065 | RGPD5 | 2 | 20 | 22 | delT | 1 |
| 374789589 | NEB | 2 | 4 | 37 | insT | 4 |
| 382950536 | LY75 | 2 | 5 | 19 | delA | 4 |
| 401635744 | TTN | 2 | 274 | 25 | delA | 1 |
| 409835043 | FAM171B | 2 | 8 | 17 | delT | 1 |
| 412064501 | COL3A1 | 2 | 14 | 21 | insA | 4 |
| 454784716 | PTMA | 2 | 4 | 14 | delT | 1 |
| 463724350[a] | AQP12B | 2 | 1 | 27 | delC | 3 |
| 463734336 | AQP12A | 2 | 2 | 14 | delG | 2 |
| 503335742 | DLEC1 | 3 | 4 | 20 | delT | 1 |
| 503335743 | DLEC1 | 3 | 4 | 20 | delA | 1 |
| 735406533 | CNOT6L | 4 | 10 | 21 | delG | 1 |
| 785877214 | LARP2 | 4 | 14 | 23 | delA | 1 |
| 798021293 | SCOC | 4 | 1 | 18 | insC | 1 |
| 810971969 | TRIM2 | 4 | 5 | 18 | delC | 1 |
| 883256725 | PRLR | 5 | 3 | 29 | delG | 4 |
| 939146286 | ANKRD32 | 5 | 16 | 15 | insC | 1 |
| 985314450[a] | LOC100288105 | 5 | 1 | 27 | delC | 3 |
| 1033746568 | BMP6 | 6 | 5 | 24 | delC | 1 |
| 1068664364 | KIAA0240 | 6 | 4 | 17 | insT | 4 |
| 1193244025 | FAM120B | 6 | 1 | 43 | insA | 1 |
| 1222659823 | KIAA0644 | 7 | 1 | 463 | delC | 4 |
| 1282877394 | CDK6 | 7 | 3 | 21 | delA | 1 |
| 1289880816 | CYP3A4 | 7 | 12 | 45 | delG | 1 |
| 1333766765 | LOC441294 | 7 | 1 | 13 | delA | 4 |
| 1340012514 | SSPO | 7 | 76 | 53 | delA | 3 |
| 1356372261 | XKR5 | 8 | 6 | 130 | delA | 4 |
| 1490189877[a] | JRK | 8 | 1 | 29 | delC | 3 |
| 1490189878[a] | JRK | 8 | 1 | 15 | delA | 2 |
| 1492082552[a] | RECQL4 | 8 | 14 | 43 | delG | 3 |
| 1505961686 | MPDZ | 9 | 2 | 28 | insG | 4 |
| 1526015264 | NFX1 | 9 | 3 | 36 | delT | 1 |
| 1573925487 | GABBR2 | 9 | 17 | 59 | insT | 4 |
| 1580509237 | ABCA1 | 9 | 4 | 24 | insT | 1 |
| 1637516682 | ARMC3 | 10 | 18 | 19 | delT | 1 |
| 1637516683 | ARMC3 | 10 | 18 | 19 | delT | 1 |
| 1657313100[a] | AGAP4 | 10 | 7 | 19 | delT | 2 |
| 1657313101[a] | AGAP4 | 10 | 7 | 14 | delT | 3 |
| 1807397040 | SYVN1 | 11 | 7 | 15 | insA | 1 |
| 1832107207 | LOC642446 | 11 | 1 | 18 | delT | 4 |
| 1855967268 | ZW10 | 11 | 8 | 40 | delC | 1 |
| 1861246651 | SLC37A4 | 11 | 3 | 143 | delC | 4 |
| 1884320486 | ATN1 | 12 | 4 | 15 | delA | 1 |
| 1929584670 | KIAA0748 | 12 | 6 | 31 | delC | 4 |
| 1955959709 | PPFIA2 | 12 | 18 | 52 | delA | 4 |
| 1994379801 | CIT | 12 | 17 | 28 | delG | 1 |
| 1995110709 | DYNLL1 | 12 | 2 | 14 | delG | 1 |
| 1997144069 | KNTC1 | 12 | 2 | 31 | delC | 4 |
| 2105624179 | RNASE4 | 14 | 1 | 15 | delC | 1 |

**Supplementary Table 6.** Continued

| Reference position | Gene | Chromosome | Coding sequence | Coverage | Allele change | Patient no. |
|---|---|---|---|---|---|---|
| 2162358340 | C14orf133 | 14 | 13 | 15 | delT | 1 |
| 2243652079 | NR2E3 | 15 | 6 | 129 | delC | 4 |
| 2256057594 | ADAMTSL3 | 15 | 12 | 26 | delT | 1 |
| 2277509768[a] | NLRC3 | 16 | 7 | 12 | delG | 2 |
| 2302633200 | EIF3C | 16 | 4 | 18 | delG | 4 |
| 2303380808 | SULT1A4 | 16 | 3 | 24 | delA | 1 |
| 2351412465 | LOC100289580 | 16 | 2 | 103 | delC | 4 |
| 2356396572 | P2RX5 | 17 | 3 | 40 | delG | 4 |
| 2376621991 | SPAG5 | 17 | 3 | 13 | delC | 1 |
| 2386619109 | CCDC49 | 17 | 5 | 14 | delT | 1 |
| 2413869089 | APOH | 17 | 5 | 24 | delC | 1 |
| 2501962862 | ZNF516 | 18 | 2 | 29 | delG | 3 |
| 2507200605 | MUM1 | 19 | 8 | 36 | delG | 1 |
| 2538892348 | C19orf55 | 19 | 9 | 20 | delG | 2 |
| 2565046537 | UBOX5 | 20 | 2 | 15 | delG | 1 |
| 2587599923 | ZNF337 | 20 | 4 | 19 | delT | 1 |
| 2598525448 | ZHX3 | 20 | 1 | 19 | delT | 1 |
| 2625038622 | NRIP1 | 21 | 1 | 24 | delG | 1 |
| 2661518554 | FAM108A5 | 22 | 2 | 13 | delG | 3 |
| 2748277559 | SPIN2B | X | 1 | 13 | delG | 2 |
| 2792445004 | TEX13A | X | 2 | 18 | delC | 3 |

[a]These indels commonly occurred in more than one HCC.

**Supplementary Table 7.** List of 40 Somatic Mutations With Amino Acid Changes Commonly Detected in Both the Tumor (at a Frequency of More Than 20% of Reads) and Matched Nontumorous Cirrhotic Liver (at a Frequency of More Than 5% of Reads) of the Same Patient

| | | | | | Tumor | | Nontumor | |
|---|---|---|---|---|---|---|---|---|
| Gene | Reference position | Chromosome | Reference nucleotide | Mutation nucleotide | Mutation frequency (%) | Patient no. | Mutation frequency (%) | Patient no. |
| LEPR | 65548341 | 1 | C | A | 25.8 | 3 | 15.0 | 3 |
| | | | | | | | 21.9 | 1 |
| ZNF408 | 1792629936 | 11 | T | A | 20.4 | 2 | 16.0 | 2 |
| | | | | | | | 15.8 | 4 |
| HRNR | 129676984 | 1 | G | C | 28.9 | 3 | 5.4 | 3 |
| PXDN | 228577682 | 2 | G | C | 45.1 | 4 | 47.2 | 4 |
| POTEF | 353150970 | 2 | T | A | 41.8 | 4 | 31.0 | 4 |
| ALPP | 455451136 | 2 | C | T | 32.5 | 4 | 37.5 | 4 |
| GPR125 | 682521774 | 4 | C | A | 38.1 | 2 | 40.0 | 2 |
| HERC6 | 746068457 | 4 | T | A | 36.5 | 4 | 44.9 | 4 |
| EGFLAM | 886579974 | 5 | T | G | 23.3 | 3 | 5.3 | 3 |
| C4A | 1057829599 | 6 | T | G | 25.0 | 2 | 11.5 | 2 |
| WISP3 | 1134999625 | 6 | T | G | 43.3 | 4 | 64.3 | 4 |
| C7orf10 | 1234451360 | 7 | T | A | 25.0 | 3 | 8.3 | 3 |
| PVRIG | 1290339880 | 7 | C | T | 23.5 | 1 | 21.3 | 1 |
| MUC17 | 1291200140 | 7 | G | A | 21.2 | 4 | 12.5 | 4 |
| PLOD3 | 1291376235 | 7 | G | C | 48.2 | 4 | 51.7 | 4 |
| COL27A1 | 1589933932 | 9 | A | G | 56.8 | 4 | 54.6 | 4 |
| AGAP9 | 1658906463 | 10 | T | G | 36.7 | 4 | 16.2 | 4 |
| POLL | 1713935693 | 10 | G | T | 44.8 | 4 | 38.6 | 4 |
| MUC5AC | 1747183167 | 11 | G | A | 43.9 | 4 | 43.8 | 4 |
| MRGPRX3 | 1764064669 | 11 | T | C | 40.0 | 4 | 42.5 | 4 |
| TMEM133 | 1843211533 | 11 | A | C | 59.5 | 4 | 83.3 | 4 |
| TMEM123 | 1844621025 | 11 | G | A | 27.3 | 2 | 7.3 | 2 |
| TMPRSS4 | 1860336319 | 11 | C | T | 54.4 | 4 | 41.3 | 4 |
| DHRS4L2 | 2108914889 | 14 | G | T | 20.5 | 3 | 11.5 | 3 |
| GOLGA6C | 2247104814 | 15 | A | T | 21.7 | 4 | 9.6 | 4 |
| PRSS22 | 2276813235 | 16 | C | T | 50.0 | 4 | 36.7 | 4 |
| FAM38A | 2351390771 | 16 | C | T | 21.4 | 4 | 54.3 | 4 |
| GGT6 | 2357265990 | 17 | G | A | 92.3 | 4 | 41.7 | 4 |
| COX10 | 2366897810 | 17 | C | T | 55.2 | 4 | 36.3 | 4 |
| KIAA0100 | 2376657621 | 17 | A | C | 47.8 | 4 | 60.0 | 4 |
| TBC1D3B | 2384202011 | 17 | C | T | 63.0 | 4 | 27.4 | 4 |
| TBC1D3D | 2385938140 | 17 | A | G | 45.9 | 4 | 21.0 | 4 |
| ERBB2 | 2387531879 | 17 | A | G | 66.7 | 4 | 54.6 | 4 |
| CSH2 | 2411602334 | 17 | C | T | 90.9 | 4 | 79.5 | 4 |
| QRICH2 | 2423941144 | 17 | T | G | 50.0 | 4 | 60.4 | 4 |
| MOCOS | 2461870479 | 18 | T | C | 72.0 | 4 | 62.2 | 4 |
| CPAMD8 | 2522819358 | 19 | G | A | 21.8 | 3 | 15.2 | 3 |
| MAP4K1 | 2541732174 | 19 | G | A | 36.0 | 4 | 54.3 | 4 |
| PSG8 | 2545901763 | 19 | C | A | 28.3 | 3 | 9.5 | 3 |
| KRTAP12-2 | 2654734983 | 21 | C | T | 59.3 | 4 | 43.5 | 4 |

NOTE. The first 2 genes listed were recurrently mutated in the nontumorous inflamed livers of 2 patients.

**Supplementary Table 8.** Overview of Selected Exome Sequencing Data From 22 Patients With HCV Infection

| | | Aligned reads | Aligned sequence (*base pairs*) | Median read depth |
|---|---|---|---|---|
| *TP53* | Tumor | 29,334 | 2,035,570 | 1476.2 |
| | Nontumor | 31,848 | 2,200,641 | 1575.3 |
| | Lymphocytes | 36,690 | 2,539,944 | 1917.2 |
| *CTNNB1* | Tumor | 90,022 | 6,215,000 | 2344.3 |
| | Nontumor | 75,785 | 5,282,450 | 1991.2 |
| | Lymphocytes | 100,430 | 7,013,325 | 2710.8 |
| *LEPR* | Tumor | 34,328 | 2,390,335 | 538.3 |
| | Nontumor | 60,128 | 4,219,089 | 1025.6 |
| | Lymphocytes | 86,830 | 6,085,511 | 1423.0 |

NOTE. Selected exome sequencing of *TP53*, *CTNNB1*, and *LEPR* was performed for 22 nontumorous cirrhotic liver tissues, 10 HCC tissues, and matched peripheral lymphocytes from each patient. Aligned reads, aligned sequences (base pairs), and median read depth are shown for each sample.

**Supplementary Table 9.** Clinical Features and Overview of Deep Sequencing Data of Patients Who Underwent Deep Sequencing of the *LEPR* Gene

| | Chronic hepatitis (n = 15) | Normal liver (n = 9) |
|---|---|---|
| Age (*y*) | 59.3 | 55.9 |
| Sex (male/female) | 6/9 | 7/2 |
| Aligned reads | 4290 | 3956 |
| Aligned sequence (*base pairs*) | 1,044,737 | 1,275,068 |
| Median read depth | 2838 | 3440 |
| No. of mutations in the *LEPR* gene | 0 | 0 |

NOTE. We determined the sequences of the *LEPR* gene in the liver of 15 noncirrhotic patients with HCV-associated chronic hepatitis. In addition, normal liver tissues were obtained from 9 liver donors at the time of the operation. Age, sex, aligned reads, aligned sequences (base pairs), median read depth, and numbers of mutations are shown.

**Supplementary Table 10.** Mean Body Weights and Serum Levels of Insulin, Triglyceride, Total Cholesterol, and Alanine Aminotransferase of C57BL/KsJ-*db/db* (*db/db*) Mice and Misty (Control) Mice After 4 Weeks of Treatment With TAA

| | *db/db* | Control |
|---|---|---|
| Body weight (*g*) | 46.5 ± 0.6 | 23.5 ± 0.4 |
| Insulin (*ng/mL*) | 30.6 ± 28.3 | 1.6 ± 0.2 |
| Triglyceride (*mg/dL*) | 95.0 ± 5.0 | 50.0 ± 20.0 |
| Total cholesterol (*mg/dL*) | 215.0 ± 15.0 | 95.0 ± 15.0 |
| Alanine aminotransferase (*IU/L*) | 1325.0 ± 1085.0 | 75.0 ± 35.0 |

NOTE. All data are presented as mean ± SD.

**Supplementary Table 11.** Categorization of the Mutated Genes Detected by Whole Exome Sequencing of the AID-Expressing Hepatocyte Cell Line Using the Kyoto Encyclopedia of Genes and Genomes Database

| Pathway | | | | |
|---|---|---|---|---|
| Metabolic pathways | ATP6V0A4 | DMGDH | HSD17B3 | PGD |
| | ATP6V1C2 | GALNT1 | HYAL2 | PHGDH |
| | BCMO1 | GATM | NDST1 | POLR3B |
| | CPS1 | HKDC1 | PAH | |
| PI3K-Akt signaling pathway | BCL2L11 | IBSP NOS3 | PRKCZ | TEK |
| | COL27A1 | | | |
| MAPK signaling pathway | FLNB | SP1 | CACNA1F | PTPN7 |
| Cytokine-cytokine receptor interaction | LEPR | TNFRSF8 | TNFRSF10A | |
| Transcriptional misregulation in cancer | EYA1 | GZMB | JMJD7-PLA2G4B | |
| Proteoglycans in cancer | FLN | ITGB3 | TIMP3 | VTN |
| PPAR signaling pathway | CPT1B | CYP4A22 | PPARD | |
| Cell cycle | E2F2 | ESPL1 | MCM7 | |
| Pathways in cancer | FLT3 | TRAF4 | PDGFA | |
| Hedgehog signaling pathway | GLI3 | LRP2 | CSNK1A1L | |
| Others | 95 genes | | | |

NOTE. The genes categorized in multiple pathways are shown in only one representative pathway. Constitutive AID expression resulted in the accumulation of nucleotide alterations in various genes, including LEPR, of the cultured hepatocyte–derived cells. Whole exome sequencing was performed on DNA derived from established non-neoplastic human primary hepatocyte cells[6] with constitutive AID expression. AID expression in the cultured hepatocytes was performed using a lentiviral system.[5] After 8 weeks of AID expression, the DNA was extracted and subjected to whole exome sequencing as described in Materials and Methods. Overall, a total of 460 nucleotide positions in 380 different genes were defined as mutated in the AID-expressing cultured hepatocytes through the variant filtering process. Among them, pathway analyses by the Kyoto Encyclopedia of Genes and Genomes revealed that many genes, including LEPR, were categorized into well-known signaling pathways: the metabolic pathway, PI3K-Akt signaling pathway, MAPK signaling pathway, cytokine-cytokine receptor interaction pathway, and transcriptional misregulation in cancer pathway. Only categorized genes are shown.

Journal of Clinical Microbiology

# Dynamics of Defective Hepatitis C Virus Clones in Reinfected Liver Grafts in Liver Transplant Recipients: Ultradeep Sequencing Analysis

**Shigeru Ohtsuru, Yoshihide Ueda, Hiroyuki Marusawa, Tadashi Inuzuka, Norihiro Nishijima, Akihiro Nasu, Kazuharu Shimizu, Kaoru Koike, Shinji Uemoto and Tsutomu Chiba**

***J. Clin. Microbiol.*** **2013, 51(11):3645. DOI: 10.1128/JCM.00676-13.**
**Published Ahead of Print 28 August 2013.**

JCM
Journals.ASM.org

# Dynamics of Defective Hepatitis C Virus Clones in Reinfected Liver Grafts in Liver Transplant Recipients: Ultradeep Sequencing Analysis

**Shigeru Ohtsuru,[a,b] Yoshihide Ueda,[a] Hiroyuki Marusawa,[a] Tadashi Inuzuka,[a] Norihiro Nishijima,[a] Akihiro Nasu,[a] Kazuharu Shimizu,[c] Kaoru Koike,[b] Shinji Uemoto,[d] Tsutomu Chiba[a]**

Department of Gastroenterology and Hepatology,[a] Department of Primary Care and Emergency Medicine,[b] and Department of Surgery,[d] Graduate School of Medicine, and Department of Nanobio Drug Discovery, Graduate School of Pharmaceutical Sciences,[c] Kyoto University, Shogoin, Sakyo-ku, Kyoto, Japan

**Hepatitis C virus (HCV) reinfects liver allografts in transplant recipients by replicating immediately after transplantation, causing a rapid increase in blood serum HCV RNA levels. We evaluated dynamic changes in the viral genetic complexity after HCV reinfection of the graft liver; we also identified the characteristics of replicating HCV clones using a massively parallel ultradeep sequencing technique to determine the full-genome HCV sequences in the liver and serum specimens of five transplant recipients with genotype 1b HCV infection before and after liver transplantation. The recipients showed extremely high genetic heterogeneity before transplantation, and the HCV population makeup was not significantly different between the liver and blood serum specimens of the individuals. Viral quasispecies complexity in serum was significantly lower after liver transplantation than before it, suggesting that certain HCV clones selectively proliferated after transplantation. Defective HCV clones lacking the structural region of the HCV genome did not increase in number, and full-genome HCV clones selectively increased in number immediately after liver transplantation. A re-increase in the same defective clone existing before transplantation was detected 22 months after transplantation in one patient. Ultradeep sequencing technology revealed that the genetic heterogeneity of HCV was reduced after liver transplantation. Dynamic changes in defective HCV clones after liver transplantation indicate that these clones have important roles in the HCV life cycle.**

The hepatitis C virus (HCV) has an approximately 9.6-kb plus-strand RNA genome that encodes the viral core, envelope glycoprotein 1 (E1), E2, and p7 structural proteins and the NS2, NS3, NS4A, NS4B, NS5A, and NS5B nonstructural proteins (1). A characteristic of HCV infection is its remarkable genetic diversity with a high degree of genetic heterogeneity in each patient, which is referred to as a quasispecies. In heterogeneous HCV clones, a dominant viral population might evolve as a result of its viral replicative fitness and concurrent immune selection pressures that drive clonal selection.

In HCV-positive liver transplant recipients, HCV reinfection of the liver allograft occurs at the time of transplantation, and replication of HCV begins immediately after transplantation. Blood serum HCV RNA levels then rapidly increase to levels that are 10- to 20-fold higher than pretransplant levels. It is thus hypothesized that specific HCV clones that have growth advantages increase after liver transplantation. Although several studies have attempted to clarify the change in genetic heterogeneity following liver transplantation, the abundant diversity and complexity of HCV have been obstacles to a detailed evaluation of viral genetic heterogeneity. The recent introduction of ultradeep sequencing technology, which is capable of producing millions of DNA sequence reads in a single run, however, is rapidly changing the landscape of genome research (2, 3).

In this study, we performed ultradeep sequencing analyses to unveil the levels of viral quasispecies of genotype 1b HCV in the liver and the serum specimens from 5 patients who underwent living donor liver transplantation (LDLT) and clarified the changes in viral genetic complexity after reinfection of HCV in the graft liver. In the analyses, we found that the population of defective HCV clones that lack structural regions of the HCV genome changed after liver transplantation. We then clarified the dynamics and characteristics of the defective HCV clones.

## MATERIALS AND METHODS

**Patients.** The participants comprised 5 Japanese adult patients with end-stage liver disease with genotype 1b HCV infection who underwent LDLT at Kyoto University Hospital between May 2006 and September 2008. Serum samples were obtained before and 1 month after liver transplantation. In addition, a blood serum sample from a patient in the chronic hepatitis phase 22 months after liver transplantation was obtained and analyzed. Liver tissue samples were obtained from 4 patients (patients 1–4) at the time of transplantation, frozen immediately, and stored at −80°C until use.

Tacrolimus with a steroid or mycophenolate mofetil was administered to induce immunosuppression in the patients. A patient who received an ABO blood type-incompatible transplant was treated with rituximab, plasma exchange, and hepatic artery or portal vein infusion with prostaglandin E1 and methylprednisolone (4).

The ethics committee at Kyoto University approved the studies (protocol no. E1211), and written informed consent for participation in this study was obtained from all patients.

**Virologic assays.** The HCV genotype was determined using a PCR-based genotyping system developed by Ohno et al. (5) to amplify the core region using genotype-specific PCR primers for the determination of the HCV genotypes 1a, 1b, 2a, 2b, 3a, 3b, 4, 5a, and 6a. The blood serum HCV RNA load was evaluated before LDLT 1 month post-LDLT and then every 3 months after LDLT using PCR and an Amplicor HCV assay (Cobas

Received 10 April 2013 Returned for modification 15 June 2013
Accepted 13 August 2013
Published ahead of print 28 August 2013
Address correspondence to Yoshihide Ueda, yueda@kuhp.kyoto-u.ac.jp.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/JCM.00676-13.

doi:10.1128/JCM.00676-13

Amplicor HCV monitor; Roche Molecular Systems, Pleasanton, CA) until April 2008 or a real-time PCR-based quantitation method for HCV (Cobas AmpliPrep/Cobas TaqMan HCV test; Roche Molecular Systems) starting May 2008.

**Direct population Sanger sequencing.** To define the representative reference sequences of full-length HCV in each clinical specimen, serum samples collected before liver transplantation were first subjected to direct population Sanger sequencing using the Applied Biosystems 3500 genetic analyzer (Applied Biosystems, Foster City, CA) (6). Total RNA was extracted from 140 μl of serum using a QIAamp viral RNA minikit (Qiagen, Valencia, CA) and reverse transcribed in a volume of 20 μl with the OneStep RNA PCR kit AMV (TaKaRa Bio, Ohtsu, Japan). The HCV genomes were amplified using Phusion high-fidelity DNA polymerase (Finnzymes, Espoo, Finland). Oligonucleotide primers were designed to amplify the first half (~5,000 bp) and latter half (~4,500 bp) of the genotype 1b HCV genome sequence. PCR products purified by the QIAquick gel extraction kit (Qiagen) were assayed for direct sequencing. The nucleotide sequences of the PCR products were determined using an ABI Prism BigDye Terminator ready reaction kit (Applied Biosystems). A blood serum sample from a healthy volunteer was used as a negative control.

**Massively parallel ultradeep sequencing.** Paired-end sequencing with multiplexed tags was carried out using the Illumina Genome Analyzer II. End repair of DNA fragments, the addition of adenine to the 3′-ends of the DNA fragments, adaptor ligation, and PCR amplification by Illumina-paired end PCR primers were performed as described previously (6, 7). Briefly, the viral genome sequences were amplified with high-fidelity PCR and sheared by nebulization using 32 lb/in$^2$ N2 for 8 min, and the sheared fragments were purified and concentrated using a QIAquick PCR purification kit (Qiagen). The overhangs resulting from the fragmentation were then converted into blunt ends using T4 DNA polymerase and Klenow enzymes, followed by the addition of terminal 3′-adenine residues. One of the adaptors containing six unique base pair (bp) tags, such as ATCACG and CGATGT (multiplexing sample preparation oligonucleotide kit; Illumina), was then ligated to each fragment using DNA ligase. Adaptor-ligated DNAs in the range of 200 to 350 bp were then size selected by agarose gel electrophoresis. These libraries were amplified independently using a minimal PCR amplification step of 18 cycles with Phusion high-fidelity DNA polymerase and then purified using a QIAquick PCR purification kit for a downstream assay. Cluster generation and sequencing were performed for 64 cycles on the Illumina Genome Analyzer II according to the manufacturer's instructions. The obtained images were analyzed and base called using the GA pipeline software version 1.4 with default settings provided by Illumina. Validation of the multiplex ultradeep sequencing of the HCV genome was performed using a plasmid encoding full-length HCV as a template, as reported previously (6). The overall error rates were determined to be, on average, 0.0010 per base pair. We also confirmed that high-fidelity PCR amplification with HCV-specific primer sets followed by multiplex ultradeep sequencing resulted in no significant increase in the error rates of viral sequencing data (ranging from 0.0012 to 0.0013 per bp; per-nucleotide error rate, 0.12% to 0.13%) (6).

**Genome Analyzer sequence data analysis.** Using the high-performance alignment software NextGENe (SoftGenetics, State College, PA), the 64-base tags obtained from the Genome Analyzer II reads were aligned to the reference HCV RNA sequences of ~9,200 bp that were determined by direct population Sanger sequencing in each clinical specimen. Entire reads were removed from the analysis when the median quality value score was <20 and when they contained >3 uncalled nucleotides. Low-quality bases were trimmed from the reads when >3 consecutive bases fell below a quality score of 16. Based on the above criteria, reads were aligned if ≥90% of their bases matched a particular position of the reference sequence. Each position of the viral genome was assigned a coverage depth representing the number of times that nucleotide position was sequenced.

**Detection of defective HCV clones.** The methods for detecting defective HCV clones were reported previously (8). Briefly, reverse transcription-PCR (RT-PCR) was performed using the OneStep RNA PCR kit (TaKaRa) with the extracted RNA from liver and blood serum as a template and two pairs of primers, 5′-CGCCGACCTCATGGGGTACA-3′ and 5′-TGGTGTACATTTGGGTGATT-3′ for the first RT-PCR (HCV-P1) and 5′-TGCTCTTTCTCTATCTTCCT-3′ and 5′-GTGATGATGCA ACCAAGTAG-3′ for the second PCR (HCV-P2). The PCR products were analyzed by electrophoresis in 0.8% agarose gels stained with ethidium bromide. Each purified DNA sample was sequenced at least three times using an ABI Prism BigDye Terminator ready reaction kit (Applied Biosystems). To determine defects in the HCV genome, the sequence of each sample was compared with the registered HCV genome sequence.

**Statistical analysis.** The viral quasispecies nature was evaluated by analyzing the genetic complexity based on the number of different sequences present in the HCV population. The genetic complexity was determined by Shannon entropy index, calculated as follows:

$$S_n = -\frac{\sum_{i=1}^{n} (\ln fi)}{N}$$

where $n$ is the number of different species identified, $fi$ is the observed frequency of the particular variant in the quasispecies, and $N$ is the total number of clones analyzed (9, 10). Statistical comparisons of the complexity between two groups were made using the Wilcoxon rank sum test or the Mann-Whitney U test. $P$ values of <0.05 were considered to be statistically significant.

## RESULTS

**Patient characteristics.** The clinical and virological characteristics of the 5 patients are summarized in Table 1. Four of the 5 recipients were male, and the median age of the patients at the time of LDLT was 52 years (range, 47 to 65 years). All patients had decompensated cirrhosis caused by chronic hepatitis C, and 3 patients had hepatocellular carcinoma before liver transplantation. Right-lobe grafts were used for all patients. All patients had an HCV genotype 1b infection. The median blood serum HCV RNA load before transplantation was 5.5 log IU/ml (range, 4.6 to 6.6 log IU/ml) and was 5.9 log IU/ml (range, 5.8 to 6.4 log IU/ml) 1 month after liver transplantation; however, this difference was not significant ($P = 0.18$).

**HCV population did not significantly differ between liver and serum samples.** To clarify the landscape of HCV heterogeneity as a quasispecies, we determined the viral full-genome sequences in liver and serum samples collected from the 5 recipients before transplantation using multiplex ultradeep sequencing and compared the results with those obtained by the direct population Sanger sequencing method. The HCV nucleotide sequence reads obtained by ultradeep sequencing were aligned to the consensus viral sequences in the serum specimen of each individual that were determined by direct population Sanger sequencing. A mean of 1,548-fold coverage was achieved at each nucleotide site of the HCV sequences in each specimen. First, the nucleotide sequence complexities expressed as the Shannon entropy index of HCV in the liver were compared with those in the serum. The overall viral complexity determined by the Shannon entropy index did not significantly differ between the liver and serum samples of each individual (see Fig. S1 in the supplemental material). Moreover, the patterns and distributions of genetic heterogeneity of the viral nucleotide sequences in the liver tissue sample were similar to those observed in the serum sample of the same patient (see Fig. S2 in the supplemental material). Next, we compared the viral genome sequences in the liver tissue with those in the serum in the same patient at the sites of the reported mutations that are related to the efficacy of interferon treatment and drug resistance against

TABLE 1 Baseline characteristics of 5 patients with chronic HCV[a] genotype 1b infection

| Patient characteristic[b] | Data for patient no[c]: | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| Age (yr) | 65 | 52 | 47 | 58 | 48 |
| Sex | Female | Male | Male | Male | Male |
| Existence of HCC | + | + | − | + | − |
| Child-Pugh score | 10 | 10 | 9 | 10 | 10 |
| MELD score | 14 | 15 | 14 | 15 | 15 |
| HCV viral load (log IU/ml) | | | | | |
| Pre-LDLT | 4.6 | 6.6 | 4.9 | 5.5 | 5.9 |
| Post-LDLT | | | | | |
| 1 mo | 5.9 | 6.1 | 5.8 | 5.8 | 6.4 |
| 22 mo | | | | | 6.5 |
| HCV infection | | | | | |
| Duration of hospital visit (yr) | 37 | 18 | 3 | 24 | 13 |
| Route of infection | Blood transfusion | Unknown | Unknown | Unknown | Unknown |
| Blood type | AB identical | A identical | A identical | A identical | A incompatible |
| Immunosuppressants | Tacrolimus, MMF | Tacrolimus, MMF | Tacrolimus, PSL | Tacrolimus, MMF | Tacrolimus, PSL |

[a] HCV, hepatitis C virus.
[b] HCC, hepatocellular carcinoma; MELD, model for end-stage liver disease; LDLT, living donor liver transplantation.
[c] MMF, mycophenolate mofetil; PSL, prednisolone.

HCV protease and polymerase inhibitors (see Table S1 in the supplemental material). The prevalences of these mutations of the HCV genome in the liver were similar to those in the serum of the same patients. These findings suggested that a similar pattern of viral heterogeneity was maintained in the liver and serum of patients with chronic HCV infection.

**Early dynamic decrease of viral complexity after liver transplantation.** To clarify the changes in the viral quasispecies after liver transplantation, we investigated the change in the viral complexities of the serum specimens before and 1 month after liver transplantation in these 5 patients. The mean coverages of 1,284-fold and 1,141-fold were mapped to each reference sequence before and after liver transplantation, respectively. We then estimated the genomic complexity by calculating the Shannon entropy index for each nucleotide position before and after liver transplantation (Fig. 1A). The level of viral complexity of the blood serum HCV significantly differed between pretransplantation and posttransplantation (mean Shannon entropy index, 0.056 versus 0.029; $P = 0.043$), demonstrating that the viral quasispecies nature after reinfection and replication in the graft liver became more homogeneous than that before transplantation. To identify the specific regions in the HCV genome that were respon-

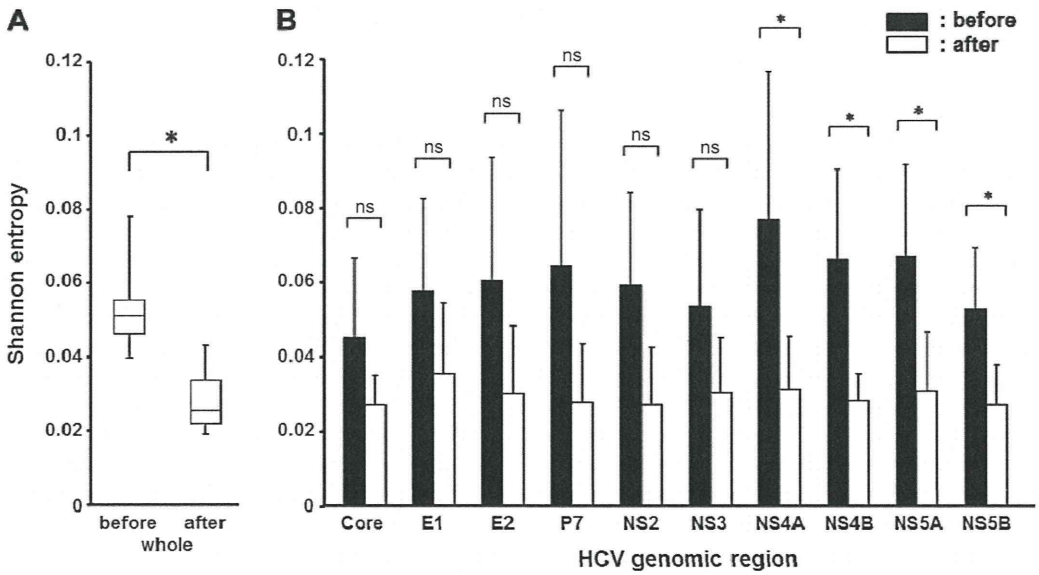


FIG 1 Changes in the genetic complexity of the HCV genome before and after liver transplantation. (A) Mean Shannon entropy index values for the overall HCV genome in 5 LDLT recipients before and after liver transplantation. (B) Mean Shannon entropy index values for each HCV genomic region before (black bars) and after (white bars) liver transplantation are shown. The error bars in panels A and B represent the standard deviation. *, $P < 0.05$; ns, nonsignificant.

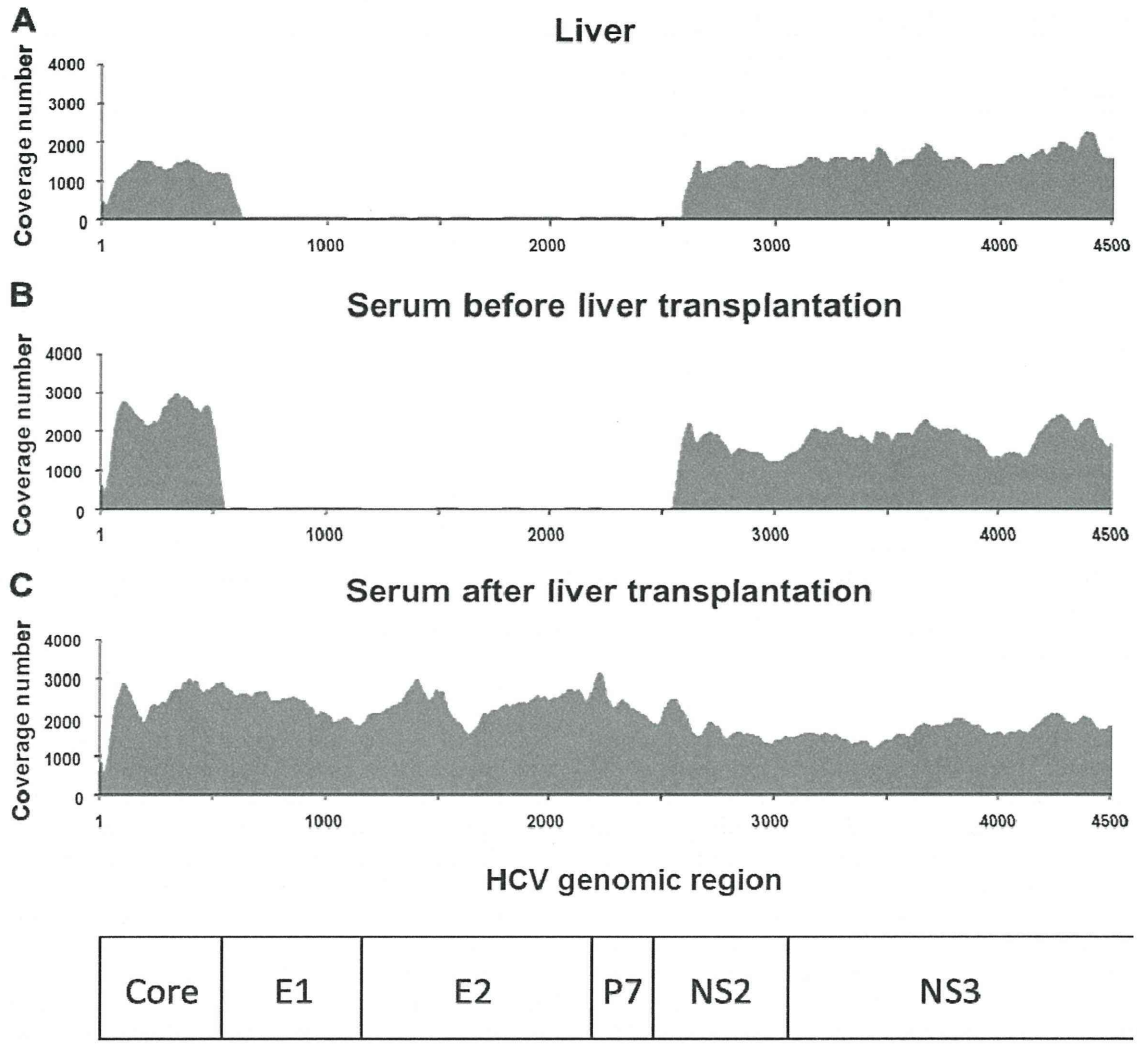


FIG 2 Dynamics of defective HCV clones indicated by coverage numbers of ultradeep sequence of HCV genome. Coverage of ultradeep sequence of HCV genome in liver (A), serum samples before liver transplantation (B), and serum samples after liver transplantation (C) for patient 1. The degree of coverage (fold) at each nucleotide site of the HCV sequence is shown. Nucleotide 1 indicates the first nucleotide of the core region of HCV RNA. Similar results were obtained in the samples from patients 2, 4, and 5.

sible for the selective increase in HCV after liver transplantation, we analyzed the changes in complexity of each region of the HCV genome (Fig. 1B). A decrease in the genetic complexity after liver transplantation was observed throughout the individual viral genetic regions. In particular, the complexity during pre- and post-transplantation was significantly different in the NS4A, NS4B, NS5A, and NS5B regions, suggesting that these regions are important for active proliferation of HCV at the early phase of reinfection in the graft liver. We then examined whether a specific nucleotide position was associated with a decrease in complexity after liver transplantation, but none of the specific nucleotide positions that changed by >50% after liver transplantation compared to before transplantation were commonly identified in the 5 patients (data not shown); this indicates that no association exists between a specific nucleotide position and the decrease in complexity after liver transplantation.

**Defective HCV clones became undetectable immediately after liver transplantation.** Using the ultradeep sequencing analyses, we found that the sequence coverage of viral genomic regions spanning from the end of the core to the middle of NS2 was smaller than those of the other regions in several liver and serum samples before liver transplantation, but this tendency was not observed in the samples after liver transplantation (Fig. 2). As we previously identified the defective HCV clones lacking the structural regions of the HCV genome in the serum samples of HCV-positive liver transplant recipients (8), we speculated that the presence of the defective HCV clones would result in the smaller coverage of E1-NS2 before transplantation, and the population of the defective clones would change after liver transplantation. Therefore, we next analyzed the population change of the defective HCV clones between before and after liver transplantation. Using RT-PCR analysis with the primers HCV-P1 and HCV-P2 (Fig. 3A), we detected both defective HCV clones and full-genome HCV clones before liver transplantation at various ratios in each sample, except for in patient 3 (Fig. 3B). The defective HCV clones became undetectable and the full-genome HCV clones became

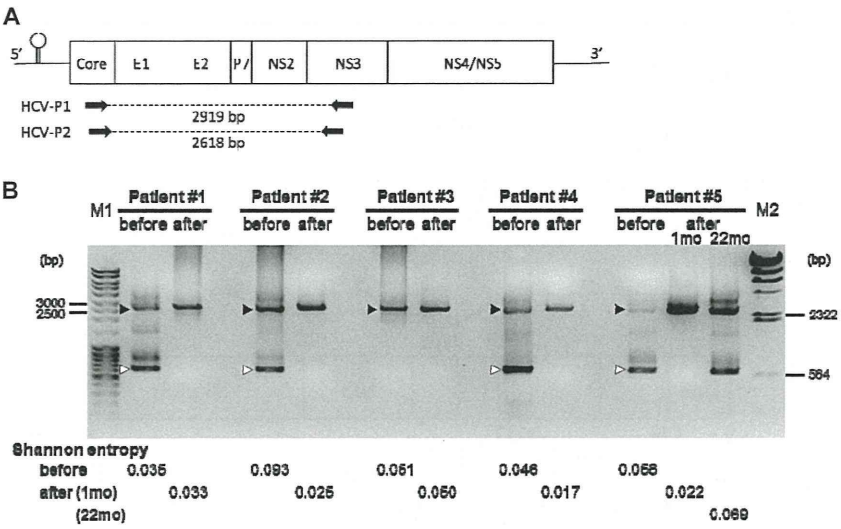


FIG 3 Dynamics of defective HCV clones based on RT-PCR analysis. (A) Schematic presentation of the HCV genome and the primer sets used in this study. (B) Results of RT-PCR analysis by using RNA samples as a template, which were extracted from blood serum before and 1 month after liver transplantation in all patients and 22 months after transplantation in patient 5. HCV-P1 and HCV-P2 (panel A) were used as primers. Lanes M1 and M2, the molecular weight markers MassRuler DNA ladder mix (Fermentas, Canada) and Lambda DNA-HindIII Digest (New England BioLabs, USA), respectively. The values shown indicate the sizes of the band in the molecular weight markers. Black arrowheads, full-length PCR fragment of 2,618 bp; white arrowheads, defective HCV clones that were confirmed by sequencing analysis. The Shannon entropy index values of these HCV specimens in the serum (before and after liver transplantation) are shown at the bottom.

dominant in the serum samples 1 month after liver transplantation, indicating that the defective HCV clones have less of a replication advantage than the full-genome clones. In patient 3, the defective HCV clones were undetectable both before and after liver transplantation.

To determine the internal structure of the deletions in the defective HCV genomes, major amplified fragments from each of the four patients with defective HCV clones before transplantation were subcloned for further sequence analyses. Schematic representations of the defective HCV RNA detected in the blood serum specimens of these patients are shown in Fig. 4. Sequence analyses revealed that the structural region was widely deleted in all of the defective HCV clones. The 3′-boundaries of the deletions were quite diverse in the clones, while the 5′-untranslated region and core regions were preserved in all four clones, as reported previously (8). Two distinct defective clones were found in patient 2. All of the deletions identified were in frame, implying that these defective HCV genomes have the potential for translation from the core to the authentic end of NS5B without a frameshift.

We then analyzed the dynamics of the defective HCV clones at the chronic hepatitis phase after liver transplantation in patient 5. As shown in the Patient 5 column in Fig. 3B, RT-PCR from a serum sample collected at 22 months after liver transplantation, when a liver biopsy specimen demonstrated findings of chronic hepatitis C with fibrosis (METAVIR score, A1 F1), showed that a defective HCV clone had reappeared. The size of the defective clone was the same as that found in the serum before transplantation, and we confirmed by sequence analysis that the deleted region of the defective HCV clone was identical to that in the pretransplant serum sample. The viral complexity analyzed by calculating the Shannon entropy index from ultradeep sequencing data also returned to the pretransplantation level at the chronic hepatitis phase (Shannon entropy values, 0.056 before transplantation, 0.022 at 1 month posttransplantation, and 0.069 at 22 months after liver transplantation). These findings indicated that the reconstitution of HCV heterogeneity occurs at the chronic hepatitis phase after liver transplantation, and the same defective HCV clone present before liver transplantation reappears at the chronic hepatitis phase after liver transplantation.

## DISCUSSION

The present study revealed two major findings from ultradeep sequencing analyses of the HCV genome sequence in liver transplant recipients before and after liver transplantation. First, the viral heterogeneity of HCV significantly decreased after liver transplantation, indicating that the clones with advantages for infection and/or replication in hepatocytes rapidly increased after liver transplantation. Second, full-genome HCV clones selectively increased, while the defective clones did not increase in number during the period immediately after liver transplantation.

The discovery of differences in the populations of HCV quasispecies between the liver and serum of the same individuals has been controversial. Most previous studies examined the HCV sequencing mainly for the hypervariable region in E2 using the Sanger sequencing method (11–13) or single-strand conformation polymorphism (12, 14, 15), but the findings were conflicting. In the present study, we obtained full-genome HCV sequences using ultradeep sequencing analysis. Our results suggested that a similar HCV population exists in the liver and blood serum, at least at the specific sites related to interferon sensitivity and drug resistance. These results are clinically important because we confirmed that the serum samples, which are easily obtained from patients, reflect the HCV population in the liver and are thus useful for analyses of resistance and sensitivity to treatment.

Differences in the HCV population between individuals can be determined by multiple factors, such as the duration of hospital visit, route of HCV infection, fibrosis progression, degree of inflammation, and the presence of hepatocellular carcinoma. In our

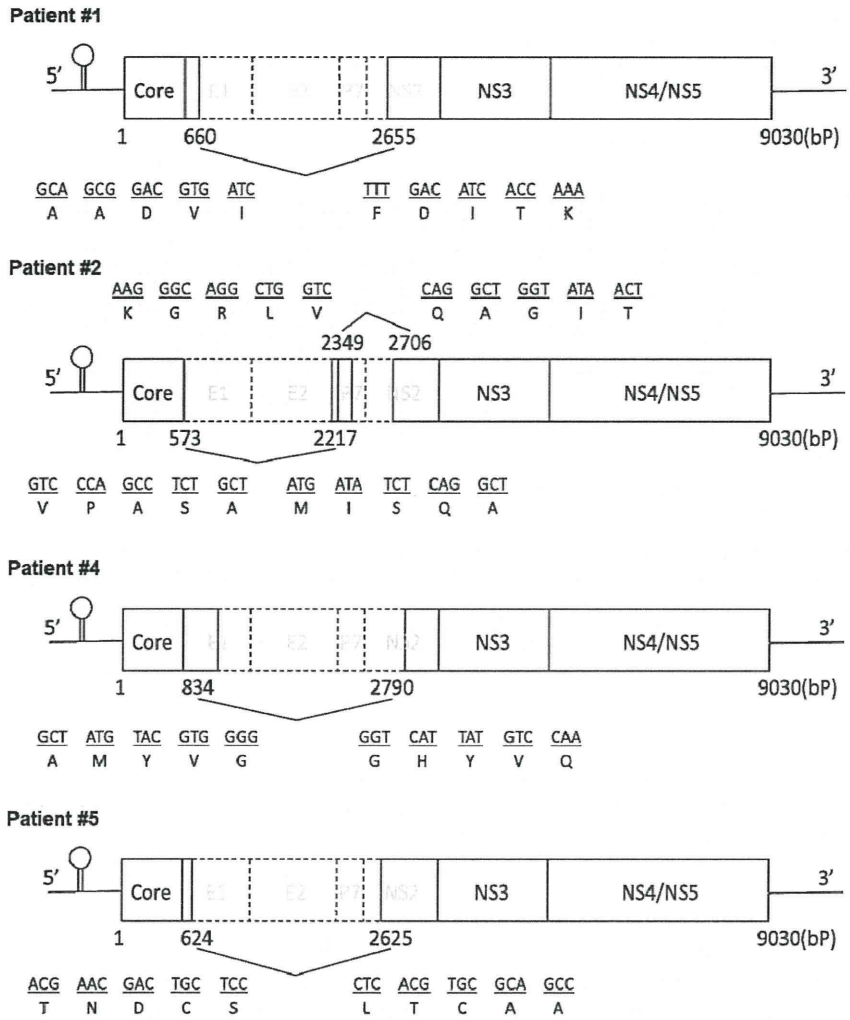


**FIG 4** Schematic presentation of major defective HCV clones in 4 patients before liver transplantation. The values in the schema indicate the nucleotide numbers from the first ATG of the core region in HCV RNA. Nucleotide and amino acid sequences before and after the deleted region of the HCV genome are shown. E1, envelope glycoprotein 1; E2, envelope glycoprotein 2; NS, nonstructural protein.

analysis, we could not find an association between these clinical characteristics and the nature of the HCV population between patients. However, we speculated that undetectable defective HCV clones present before liver transplantation in patient 3 might be associated with a shorter duration of HCV infection. In patients 1 and 3, the difference in viral complexity as measured by the Shannon entropy index values between before and after liver transplantation was small. The reason is unclear at present, but differences in the clinical features of infected patients might affect the results. Further large-scale investigations may reveal the relationship between clinical features in patients and the nature of a specific HCV population.

Our large-scale analysis using ultradeep sequencing demonstrated that the complexity of all regions of the HCV genome was dramatically reduced 1 month after liver transplantation compared with the pretransplantation level of complexity. This finding is consistent with findings from previous reports using Sanger sequencing methods that showed that heterogeneity is decreased in the hypervariable region of E2 of HCV after liver transplantation (16, 17). Gretch et al. (16) analyzed HCV quasispecies before and after liver transplantation by comparing the differences in the hypervariable region of the HCV genome in 5 transplant recipients. They found that different HCV clones were present in pretransplant blood serum and relatively homogeneous quasispecies variants emerged after liver transplantation in all 5 cases. Hughes et al. (17) demonstrated that the viral complexity of the hypervariable region 1 in postperfusion liver tissue at 2.5 h after liver transplantation was significantly lower than that in explanted liver and in pretransplant serum, although there was no significant difference in the complexity between the explanted liver and pretransplant serum. Our present data confirmed the results of these previous studies and added new information from the full-genome ultradeep sequencing. In particular, our data demonstrated a new aspect of the analyses of full-genome and defective HCV clones, because the defective HCV clones lack hypervariable regions that were analyzed in previous studies. Interestingly, our analysis revealed significant decreases in complexity in the NS4A, NS4B, NS5A, and NS5B regions after transplantation, although a decreasing trend was detected in all regions of the HCV genome. Because the region from NS4A to NS5B has important roles in

HCV replication (18–20), a decrease in the complexity of the NS4A to NS5B sequence after liver transplantation might indicate the presence of the specific NS4A to NS5B sequence in the HCV genome that confers advantages in the reinfection and/or replication processes. Therefore, we attempted to identify the specific HCV genome sequences with such advantages. However, we could not identify a common feature of the HCV genomic changes in amplified HCV clones after liver transplantation among the 5 cases tested. This may be due to differences between individuals in the relative fitness of a viral subpopulation in its host, which is determined by multiple factors, including infection capacity, replication ability, and mechanisms by which to escape from immune pressure.

We previously identified defective HCV clones in the blood serum of patients after liver transplantation (8). Other groups also reported that defective HCV clones exist in the liver and serum of patients with chronic hepatitis C and patients with immunosilent infections (21–25). These reports demonstrated that deletions in the HCV genome were present mainly in the structural region, while the 5′-untranslated region, the core, and NS3 to NS5B regions were preserved, and that most of the deletions were in frame, indicating that the preserved regions can be translated to the authentic terminus. Indeed, Sugiyama et al. (24) recently demonstrated that the defective genome can be translated, self-replicated, and encapsidated as an infectious particle by *trans*-complementation of the structural proteins *in vitro*. Pacini et al. (23) also reported that defective HCV clones show robust replication, efficient *trans*-packaging, and infection of cultured cells. These data suggest that the abilities of defective HCV genomes to infect, replicate, and be encapsidated do not differ from those of full-genome HCV. The *in vivo* data reported here, however, clearly reveal that the amount of defective HCV clones was lower than that of full-genome HCV after liver transplantation, although the reason for this remains unknown. One possibility is that the capabilities to infect, replicate, or be encapsidated differ between defective HCV and full-genome HCV *in vivo*. It is noteworthy that an identical defective HCV clone that was detected before transplantation reappeared during the chronic hepatitis phase after transplantation in patient 5. This finding suggests that the defective clone in the blood serum also infected the graft liver, replicated, and was encapsidated in the graft liver after liver transplantation. Therefore, the speed of these steps would differ between defective HCV clones and full-genome HCV clones.

The present study revealed a limitation of the massively parallel ultradeep sequencing technology in the analyses of viral quasispecies. Because the massively parallel ultradeep sequencing platform is based on multitudinous short reads, it is difficult to separately evaluate the association between nucleotide sites that are mapped to different viral genome regions in a single viral clone. Indeed, it is difficult to clarify the potential mutational linkage between different viral genomic regions because of the short read lengths of the shotgun sequencing approach.

In conclusion, after liver transplantation, viral heterogeneity decreased significantly and the number of full-genome HCV clones increased immediately, whereas the defective HCV clones began to increase in number over a longer period. Further analysis will reveal the significance of the changes in defective HCV clones after liver transplantation.

## ACKNOWLEDGMENTS

This work was supported by Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research no. 21229009 and 23590972 and Health and Labor Sciences research grants for research on intractable diseases and research on hepatitis from the Ministry of Health, Labor, and Welfare, Japan.

## REFERENCES

1. **Shimotohno K.** 1995. Hepatitis C virus as a causative agent of hepatocellular carcinoma. Intervirology **38**:162–169.
2. **Mardis ER.** 2009. New strategies and emerging technologies for massively parallel sequencing: applications in medical research. Genome Med. **1**:40. doi:10.1186/gm40.
3. **Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437**:376–380.
4. **Ueda Y, Marusawa H, Kaido T, Ogura Y, Ogawa K, Yoshizawa A, Hata K, Fujimoto Y, Nishijima N, Chiba T, Uemoto S.** 2012. Efficacy and safety of prophylaxis with entecavir and hepatitis B immunoglobulin in preventing hepatitis B recurrence after living-donor liver transplantation. Hepatol Res. **43**:67–71.
5. **Ohno O, Mizokami M, Wu RR, Saleh MG, Ohba K, Orito E, Mukaide M, Williams R, Lau JY.** 1997. New hepatitis C virus (HCV) genotyping system that allows for identification of HCV genotypes 1a, 1b, 2a, 2b, 3a, 3b, 4, 5a, and 6a. J. Clin. Microbiol. **35**:201–207.
6. **Nasu A, Marusawa H, Ueda Y, Nishijima N, Takahashi K, Osaki Y, Yamashita Y, Inokuma T, Tamada T, Fujiwara T, Sato F, Shimizu K, Chiba T.** 2011. Genetic heterogeneity of hepatitis C virus in association with antiviral therapy determined by ultra-deep sequencing. PLoS One **6**:e24907. doi:10.1371/journal.pone.0024907.
7. **Nishijima N, Marusawa H, Ueda Y, Takahashi K, Nasu A, Osaki Y, Kou T, Yazumi S, Fujiwara T, Tsuchiya S, Shimizu K, Uemoto S, Chiba T.** 2012. Dynamics of hepatitis B virus quasispecies in association with nucleos(t)ide analogue treatment determined by ultra-deep sequencing. PLoS One **7**:e35052. doi:10.1371/journal.pone.0035052.
8. **Iwai A, Marusawa H, Takada Y, Egawa H, Ikeda K, Nabeshima M, Uemoto S, Chiba T.** 2006. Identification of novel defective HCV clones in liver transplant recipients with recurrent HCV infection. J. Viral Hepat. **13**:523–531.
9. **Fishman SL, Branch AD.** 2009. The quasispecies nature and biological implications of the hepatitis C virus. Infect. Genet. Evol. **9**:1158–1167.
10. **Wolinsky SM, Korber BT, Neumann AU, Daniels M, Kunstman KJ, Whetsell AJ, Furtado MR, Cao Y, Ho DD, Safrit JT.** 1996. Adaptive evolution of human immunodeficiency virus-type 1 during the natural course of infection. Science **272**:537–542.
11. **Cabot B, Martell M, Esteban JI, Sauleda S, Otero T, Esteban R, Guàrdia J, Gómez J.** 2000. Nucleotide and amino acid complexity of hepatitis C virus quasispecies in serum and liver. J. Virol. **74**:805–811.
12. **Jang SJ, Wang LF, Radkowski M, Rakela J, Laskus T.** 1999. Differences between hepatitis C virus 5′ untranslated region quasispecies in serum and liver. J. Gen. Virol. **80**(Pt 3):711–716.
13. **Sakai A, Kaneko S, Honda M, Matsushita E, Kobayashi K.** 1999. Quasispecies of hepatitis C virus in serum and in three different parts of the liver of patients with chronic hepatitis. Hepatology **30**:556–561.
14. **De Mitri MS, Mele L, Chen CH, Piccinini A, Chianese R, D'Errico A, Alberti A, Pisi E.** 1998. Comparison of serum and liver hepatitis C virus quasispecies in HCV-related hepatocellular carcinoma. J. Hepatol. **29**: 887–892.
15. **Sakamoto N, Enomoto N, Kurosaki M, Asahina Y, Maekawa S, Koizumi K, Sakuma I, Murakami T, Marumo F, Sato C.** 1995. Comparison of the hypervariable region of hepatitis C virus genomes in plasma and liver. J. Med. Virol. **46**:7–11.
16. **Gretch DR, Polyak SJ, Wilson JJ, Carithers RL, Jr, Perkins JD, Corey L.** 1996. Tracking hepatitis C virus quasispecies major and minor variants in

symptomatic and asymptomatic liver transplant recipients. J. Virol. **70:** 7622–7631.
17. **Hughes MG, Jr, Rudy CK, Chong TW, Smith RL, Evans HL, Iezzoni JC, Sawyer RG, Pruett TL.** 2004. E2 quasispecies specificity of hepatitis C virus association with allografts immediately after liver transplantation. Liver Transpl. **10:**208–216.
18. **Gao L, Aizaki H, He JW, Lai MM.** 2004. Interactions between viral nonstructural proteins and host protein hVAP-33 mediate the formation of hepatitis C virus RNA replication complex on lipid raft. J. Virol. **78:** 3480–3488.
19. **Moradpour D, Brass V, Bieck E, Friebe P, Gosert R, Blum HE, Bartenschlager R, Penin F, Lohmann V.** 2004. Membrane association of the RNA-dependent RNA polymerase is essential for hepatitis C virus RNA replication. J. Virol. **78:**13278–13284.
20. **Shimakami T, Hijikata M, Luo H, Ma YY, Kaneko S, Shimotohno K, Murakami S.** 2004. Effect of interaction between hepatitis C virus NS5A and NS5B on hepatitis C virus RNA replication with the hepatitis C virus replicon. J. Virol. **78:**2738–2748.
21. **Bernardin F, Stramer SL, Rehermann B, Page-Shafer K, Cooper S, Bangsberg DR, Hahn J, Tobler L, Busch M, Delwart E.** 2007. High levels of subgenomic HCV plasma RNA in immunosilent infections. Virology **365:**446–456.
22. **Noppornpanth S, Smits SL, Lien TX, Poovorawan Y, Osterhaus ADME, Haagmans BL.** 2007. Characterization of hepatitis C virus deletion mutants circulating in chronically infected patients. J. Virol. **81:**12496–12503.
23. **Pacini L, Graziani R, Bartholomew L, De Francesco R, Paonessa G.** 2009. Naturally occurring hepatitis C virus subgenomic deletion mutants replicate efficiently in Huh-7 cells and are *trans*-packaged *in vitro* to generate infectious defective particles. J. Virol. **83:**9079–9093.
24. **Sugiyama K, Suzuki K, Nakazawa T, Funami K, Hishiki T, Ogawa K, Saito S, Shimotohno KW, Suzuki T, Shimizu Y, Tobita R, Hijikata M, Takaku H, Shimotohno K.** 2009. Genetic analysis of hepatitis C virus with defective genome and its infectivity *in vitro*. J. Virol. **83:** 6922–6928.
25. **Yagi S, Mori K, Tanaka E, Matsumoto A, Sunaga F, Kiyosawa K, Yamaguchi K.** 2005. Identification of novel HCV subgenome replicating persistently in chronic active hepatitis C patients. J. Med. Virol. **77:**399–413.