Moi je ferais comme ça :

Première méthode

Il suffit de remarquer que T croissante et $T_{\#}\mu = \nu$ impliquent que T optimise le coût quadratique. Il y a unicité pour l'optimum (c'est un cas particulier du théorème de Brenier). Donc il y a unicité pour T. Pas besoin de passer par l'optimalité pour $h(x - y)$.

Deuxième méthode

Définissons $F(x) = \mu([a, x])$ et $G(y) = \nu([b, y])$, où a et b sont les bornes inférieures des intervalles (éventuellement $-\infty$). En particulier F et G sont continues à droite et $G_{-}(y) := \nu([b, y[)$ coïncide avec sa limite gauche en tout point.

Si T est croissant on a $T^{-1}([b, T(x)]) \supset [a, x]$. Donc $G(T(x)) \geq F(x)$. De plus, $T^{-1}([b, T(x)[) \subset [a, x[\subset [a, x]$, donc $G_{-}(T(x)) \leq F(x)$.

En particulier, pour tout x, le point $T(x)$ doit être tel que $F(x) \in [G_{-}(T(x)), G(T(x))]$. La question est : peut-il y avoir deux valeurs différentes $t_1 <_2$ telles que $F(x) \in [G_{-}(t_1), G(t_1)] \cap [G_{-}(t_2), G(t_2)]$? si c'est le cas on aurait $G_{-}(t_2) \leq G(t_1)$ mais alors G, qui est croissante serait constante (et égale à $G_{-}(t_2) = G(t_1) = F(x)$) sur tout $[t_1, t_2[$.

Donc les points d'ambiguité sont les images réciproques par F des valeurs prises par G sur des plateaux. Ces plateaux sont dénombrables et l'ensemble A des valeurs qu'on y prend l'est aussi.

Or, on a, pour toute valeur l, $\mu(F^{-1}(\{l\})) = 0$, puisque cette image réciproque est un intervalle sur lequel F est constante, ce qui signifie qu'il pourrait juste y avoir de la masse au début, mais μ n'a pas d'atomes ... Par conséquent, $\mu(F^{-1}(A)) = 0$, ce qui signifie que les points d'ambiguité sont négligeables.