

## **Supporting Information**

### **CHHM: a Manually Curated Catalogue of Human Histone Modifications Revealing Hotspot Regions and Unique Distribution Patterns**

Wendong Ma<sup>1\*</sup>, Xiaofan Ding<sup>1\*</sup>, Jijia Xu<sup>1</sup> and Terence Chuen Wai Poon<sup>1#</sup>

<sup>1</sup> Institute of Translational Medicine, Centre for Precision Medicine Research and Training, MoE Frontiers Science Center for Precision Oncology, Faculty of Health Sciences, University of Macau, Macau, 999078, China

\*These authors contributed equally to this work.

**#Corresponding author:**

Terence Chuen Wai Poon

Associate Professor

Faculty of Health Sciences, University of Macau, Macau, China

Office Tel: +853 8822-4501

Email: [tcwpoon@um.edu.mo](mailto:tcwpoon@um.edu.mo)

## Table of Contents

**Table S1.** Summary of identified human histone variants. ----- Page S3

**Table S2.** The list of representative human histone variants for hotspot analysis. -- Page S7

**Figure S1.** The distribution patterns of various sites of modifications among human histone families H1, H2A, H2B, H3, and H4, at confidence levels of CL2 to CL3. ----- Page S9

**Table S1.** Summary of identified human histone variants. Human histone variants include 11 variants of human H1, 21 variants of human H2A, 21 variants of human H2B, 9 variants of human H3, and 2 variants of human H4. Protein information is retrieved from UniProt, accessed in Dec 2022.

<b>UniProt Primary Accession</b>	<b>UniProt Accession Suffix</b>	<b>Protein Name</b>	<b>Gene Name (Synonyms Name)</b>	<b>Chromosomal Location</b>	<b>Family</b>
<b>P07305</b>	H10_HUMAN	Histone H1.0	<i>H1-0 (H1F0, H1FV)</i>	22q13.1	H1
<b>Q02539</b>	H11_HUMAN	Histone H1.1	<i>H1-1 (H1F1, HIST1H1A)</i>	6p22.2	H1
<b>P16403</b>	H12_HUMAN	Histone H1.2	<i>H1-2 (H1F2, HIST1H1C)</i>	6p22.2	H1
<b>P16402</b>	H13_HUMAN	Histone H1.3	<i>H1-3 (H1F3, HIST1H1D)</i>	6p22.2	H1
<b>P10412</b>	H14_HUMAN	Histone H1.4	<i>H1-4 (H1F4, HIST1H1E)</i>	6p22.2	H1
<b>P16401</b>	H15_HUMAN	Histone H1.5	<i>H1-5 (H1F5, HIST1H1B)</i>	6p22.1	H1
<b>P22492</b>	H1T_HUMAN	Histone H1t	<i>H1-6 (H1FT, H1T, HIST1H1T)</i>	6p22.2	H1
<b>Q75WM6</b>	H1FNT_HUMAN	Testis-specific H1 histone	<i>H1-7 (H1FNT, HANP1)</i>	12q13.11	H1
<b>Q8IZA3</b>	H18_HUMAN	Histone H1.8	<i>H1-8 (H1FOO, H1OO, OSH1)</i>	3q22.1	H1
<b>P60008</b>	HILS1_HUMAN	Putative spermatid-specific linker histone H1-like protein	<i>H1-9P (H1-9, HILS1)</i>	17q21.33	H1
<b>Q92522</b>	H1X_HUMAN	Histone H1.10	<i>H1-10 (H1FX)</i>	3q21.3	H1
<b>P0C0S8</b>	H2A1_HUMAN	Histone H2A type 1	<i>H2AC11 (H2AFP, HIST1H2AG); H2AC13 (H2AFC, HIST1H2AI); H2AC15 (H2AFD, HIST1H2AK); H2AC16 (H2AFI, HIST1H2AL); H2AC17 (H2AFN, HIST1H2AM)</i>	6p22.1	H2A
<b>Q96QV6</b>	H2A1A_HUMAN	Histone H2A type 1-A	<i>H2AC1 (H2AFR, HIST1H2AA)</i>	6p22.2	H2A
<b>P04908</b>	H2A1B_HUMAN	Histone H2A type 1-B/E	<i>H2AC4 (H2AFM, HIST1H2AB); H2AC8 (H2AFA, HIST1H2AE)</i>	6p22.2	H2A
<b>Q93077</b>	H2A1C_HUMAN	Histone H2A type 1-C	<i>H2AC6 (H2AFL, HIST1H2AC)</i>	6p22.2	H2A
<b>P20671</b>	H2A1D_HUMAN	Histone H2A type 1-D	<i>H2AC7 (H2AFG, HIST1H2AD)</i>	6p22.2	H2A

<b>Q96KK5</b>	H2A1H_HUMAN	Histone H2A type 1-H	<i>H2AC12 (HIST1H2AH, HIST1H2AI)</i>	6p22.1	H2A
<b>Q99878</b>	H2A1J_HUMAN	Histone H2A type 1-J	<i>H2AC14 (H2AFE, HIST1H2AJ)</i>	6p22.1	H2A
<b>Q6FI13</b>	H2A2A_HUMAN	Histone H2A type 2-A	<i>H2AC18 (H2AFO, HIST2H2AA, HIST2H2AA3); H2AC19 (HIST2H2AA4)</i>	1q21.2	H2A
<b>Q8IUE6</b>	H2A2B_HUMAN	Histone H2A type 2-B	<i>H2AC21 (HIST2H2AB)</i>	1q21.2	H2A
<b>Q16777</b>	H2A2C_HUMAN	Histone H2A type 2-C	<i>H2AC20 (H2AFQ, HIST2H2AC)</i>	1q21.2	H2A
<b>Q7L7L0</b>	H2A3_HUMAN	Histone H2A type 3	<i>H2AC25 (H2AW, HIST3H2A)</i>	1q42.13	H2A
<b>P0C5Y9</b>	H2AB1_HUMAN	Histone H2A-Bbd type 1	<i>H2AB1 (H2AFB1)</i>	Xq28	H2A
<b>P0C5Z0</b>	H2AB2_HUMAN	Histone H2A-Bbd type 2/3	<i>H2AB2 (H2AFB2); H2AB3 (H2ABBD, H2AFB, H2AFB3)</i>	Xq28	H2A
<b>Q9BTM1</b>	H2AJ_HUMAN	Histone H2A.J	<i>H2AJ (H2AFJ)</i>	12p12.3	H2A
<b>A0A3B3IU63</b>	H2AL3_HUMAN	Histone H2A-like 3	<i>H2AL3 (H2AL1RP)</i>	Xp11.4	H2A
<b>Q71UI9</b>	H2AV_HUMAN	Histone H2A.V	<i>H2AZ2 (H2AFV, H2AV)</i>	7p13	H2A
<b>P16104</b>	H2AX_HUMAN	Histone H2AX	<i>H2AX (H2AFX)</i>	11q23.3	H2A
<b>O75367-2</b>	H2AY_HUMAN	Core histone macro-H2A.1	<i>MACROH2A1 (H2AFY)</i>	5q31.1	H2A
<b>Q9P0M6</b>	H2AW_HUMAN	Core histone macro-H2A.2*	<i>MACROH2A2 (H2AFY2)</i>	10q22.1	H2A
<b>P0C0S5</b>	H2AZ_HUMAN	Histone H2A.Z	<i>H2AZ1 (H2AFZ, H2AZ)</i>	4q23	H2A
<b>O75409</b>	HYPM_HUMAN	Huntingtin-interacting protein M*	<i>H2AP (CXorf27, HYPM)</i>	Xp11.4	H2A
<b>Q96A08</b>	H2B1A_HUMAN	Histone H2B type 1-A	<i>H2BC1 (HIST1H2BA, TSH2B)</i>	6p22.2	H2B
<b>P33778</b>	H2B1B_HUMAN	Histone H2B type 1-B	<i>H2BC3 (H2BFF, HIST1H2BB)</i>	6p22.2	H2B
<b>P62807</b>	H2B1C_HUMAN	Histone H2B type 1-C/E/F/G/I	<i>H2BC4 (H2BFL, HIST1H2BC); H2BC6 (H2BFH, HIST1H2BE); H2BC7 (H2BFG, HIST1H2BF); H2BC8 (H2BFA, HIST1H2BG); H2BC10 (H2BFK, HIST1H2BI)</i>	6p22.2	H2B
<b>P58876</b>	H2B1D_HUMAN	Histone H2B type 1-D	<i>H2BC5 (H2BFB, HIRIP2, HIST1H2BD)</i>	6p22.2	H2B
<b>Q93079</b>	H2B1H_HUMAN	Histone H2B type 1-H	<i>H2BC9 (H2BFJ, HIST1H2BH)</i>	6p22.2	H2B
<b>P06899</b>	H2B1J_HUMAN	Histone H2B type 1-J	<i>H2BC11 (H2BFR, HIST1H2BJ)</i>	6p22.1	H2B

<b>Q60814</b>	H2B1K_HUMAN	Histone H2B type 1-K	<i>H2BC12 (H2BFT, HIRIP1, HIST1H2BK)</i>	6p22.1	H2B
<b>Q99880</b>	H2B1L_HUMAN	Histone H2B type 1-L	<i>H2BC13 (H2BFC, HIST1H2BL)</i>	6p22.1	H2B
<b>Q99879</b>	H2B1M_HUMAN	Histone H2B type 1-M	<i>H2BC14 (H2BFE, HIST1H2BM)</i>	6p22.1	H2B
<b>Q99877</b>	H2B1N_HUMAN	Histone H2B type 1-N	<i>H2BC15 (H2BFD, HIST1H2BN)</i>	6p22.1	H2B
<b>P23527</b>	H2B1O_HUMAN	Histone H2B type 1-O	<i>H2BC17 (H2BFH, H2BFN, HIST1H2BO)</i>	6p22.1	H2B
<b>Q6DN03</b>	H2B2C_HUMAN	Putative histone H2B type 2-C	<i>H2BC20P (HIST2H2BC)</i>	1q21.2	H2B
<b>Q6DRA6</b>	H2B2D_HUMAN	Putative histone H2B type 2-D	<i>H2BC19P (HIST2H2BD)</i>	1q21.2	H2B
<b>Q16778</b>	H2B2E_HUMAN	Histone H2B type 2-E	<i>H2BC21 (H2BFQ, HIST2H2BE)</i>	1q21.2	H2B
<b>Q5QNW6</b>	H2B2F_HUMAN	Histone H2B type 2-F	<i>H2BC18 (HIST2H2BF)</i>	1q21.2	H2B
<b>Q8N257</b>	H2B3B_HUMAN	Histone H2B type 3-B	<i>H2BC26 (H2BU1, HIST3H2BB)</i>	1q42.13	H2B
<b>P0C1H6</b>	H2BFM_HUMAN	Histone H2B type F-M	<i>H2BW2 (H2BFM)</i>	Xq22.2	H2B
<b>P57053</b>	H2BFS_HUMAN	Histone H2B type F-S	<i>H2BC12L (H2BFS, H2BS1)</i>	21q22.3	H2B
<b>A0A2R8Y619</b>	H2BK1_HUMAN	Histone H2B type 2-K1	<i>H2BK1 (H2BE1)</i>	7q36.1	H2B
<b>P0DW85</b>	H2BN1_HUMAN	Histone H2A.N	<i>H2BN1</i>	17q11.2	H2B
<b>Q7Z2G1</b>	H2BWT_HUMAN	Histone H2B type W-T	<i>H2BW1 (H2BFWT)</i>	Xq22.2	H2B
<b>P68431</b>	H31_HUMAN	Histone H3.1	<i>H3C1 (H3FA, HIST1H3A); H3C2 (H3FL, HIST1H3B); H3C3 (H3FC, HIST1H3C); H3C4 (H3FB, HIST1H3D); H3C6 (H3FD, HIST1H3E); H3C7 (H3FI, HIST1H3F); H3C8 (H3FH, HIST1H3G); H3C10 (H3FK, HIST1H3H); H3C11 (H3FF, HIST1H3I); H3C12 (H3FJ, HIST1H3J)</i>	6p22.2, 6p22.1 (H3C10, H3C11, H3C12)	H3
<b>Q71DI3</b>	H32_HUMAN	Histone H3.2	<i>H3C13 (HIST2H3D); H3C14 (H3F2, H3FM, HIST2H3C); H3C15 (HIST2H3A)</i>	1q21.2	H3
<b>P84243</b>	H33_HUMAN	Histone H3.3	<i>H3-3A (H3.3A, H3F3, H3F3A); H3-3B (H3.3B, H3F3B)</i>	1q42.12 (H3-3A), 17q25.1 (H3-3B)	H3
<b>Q16695</b>	H31T_HUMAN	Histone H3.1t	<i>H3-4 (H3FT, HIST3H3)</i>	1q42.13	H3

<b>Q6NXT2</b>	H3C_HUMAN	Histone H3.3C	<i>H3-5 (H3F3C)</i>	12p11.21	H3
<b>Q5TEC6</b>	H37_HUMAN	Histone H3-7	<i>H3-7 (H3-2, HIST2H3PS2)</i>	1q21.1	H3
<b>P0DPK2</b>	H3Y1_HUMAN	Histone H3.Y	<i>H3Y1</i>	5p15.1	H3
<b>P0DPK5</b>	H3Y2_HUMAN	Histone H3.X	<i>H3Y2</i>	5p15.1	H3
<b>P49450</b>	CENPA_HUMAN	Histone H3-like centromeric protein A	<i>CENPA</i>	2p23.3	H3
<b>P62805</b>	H4_HUMAN	Histone H4	<i>H4C1 (H4/A, H4FA, HIST1H4A); H4C2 (H4/I, H4FI, HIST1H4B); H4C3 (H4/G, H4FG, HIST1H4C); H4C4 (H4/B, H4FB, HIST1H4D); H4C5 (H4/J, H4FJ, HIST1H4E); H4C6 (H4/C, H4FC, HIST1H4F); H4C8 (H4/H, H4FH, HIST1H4H); H4C9 (H4/M, H4FM, HIST1H4I); H4C11 (H4/E, H4FE, HIST1H4J); H4C12 (H4/D, H4FD, HIST1H4K); H4C13 (H4/K, H4FK, HIST1H4L); H4C14 (H4/N, H4F2, H4FN, HIST2H4, HIST2H4A); H4C15 (H4/O, H4FO, HIST2H4B); H4C16 (H4-16, HIST4H4)</i>	6p22.2, 6p22.1 (H4C9, H4C11, H4C12, H4C13), 1q21.2 (H4C14, H4C15), 12p12.3 (H4C16)	H4
<b>Q99525</b>	H4G_HUMAN	Histone H4-like protein type G	<i>H4C7 (H4/L, H4FL, HIST1H4G)</i>	6p22.2	H4

\* Q9P0M6|H2AW\_HUMAN and O75409|HYPM\_HUMAN are not counted into H2A family in UniProt.

**Table S2.** The list of representative human histone variants for hotspot analysis. The representative variants refer to variants with high homology, consisting of 5 variants in H1, 12 variants in H2A, 15 variants in H2B, 7 variants in H3, and 2 (all) variants in H4.

<b>UniProt Primary Accession</b>	<b>Protein Name</b>	<b>Family</b>	<b>Number of Amino Acids</b>
Q02539	Histone H1.1	H1	215
P16403	Histone H1.2	H1	213
P16402	Histone H1.3	H1	221
P10412	Histone H1.4	H1	219
P16401	Histone H1.5	H1	226
P0C0S8	Histone H2A type 1	H2A	130
Q96QV6	Histone H2A type 1-A	H2A	131
P04908	Histone H2A type 1-B/E	H2A	130
Q93077	Histone H2A type 1-C	H2A	130
P20671	Histone H2A type 1-D	H2A	130
Q96KK5	Histone H2A type 1-H	H2A	128
Q99878	Histone H2A type 1-J	H2A	128
Q6F113	Histone H2A type 2-A	H2A	130
Q8IUE6	Histone H2A type 2-B	H2A	130
Q16777	Histone H2A type 2-C	H2A	129
Q7L7L0	Histone H2A type 3	H2A	130
Q9BTM1	Histone H2A.J	H2A	129
Q96A08	Histone H2B type 1-A	H2B	127
P33778	Histone H2B type 1-B	H2B	126
P62807	Histone H2B type 1-C/E/F/G/I	H2B	126
P58876	Histone H2B type 1-D	H2B	126
Q93079	Histone H2B type 1-H	H2B	126
P06899	Histone H2B type 1-J	H2B	126
O60814	Histone H2B type 1-K	H2B	126
Q99880	Histone H2B type 1-L	H2B	126
Q99879	Histone H2B type 1-M	H2B	126
Q99877	Histone H2B type 1-N	H2B	126
P23527	Histone H2B type 1-O	H2B	126
Q16778	Histone H2B type 2-E	H2B	126
Q5QNW6	Histone H2B type 2-F	H2B	126
Q8N257	Histone H2B type 3-B	H2B	126
P57053	Histone H2B type F-S	H2B	126
P68431	Histone H3.1	H3	136
Q71DI3	Histone H3.2	H3	136
P84243	Histone H3.3	H3	136
Q16695	Histone H3.1t	H3	136
Q6NXT2	Histone H3.3C	H3	135

Q5TEC6	Histone H3-7	H3	136
P0DPK2	Histone H3.Y	H3	136
P62805	Histone H4	H4	103
Q99525	Histone H4-like protein type G	H4	98



**Figure S1.** The prevalence of various amino acid residues with modifications in different human histone families H1, H2A, H2B, H3, and H4, at confidence levels of CL2 to CL3. K, R, S, T, Y, M, C, Q, E, N, P, and Nt refer to lysine, arginine, serine, threonine, tyrosine, methionine, cysteine, glutamine, glutamic acid, asparagine, proline, and protein *N*-terminus, respectively.

