

520

521 **Supplementary Figure 1. Associations between T-cell Repertoire Metrics and Clinical**
522 **Variables.** Fisher's Exact Tests were used to compare two categorical variables, Mann-
523 Whitney/Kruskal-Wallis tests were used to compare Categorical vs. numerical variables, and
524 Spearman tests were used to compare two numerical variables. P values are indicated in each
525 cell, and tests with $p < 0.05$ are highlighted in black text

526

527 **Supplementary Figure 2. Effects of Donor Type and Conditioning Intensity on the T-cell**
528 **Repertoire** **A.** Effect of Donor type on the T-cell repertoire post-transplant. Only subjects with
529 matched related donors (MRD) and matched unrelated donors (MUD) are shown, as there were
530 only 6 mismatched unrelated donors, and 3 mismatched related donors. * $p < 0.05$ by Mann-
531 Whitney-U test. **B.** Effect of conditioning intensity on the T-cell repertoire post-transplant. * $p <$
532 0.05 by Mann-Whitney-U test.

533

534 **Supplementary Figure 3. Cumulative incidence of acute GVHD stratified by clonal**
535 **expansion and additional clinical metrics** **A.** Cumulative incidence of acute GVHD stratified

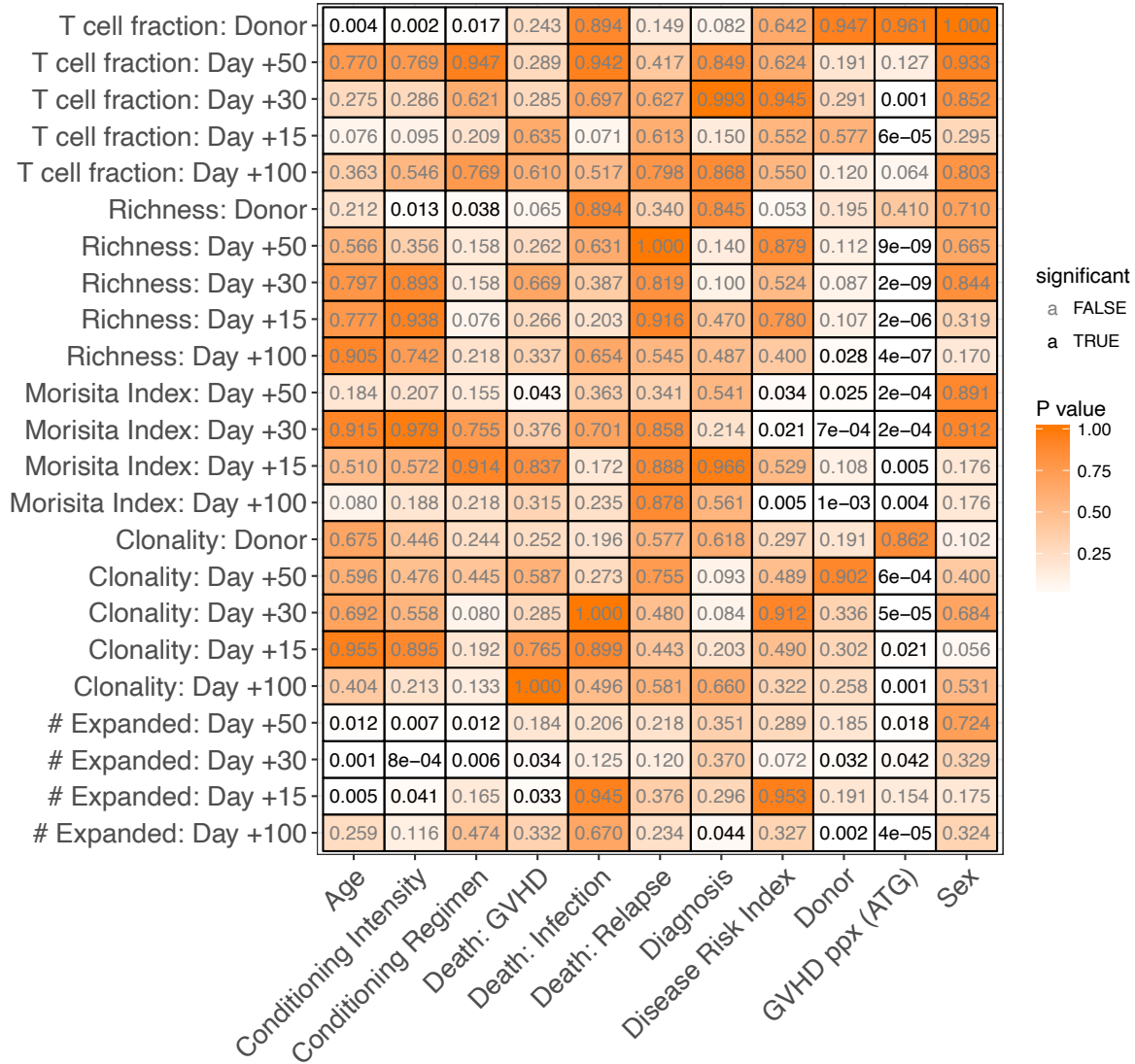
536 by clonal expansion (high, greater than the median; low, less than the median) and use of ATG.
537 B. Cumulative incidence of acute GVHD stratified by clonal expansion (high, greater than the
538 median; low, less than the median) and conditioning intensity C. Cumulative incidence of acute
539 GVHD stratified by clonal expansion (high, greater than the median; low, less than the median)
540 and donor type (MRD = matched related donor; MUD = matched unrelated donor)
541

542 **Supplementary Figure 4. Association of clinical and immunoSEQ metrics with overall**
543 **survival** A. immunoSEQ and clinical variables with single variable Cox Regression $P < 0.2$ are
544 shown. 95% confidence intervals are shown.
545

546 **Supplementary Figure 5. Other associations with transplant outcome** A. Kaplan Meier curve
547 of overall survival stratified by richness at Day +30 (high, greater than the median; low, less than
548 the median) and use of ATG. B. Kaplan Meier Curve of overall survival stratified by richness at
549 Day +30 (high, greater than the median; low, less than the median) after removing subjects given
550 ATG.
551

552 **Supplementary Figure 6. Association of CMV reactivation with TCR repertoire changes.**
553 A) TCR repertoire metrics and clinical variables with single variable Cox Regression $p < 0.2$ are
554 shown with 95% confidence interval for association of CMV reactivation before day +150. B)
555 Cumulative incidence curve of time to CMV reactivation stratified by high (above the median)
556 and low (below the median) C) Productive clonality at various time points stratified by CMV
557 reactivation status.

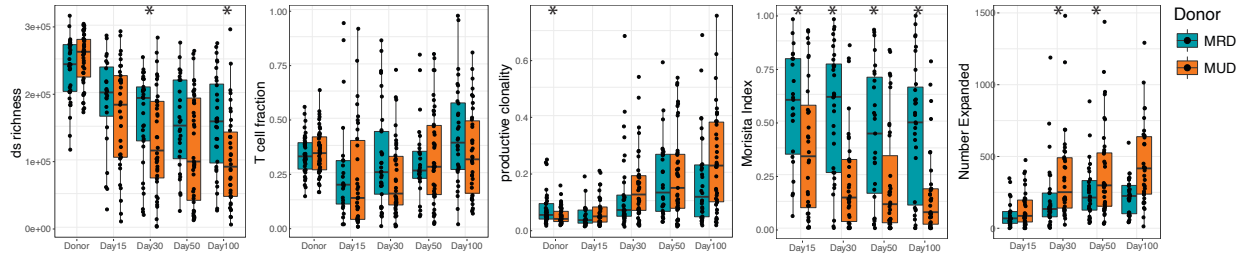
575 **Supplementary Figure 1.**



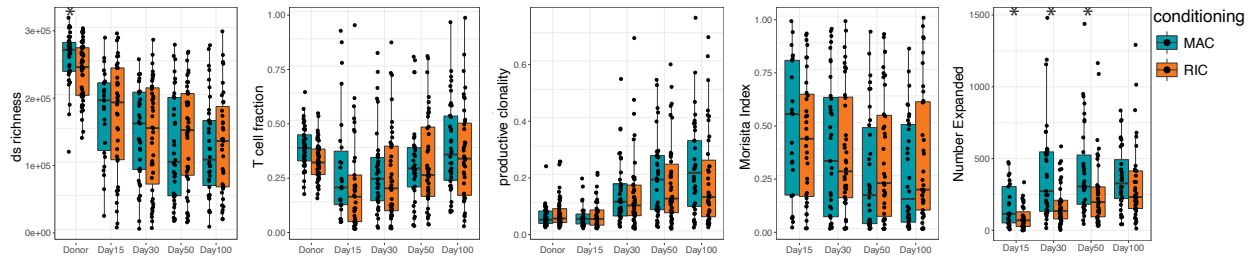
576
577
578
579

580 **Supplementary Figure 2.**

A

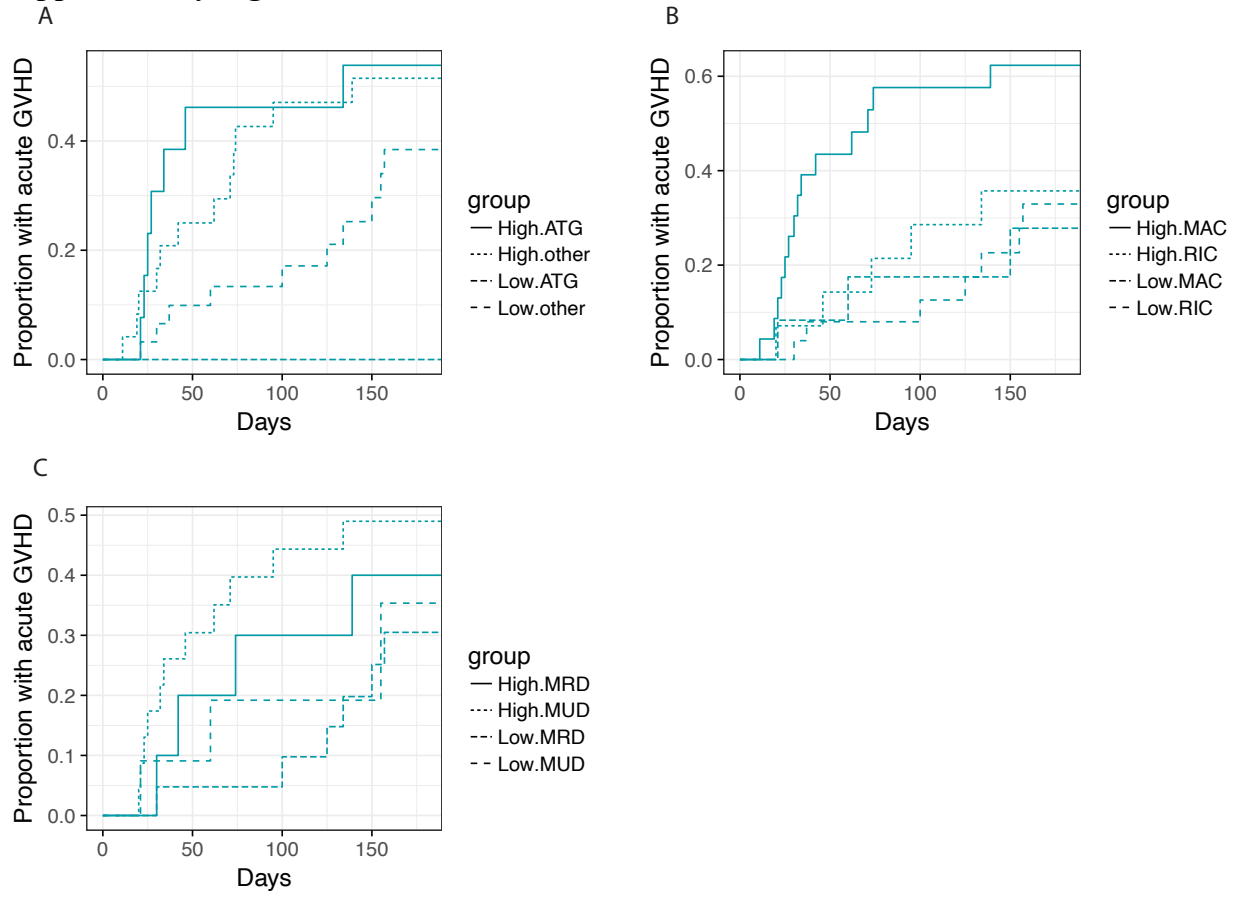


B



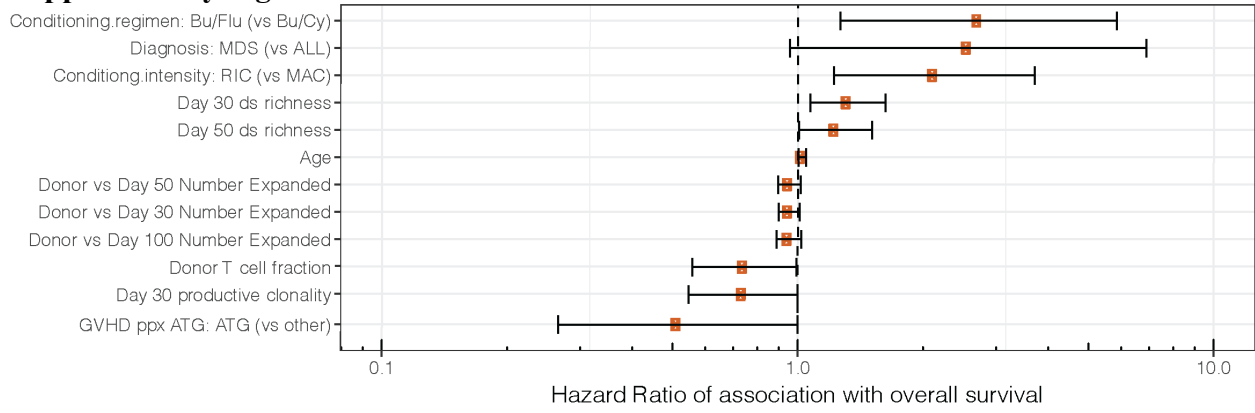
581
582

583 **Supplementary Figure 3.**



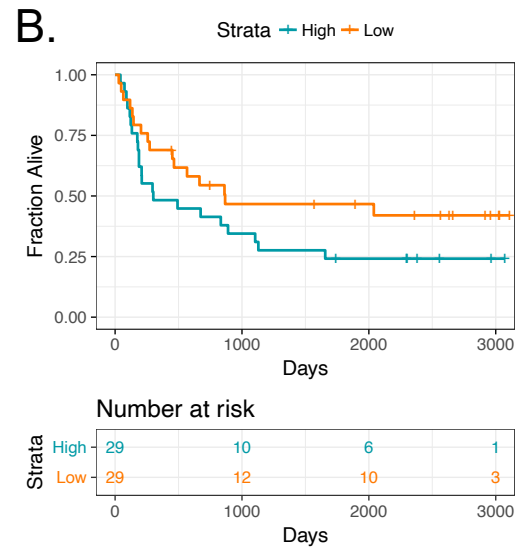
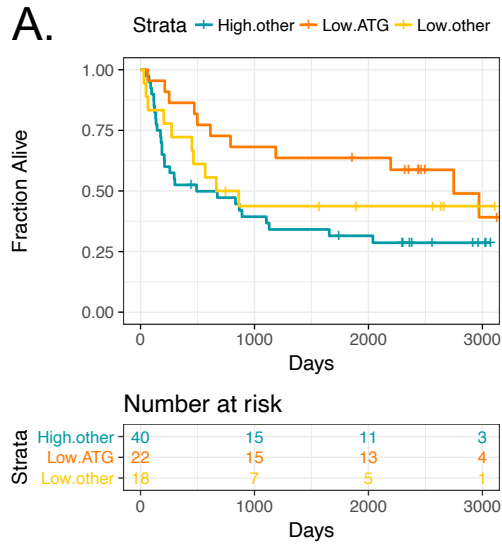
584
585

586 **Supplementary Figure 4**



587
588

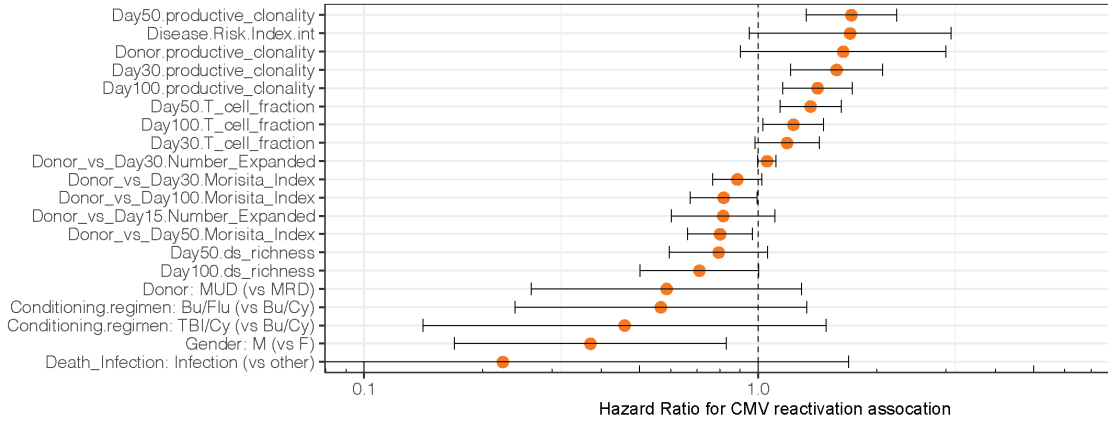
589 **Supplementary Figure 5**



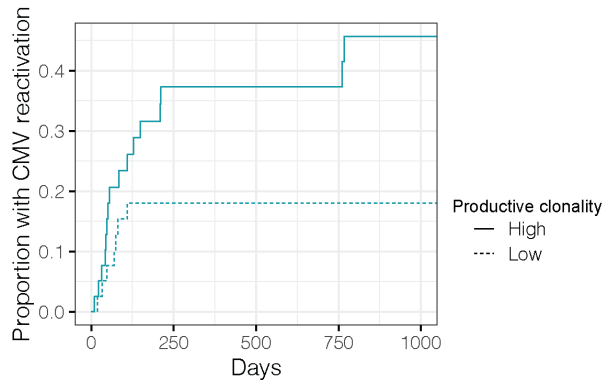
590
591

592 **Supplementary Figure 6**
 593

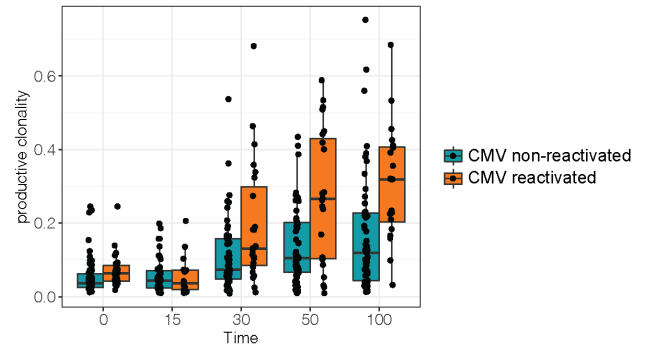
A



B



C



594
 595
 596

603

604 **Supplementary Data**

605

606 1. Definitions

607

608 **Clonality**

609 Clonality was defined as 1–Peilou’s evenness¹². Clonality was calculated on productive
610 rearrangements by:

$$611 \text{ clonality} = 1 + \frac{\sum_i^N p_i \log_2(p_i)}{\log_2(N)}$$

612

613 where p_i is the proportional abundance of rearrangement i and N is the total number of
614 rearrangements.

615

616 **Number of Expanded Clones Compared to Donor**

617 A standard binomial two-sided test of the null hypothesis that the probability of success in a
618 Bernoulli experiment is p is computed for each clone. In this framework:

619

$$620 \Pr(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$621 p = \frac{N1}{N1+N2} \text{ (probability of success)}$$

622 $N1$ = Total templates in sample 1

623 $N2$ = Total templates in sample 2

624 k = Number of templates of the given clone in sample 1 (number of successes)

625 $n2$ = Number of templates of the given clone in sample 2

626 $n = k + n2$ (number of trials)

627

628 Clones passing a Benjamini-Hochberg multiple test correction at $\alpha=0.01$ are considered
629 significantly expanded or contracted, depending on whether they have a higher frequency in the
630 post-transplant sample compared to the donor. The number of expanded clones was simply the
631 number of all significant clones at a higher frequency in the post-transplant sample compared to
632 the donor.

633

634

635 **Morisita’s Index**

636 Morisita’s Index was defined by:

$$MI = \frac{2 \sum_i^S a_i b_i}{\left(\frac{\sum_i^S a_i^2}{A^2} + \frac{\sum_i^S b_i^2}{B^2} \right) AB}$$

637

638 where a_i and b_i are the number of templates of clone i in sample A and sample B, respectively. A
639 and B are the total number of templates in sample A and sample B, respectively.

640

641