

Haplotype frequency sets in the Search & Match Service v1 (ARCHIVE)

THIS IS INFO REGARDING V1 OF SEARCH AND MATCH WHICH USES OPTIMATCH. Please check [Haplotype frequency sets in the HAP-E matching algorithm of the Search & Match Service](#) for the information regarding V2 which uses Hap-E.

With the last release of our the Search & Match Service, WMDA introduces a new haplotype frequency set configuration to better serve all patients and provide increased accuracy when matching donors and cord blood units to patients. The Search & Match Service currently contains no ethnicity information for donors and therefore the new frequency sets will be calculated and based upon geographical and available data of the donors of an organisation. These frequency sets utilize high resolution typing results from organisations to extrapolate the haplotypes of the region or organisation. Thus, an organisation or geographical region must meet a minimum threshold of number of donors and availability of high resolution typing to build usable and valuable frequencies. For cord blood units, we will use the same sets that were determined in the donor populations.

The inclusion criteria aims to balance a high number of donors, the quality of the HLA types and the complexity of the haplotype frequency estimation. The number and HLA types of the loci that are included in the estimation can be independent of the loci of the haplotypes actually estimated if the quality of the former is appropriate.

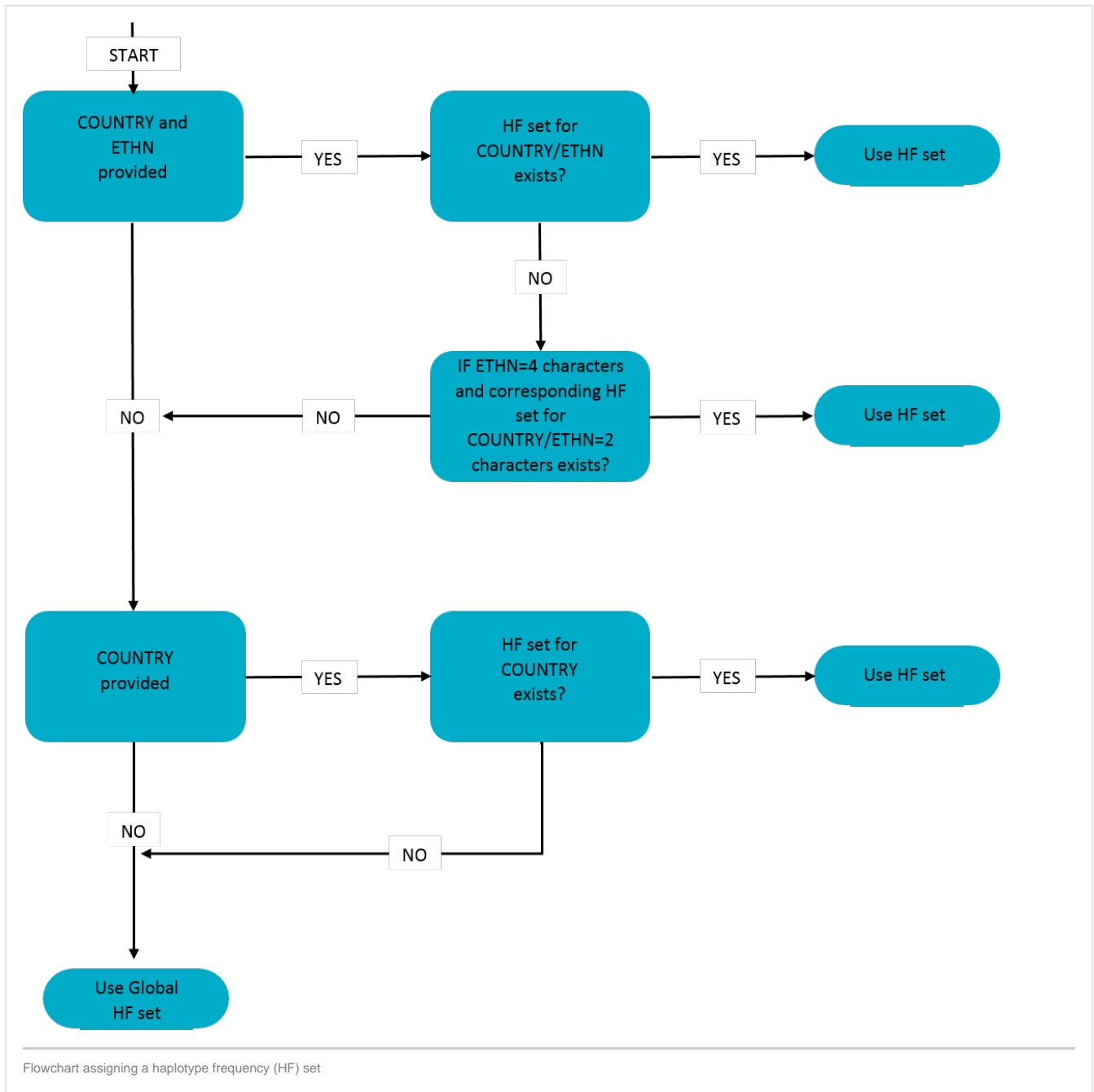
5-locus Haplofrequency estimation was performed on all organisations/populations with at least 5000 donors in HLA-A, -B, -C and -DRB1 types in 90% high resolution. Organisations in the same country/within the same population are usually combined, e.g. AT-ABMDR (Austrian Bone Marrow Donors) and AT-Verein (Austria- Verein geben fur Lebel) but *not* for US-NMDP (general US population) and US-GOL (predominantly Ashkenazi Jews). The 90% criteria is based on the global frequency set and includes all allele codes whose top 90% cumulated allele frequencies are in high resolution.

The current global frequency set will continue to be used for donors of organisations if it is the best representation. The number of this frequency set is 999.

Below, you can find a figure how the haplotype frequency (HF) sets are assigned to the donors and cord blood units. The system first determines if both COUNTRY and ETHNICITY are provided. If yes, then it will check if an ethnicity specific set for this country exists. If the donor for example is coming from USA and has ethnicity HISA, then we do not have a specific set available. However, there is a more broad Hispanic HF set available (USA-HI) and this will then be used.

If there is no ethnicity available for the donor or no ethnicity specific HF set exists for that country, the system will check if a country specific HF set is available or if the country is part of one of the regional HF sets.

If no country specific HF set is available, the Search & Match Service will use the global consensus HF set.



In the table below, you can find the haplotype frequency set numbers and which countries/organisations were included as well as the total sample size. In the last part from the table we included regional sets, some country/ethnicity specific sets and the global consensus set.

The regional sets are defined as follows:

East Asia (eas): CN (CN+CN1), HK, TW, JP (Set number: 32)

Eastern Europe (eeu): AT (A+A2), CZ (CS+CS2), CY (CY+CY2), GR2, PL (PL3+PL5+PL6), TR (TRAN+TRIS+TRKK), BG, HU, HR, LT, MK, RU (R2+R4), RO, RS, SK, SI (Set number: 33)

South America (sam): AR, BR, UY (Set number: 34)

			INPUT DATA
--	--	--	-------------------

HF set number (pop_id)	Determination date	Sample size	ISO country code	I ONS	Search & Match Service_Registry_codes	Old BMDW Registry codes
1	2017-04-25	26,352	AR	51 17	AR-INCUCAI	AR
2	2017-04-25	21,709	AT	26 14 49 61	AT-ABMDR AT-Verein	A A2
3	2017-04-25	38,548	AU + NZ	77 48 82 61	AU-ABMDR NZ-NZBMDR	AUS NZ
4	2017-04-25	49,392	BE	42 01	BE-MDPB	B
5	2017-04-25	82,656	BR	87 66	BR-REDOME	BR
6	2017-04-25	214,836	CA	51 03 69 12	CA-One Match CA-HemaQuec	CND CND3
7	2017-04-25	89,533	CH	93 41	CH-SBSC	CH
8	2017-04-25	591,099	CN + HK + TW	21 97 66 81 40 70 34 58	CN-CMDP CN-Sunsh HK-HKBMDR TW-Tzu Chi	CN CN1 HK TW
9	2017-04-25	27,230	CY	42 78 97 51	CY-Par BMDR CY-CBMDR	CY CY2
10	2017-04-25	36,878	CZ	47 53 54 40	CZ-CSCR CZ-CNMDR	CS CS2
11	2017-04-25	6,273,693	DE	69 39	DE-ZKRD	D
12	2017-04-25	20,364	DK	20 15 74 84	DK-DSCDW DK-DSDE	DK DK2
13	2017-04-25	59,177	ES	78 13	ES-REDMO	E(+E2)
14	2017-04-25	17,377	FI	97 38	Fi-FSCR	FI
15	2017-04-25	173,541	FR	18 04	FR-FGM	F

16	2017-04-25	1,050,556	GB + IE	63 54 17 26 27 31 99 68 55 90	GB-Anthony GB-WBMDR GB-BBMR GB-DKMS IE-IUBMR	GB GB3 GB4 GB5 IRL
17	2017-04-25	38,855	GR	14 61	GR-CBMDP	GR2
18	2017-04-25	377,768	IL	52 39 49 87 40 68	IL-Hadassah IL-Ezer Miz. IL-SHBB	IL IL2 IL3
19	2017-04-25	208,655	IN	84 86 28 24 41 31 45 96	IN-Datri IN-GeneBand IN-MDR IN-BMST	IN2 IN3 IN4 IN5
20	2017-04-25	150578	IT	74 50	IT-IBMDR	I
21	2017-04-25	129,272	NL	81 39	NL-Matchis	NL
22	2017-04-25	20,925	NO	72 14	NO-NBMDR	N
23	2017-04-25	1,167,906	PL	39 18 53 91 74 14	PL-ALF PL-Poltranspl PL-DKMS	PL3 PL5 PL6
24	2017-04-25	158,284	PT	72 58	PT-Cedace	P
25	2017-04-25	25,498	SA	21 07	SA-SSCDR	SA
26	2017-04-25	45,892	SE	52 85	SE-Tobias	S
27	2017-04-25	62,785	SG	37 85	SG-BMDP	SG
28	2017-04-25	40,149	TH	83 62	TH-TSCDR	TH
29	2017-04-25	32,572	TR	38 93 55 09 35 03	TR-TRAN TR-TRIS TR-TURKOK	TRAN TRIS TRKK
30	2017-04-25	4,068,326	US	35 53	US-NMDP	USA1
31	2017-04-25	163,588	US	10 33	US-GOL	USA4
32	2017-04-25	792,159	CN+HK+TW+JP	eas	<i>Search&Match:eas</i>	CN+CN1+HK+TW+JP

33	2017-04-25	1,369,965	AT+CZ+CY+GR+PL+TR+BG+HU+HR+LT+MK+RU+RO+RS+SK+SI	eeu	Search&Match:eeu	A+A2+CS+CS2+CY+CY2+GR+GR2+PL3+PL5+PL6+TRAN+TRIS+TRKK+BG+H+HR+LT+MK+R2+R4+RO+RS+SK+SLO
34	2017-04-25	109,841	AR+BR+UY	sam	Search&Match:sam	AR+BR+UY
100	2016-09-30	151,204	DE-ASSW	-	DE-ASSW subset	DE-ASSW
101	2017-04-05	656,591	US-AF (ZKRD)	-	US-AF subset (ZKRD_estimation)	USA1-AF
102	2017-04-05	796,780	US-AS (ZKRD)	-	US-AS subset (ZKRD_estimation)	USA1-AS
103	2017-04-05	3,740,668	US-CA (ZKRD)	-	US-CA subset (ZKRD_estimation)	USA1-CA
104	2017-04-05	1,002,893	US-HI (ZKRD)	-	US-HI subset (ZKRD_estimation)	USA1-HI
999	2015-12-16	11,430,561			Global consensus	All Search & Match Service donor registries with a sufficient number of donors with the defined level of HLA typing.

Since WMDA is doing these calculations on behalf of organisations, it is important for organisations to submit high resolution typing when available and donor ethnicity data. This greatly improves the haplofrequency sets generated and allows for individual donor match grades to be more accurate.

It is important to know that your organisations frequency set is changeable!

You can request that your donors be returned to the global consensus haplofrequency set or to a regional set. Your request should include the reasons why you believe that another set is more applicable for your donors. Additionally, if your organisation has a frequency set available that you would like to see utilized for the donors of your organisation we can implement it as well. In this case, your request should include at least a description about your population, the sample size, inclusion and exclusion criteria, calculation method, and the reasons why your haplofrequency set would be better than the current set applied by WMDA. Both requests will be reviewed by the WMDA Bioinformatics Working Group. Please send requests to support@wmda.info. In the future, you'll be able to dictate a frequency set on a per donor basis.

Date	Version	Description	Author
2017-10-31	1.0	Replacement BMDW global haplotype frequency set for more specific sets	JK
2018-01-24	1.1	Modification some sets; introduced during OptiMatch version 3.31.0	JK
2019-02-15	1.2	Replaced BMDW by WMDA / Search & Match Service; updated email address	JK